Clinical Pragmatism Mapping of ICPC-2 PLUS to SNOMED CT

Dr Hywel Lloyd, Director of Informatics,

BPAC Clinical Solutions, Health Terminology New Zealand.
Honorary Senior Lecturer, Department of General Practice, University of Otago, Dunedin

Aim: To provide a clinically safe, single direction, one to one translation of ICPC-2 PLUS to SNOMED CT for the context of Obstetric Care.

Objective: To automatically map code 'descriptions' for ICPC-2 Plus codes to SNOMED CT. When no maps exist and the codes are required for the obstetric care system then undertake a manual clinical mapping process to ensure a clinically safe application.

Purpose

Best Practice Clinical Solutions (BPAC – CS) required the translation of ICPC-2 Plus clinical codes to SNOMED CT for our project with the Australian Department of Health on the iCareNet product. iCareNet is a cloud based EHR with the concept of Plans of Care being the foundation for making collaborative integrated care possible across the Northern Territory health system. The clinical context for the integration was in perinatal care in primary care. The requirement was to map, a one to one association of all ICPC-2 PLUS codes to SNOMED CT. We had a specific focus on diagnostic codes and procedures, with reason for encounter and symptom codes being less important.

The aim was to provide a clinically safe representation of past and current problems from PCIS and Communicare to the iCareNet platform and use the diagnostic codes to drive electronic decision support for clinicians using iCareNet.

Background

International classification of process in primary care (ICPC) was conceived by WONCA, AHA and WHO. It was an amalgamation of three classifications; the 'diagnostic codes' from International classification of Health Problem in Primary care (ICHPPC 1974), the 'Symptoms and Complaints' Reason for Encounter Classification (RFEC 1981) and the process of care classification IC-Process-PC. ICPC was published in 1987. ICPC was originally designed for paper based data collection and analysis for categorising data for health statistics, and may lack the specificity needed for clinical records. The structure of ICPC is illustrated in Table 1.

Table 1 ICPC Structure, with ICPC-2 Plus code counts by chapter and rubric.

ICPC 2 Bi Axial Structure																				
Components	Rubric	Chapters: 15 Somatic, 1 Psychological, 1 Social are all based on body systems																		
		A	B	D	E	K	L	M	N	P	R	S	T	U	W	X	Y	Z		
Symptoms and Complaints	01–29	161	23	205	71	48	63	177	130	295	134	197	73	61	92	150	99	244		
Diagnostic and Preventative Procedures	30–49	356	58	115	40	36	71	153	45	43	62	30	73	54	75	50	47	24		
Treatment Procedures, medication	50–59	87	27	71	39	39	63	160	40	32	50	72	38	57	47	57	35	15		
Test Results	60-61	6	2	2	2	2	2	2	2	2	4	2	2	4	3	2	2			
Administrative	62	44	1	1	1	3	1	2	1	1	2	1	1	2	2	2	2	6		
Referral and other reasons for encounter	63–69	82	9	14	7	7	12	10	10	17	8	7	9	8	13	9	9	21		
Diseases	70–99	427	130	355	147	85	320	499	178	139	177	257	171	132	133	176	80	0		

ICPC 2 Chapters							
A	General	H	Ear	P	Psychological	U	Urinary
B	Blood, blood forming	K	Circulatory	R	Respiratory	W	Pregnancy
D	Digestive	L	Musculoskeletal	S	Skin	X	Female genital
F	Eye	N	Neurological	T	End/met/nut	Y	Male genital
						Z	Social

ICPC-2 Plus is intentionally primary care focussed and structured, which raises some challenges when translating to a more disease focused ontology.

The Family Medicine Research Centre (FMRC), at the University of Sydney, recognised the need for an acceptable and usable terminology for general practice computerised clinical systems. The FMRC designed an Australian general practice terminology, classified to ICPC, based on terms recorded in over one million general practice encounter records. The terminology was released in 1995 and called ICPC PLUS.

