

Aligning an administrative procedure coding system with SNOMED CT

Stefan SCHULZ^{a,c} Johannes STEFFEL^a, Peter POLSTER^a,
Matvey PALCHUK^b, Philipp DAUMKE^a

^a *Averbis GmbH, Freiburg, Germany*

^b *TriNetX, Cambridge, MA, USA*

^c *Institute for Medical Informatics, Statistics and
Documentation, Medical University of Graz, Austria*

Context

- Secondary use of clinical data – major promise for a variety of use cases (observational research, decision support, prediction, QA)
- Most clinical data are unstructured
- Structured data mainly used for administrative purposes
- Situation in German Hospitals:
DRG (Diagnosis-related group) system calculates rates for in-patient care according to
 - Main diagnoses
 - Major procedures (inexpensive ones not mapped)
 - Additional features, e.g. ventilation time
- ICD-10-GM and OPS provided by

**ICD-10 GM – extension
of WHO classification**

**OPS – national
procedure coding system
– only in German**

Challenge

- How to reach international interoperability of clinical procedures?
- Mapping OPS codes to the international standard **SNOMED CT**
- Questions:
 - How compatible are the two terminologies?
 - How good is the mapping?
 - What effort is needed?

OPS classification

35,641 codes in a single hierarchy (7 levels)

SNOMED CT procedure subhierarchy

81,462 codes in multiple hierarchies

- **So what ?**

Ontological background

	Procedure classification	Procedure ontology
Semantics	Closed world (disjoint classes)	Open world, classes (extension of concepts) often overlap
Structure	Single hierarchy, extensional	Multiple hierarchy, intensional
Constructors	Subclass, exclusion rules	Subclass, Equivalence, conjunction, existential quantification
Ontological commitment	Nowhere really defined – most likely purpose-oriented standardised information objects related to medical procedures	Classes of medical procedures, i.e. (parts of) actions performed on a patient by a health professional
Purpose	Provision of correlates to medical procedures and their parts, from which the monetary value of a procedure can be derived	Provision of representational units within a formal account of the electronic health record

Characteristics OPS / SNOMED CT

	OPS classification	SNOMED CT procedure subhierarchy
Size	35,641 codes distributed across seven hierarchical levels	81,462 codes distributed across seven hierarchical levels

Characteristics OPS / SNOMED CT

	OPS classification	SNOMED CT procedure subhierarchy
Size	35,641 codes distributed across seven hierarchical levels	81,462 codes distributed across seven hierarchical levels
Example 1	Video-EEG (10/20 Electrodes). Excl.: Video-EEG during pre and intraoperative epilepsy assessment	Video electroencephalogram
Example 2	Right heart catheterization: Other	Catheterization of right heart
Example 3	Individual psychotherapy delivered by physician – three sessions per week	Individual psychotherapy (regime/therapy)

8-98 Sonstige multimodale Komplexbehandlung

8-98b Andere neurologische Komplexbehandlung des akuten Schlaganfalls

Exkl.: Neurologische Komplexbehandlung des akuten Schlaganfalls ([8-981](#) ff.)

Info: Diese Codes können auch beim Vorliegen einer TIA angegeben werden

Besteht über die Therapiemöglichkeiten der vorhandenen Schlaganfalleinheit hinaus die Indikation zu einer Behandlung auf der Intensivstation, kann, wenn die Mindestmerkmale dieses OPS-Kodes erfüllt sind, die dortige Behandlungszeit auch für die Kodierung der neurologischen Komplexbehandlung des akuten Schlaganfalls berücksichtigt werden, auch wenn auf der Intensivstation nicht ausschließlich Patienten mit einem akuten Schlaganfall behandelt werden