ICPC-2 PLUS the second edition, which is the BEACH coding system, allows health professionals to record symptoms, diagnosis (problems and labels) past health issues and processes (procedures, counselling and referrals). ICPC-2 PLUS is supported and regularly updated by FMRC approximately three times per year to incorporate user feedback and suggestions for new content.

Wang and Patric *et al*, in 2008, undertook a computational translation of ICPC-2 PLUS to SNOMED CT, see Table 2. The mappings, as far as we could ascertain, are not available for implementation. They are computer generated matches that have not been clinically validated.

Table 2 Previous mapping of ICPC-2 Plus by Wang & Patrick 2008

	A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT				
	Mapping Steps				
	Total Codes	Unified medical Language System metathesaurus mapping	Lexical Mapping (WordNet)	Post-coordination	Mapped Codes
FMRC, University of Sydney	7,410	53.0%	60.31%	80.58%	5,971

BMC Medical Informatics and Decision Making 2008, 8(Suppl 1):S5 Yefeng Wang, Jon Patrick, Graeme Miller, Julie O'Halloran
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2582792/

Method

We decided to combine computer generated direct matching of terms with a clinical code mapping process (CCMP). We built an application for clinical users to select appropriate SNOMED CT conceptIDs from matched ICPC- 2 PLUS terms. All resultant matches and selections were independently checked by three clinical informaticians within the BPAC-CS team. We uploaded the Australian AU1000036_20150531 SNOMED CT release and ICPC-2 PLUS code set from the FRMC web site to SQL Server 2012. The direct mapping process involved matching ICPC-2 PLUS Term Description or the nalan50 terms to match SNOMED CT Description fields. We undertook a number of techniques to improve the matching.

The automated matching results were then verified by our clinical informatics team. The remaining unmatched codes were then matched using the code CCMP. We utilised the free text search capability within SQL Server 2012. Potential matches were ranked using the FREETEXTTABE RANK functionality. The RANKing is based on the OKAPI BM25 ranking formula. We adjusted the RANK score as longer length SNOMED CT descriptions ranked too high in the pure FREETEXTTABKLE RANK score.

The Application presented to the clinician a list of ICPC-2 PLUS codes to be mapped and for each ICPC-2 PLUS code to a list of returned SNOMED CT codes in RANK and Strike Rate score order. We took a pragmatic approach to prioritise an initial ICPC-2 PLUS maternity reference set to be included within the iCareNet system. The decision was made to include pre-coordinated codes only, and not to undertake post-coordinated mapping.

Results

Through a dual process of automatic direct matching of terms or SNOMED CT synonyms, and manually selected codes, by our clinical informaticians, we mapped 5625 ICPC-2 PLUS codes to SNOMED CT, 68.1% of the 8258 total codes, Table 3. The BPAC CS informaticians, in collaboration with the clinical team fromNorthern Territory, identified 4477 ICPC-2 Plus codes to be consumed by the iCareNet system. We mapped 4391 (98.1%) of these codes. We decided not to undertake post-co-ordinated mapping. Thirty one ICPC-2 PLUS codes were mapped to two SNOMED CT conceptIDs.

Automatic matching with the SNOMED CT descriptions and synonyms directly to the ICPC-2 PLUS terms achieved a 37.1% coverage. This was significantly less than Wang and Patrick (60.31%) using the WordNet lexical matching process. The automatic matching process did not deliver accurate mappings on review by the clinical team. Of the 2618 matches, 635 (24.3%) were manually re-mapped by the informaticians after clinical review of automatic mapping. Ten diagnostic codes were not mapped to SNOMED CT. Nine of these have since been mapped to a higher hierarchical level than desired SNOMED CT code, see Table 4.

Table 3 The number of ICPC-2 Plus codes mapped comparing the automatic codes with the clinician selected codes.

		ICPC-2 PLUS			ICPC-2PLUS iCareNet Subset			Coverage	
Components	Rubric	Total Codes	Auto Direct Match	CCSP	Total Codes	Auto Direct Match	CCSP	ICPC2 PLUS	iCareNet Codes
Symptoms and Complaints	01–29	2223	690	798	931	264	643	66.9%	97.4%
Diagnostic and Preventative Procedures	30–49	1332	106	395	419	58	340	37.6%	95.0%
Treatment Procedures, medication	50–59	929	181	273	372	100	251	48.9%	94.4%
Test Results	60–61	43	0	9	7	0	7	20.9%	100.0%
Administrative	62	73	2	4	5	1	4	8.2%	100.0%
Referral and other reasons for encounter	63–69	252	7	112	119	4	105	47.2%	91.6%
Diseases	70–99	3406	1632	1416	2624	1295	1319	89.5%	99.6%
Totals		8258	2618	3007	4477	1722	2669	68.1%	98.1%

Table 4 Disease codes not initially mapped to SNOMED CT.

ICPCCode	ICPC Code	CHAPTER	RUBRIC	Term Code	Term description
B99018	B99	B	99	18	Disease; blood forming organs
D72J94	D72	D	72	J94	HepB:Infected:NOT on Treatment
D72J95	D72	D	72	J95	HepB:Infected: ON Treatment
D72J96	D72	D	72	J96	HepB:Immune by Exposure
K71J96	K71	K	71	J96	Rheumatic heart; Prophylaxis
K71J97	K71	K	71	J97	Rheumatic Heart (P1); Severe
K71J98	K71	K	71	J98	Rheumatic Heart (P2); Moderate
K71J99	K71	K	71	J99	Rheumatic Heart (P3); Mild
P71010	P71	P	71	10	Disorder;mental;post brain dam
P79017	P79	P	79	17	Homophobia

Discussion

The methods for mapping ICPC-2 PLUS codes to SNOMED CT we undertook has highlighted the need for clinical validation of machine matched codes, with 24% of codes required to be remapped. This may explain why previous work in this area does not appear to have been released for clinical use. The mapping of high temperature to a physical force rather than to a finding demonstrates the inaccuracies inherent in an automated approach to code translations. Mismatches also occur when duplicate synonyms exist at different levels of the SNOMED CT hierarchy in the examples in Table 5.

There is a need for the SNOMED CT synonyms to be expanded to incorporate a larger number of synonyms. A consistent standardised approach to how synonyms are constructed within similar hierarchies, see Figure 1, would be helpful. Improved synonym structure and more synonyms including the WordNet lexicon would improve future mapping projects and also improve text searches within health information systems. More synonyms would reduce the requirements for system suppliers to introduce additional key word searching, external to the SNOMED CT ontology.

We were unable to map a number of ICPC-2 Plus codes to SNOMED CT as they were particular to Australia. These codes would require a specific code addition to the Australian reference set.

Table 5. Examples of duplicate synonyms exist at diffrenet levels of the SNOMED CT hierachy.

Row	conceptID	Description	FSN	Code	Description
	285717004	High temperature	High temperature (physical force)	A03006	Temperature;high
	164303007	O/E – temperature elevated	On examination – temperature elevated (finding)	A03006	Temperature;high
93571013	56267009	Vascular dementia	Multi-infarct dementia (disorder)	P70011	Dementia vascular
2770951017	429998004	Vascular dementia	Vascular dementia (disorder)	P70011	Dementia vascular

56267009 multi-infarct dementia has a synonym of vascular dementia. This is further down the hierarchy compared to 429998004 vascular dementia as ICPC does not code any other type of vascular dementia so pick the code higher up the tree. Map to 429998004.

Figure 1

Conclusion

The work we have undertaken has achieved a clinically validated ICPC-2 PLUS map to SNOMED CT for the implementation of an antenatal record system for the NT government. Additional work is required to expand the Australian reference set to accomplish a full translation of ICPC-2 Plus to SNOMED CT.