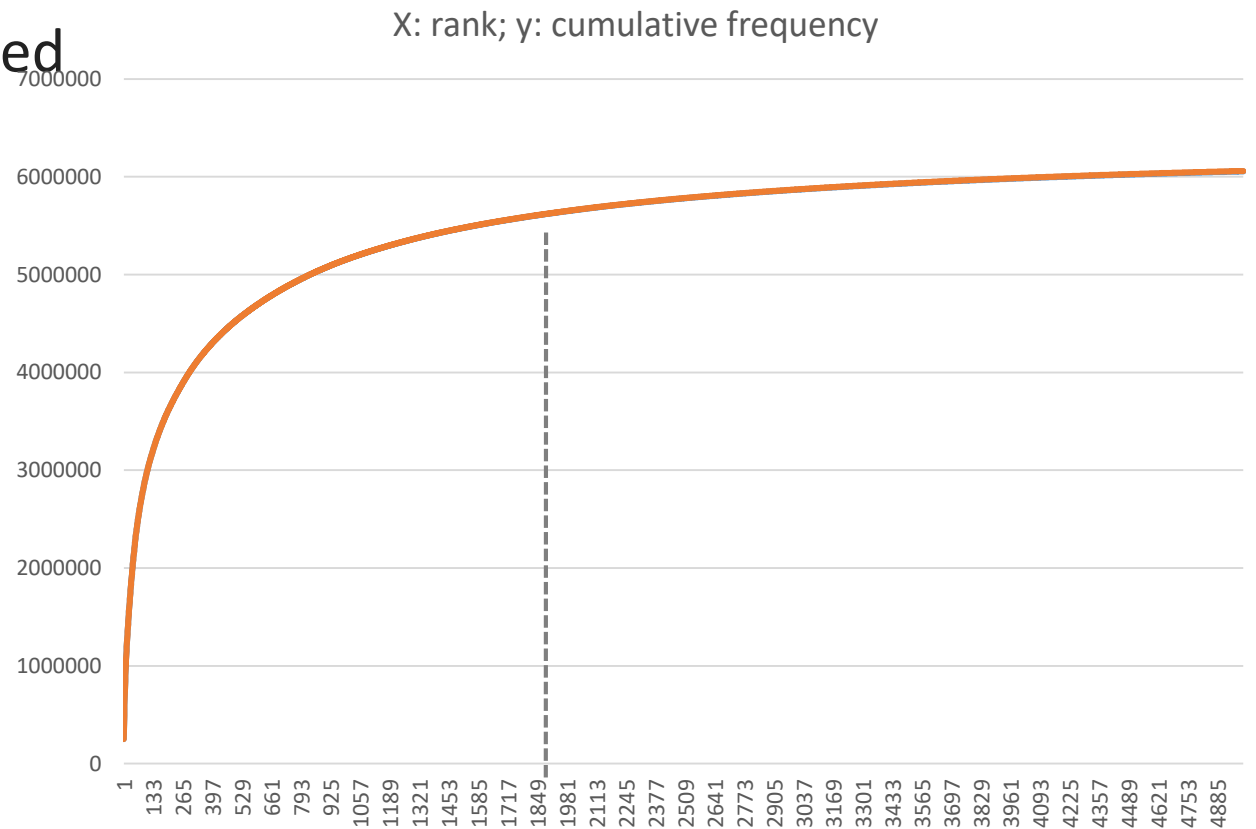
Mindestmerkmale: Behandlung auf einer spezialisierten Einheit durch ein multidisziplinäres, auf die Schlaganfallbehandlung spezialisiertes Team unter fachlicher Behandlungsleitung durch einen Facharzt für Neurologie oder einen Facharzt für Innere Medizin (in diesem Fall muss im Team der neurologische Sachverstand kontinuierlich eingebunden sein) mit: 24-stündiger ärztlicher Anwesenheit (auch als Bereitschaftsdienst)

Mindestmerkmale: Behandlung auf einer spezialisierten Einheit durch ein multidisziplinäres, auf die Schlaganfallbehandlung spezialisiertes Team unter fachlicher Behandlungsleitung durch einen Facharzt für Neurologie oder einen Facharzt für Innere Medizin (in diesem Fall muss im Team der neurologische Sachverstand kontinuierlich eingebunden sein) mit: 24-Stunden-Monitoring von mindestens 6 der folgenden Parameter: Blutdruck, Herzfrequenz, EKG, Atmung, Sauerstoffsättigung, Temperatur, intrakranieller Druck, EEG, evozierte Potentiale. Blutdruck, Temperatur und evozierte Potentiale können auch nichtautomatisiert bestimmt werden. Das Monitoring darf nur zur Durchführung spezieller Untersuchungen oder Behandlungen unterbrochen werden. Alle Parameter müssen im Abstand von 4 Stunden oder häufiger erhoben und dokumentiert werden

Mindestmerkmale: Behandlung auf einer spezialisierten Einheit durch ein multidisziplinäres, auf die Schlaganfallbehandlung spezialisiertes Team unter fachlicher Behandlungsleitung durch einen Facharzt für Neurologie oder einen Facharzt für Innere Medizin (in diesem Fall muss im Team der neurologische Sachverstand kontinuierlich eingebunden sein) mit: 6-stündlicher (maximaler Abstand nachts 8 Stunden) Überwachung und Dokumentation des neurologischen Befundes durch einen Arzt zur Früherkennung von Schlaganfallprogression, -rezidiv und anderen Komplikationen

Resources

- Approx. 6.2M token encodings from German university hospitals
- 20,501 out of 35,641 codes used; 3,288 only once!
 - 42.5% of code types – never used
 - Only 48.3 of code types – used more than once
- 90% percentile: 2125 codes
- Mapping (sponsored by TriNetX):
 - Phase 1: mapping #1 - #1000
 - Phase 2: mapping #1001 - #2125



Organisation of mapping

- Main author together with two MD students: 8 hrs / week over 3 months
- Step-wise approach (codes #1 - #1000 randomly ordered):
 - #1 - #1000 Collaborative exploratory mapping, formulation of mapping guidelines, e.g. criteria of exclusion, timeout at complicated mappings
 - #101 - #300 half of codes mapped by one student, revised by the other
 - #301 - #400 independent mapping, IAA testing, adjudication
 - #401 - #800 half of codes mapped by one student, revised by the other, discussion of controversial cases, adjustment of mapping guidelines
 - #801 - #1000 independent mapping, IAA testing. Final decisions about “difficult” codes. Thorough revision by first author.

Mapping guidelines

- Logical combinations allowed (between up to four SNOMED CT concepts):
 - OR – logical disjunction between two SNOMED CT concepts
 - AND – logical conjunction between two SNOMED CT concepts
 - ADD – addition of two SNOMED CT concepts
- No further SNOMED CT post-co-ordination allowed

- Scoring

Meaning regarding source OPS code (S) and target code or expression (T)

Exact	T holds for the same (individual) procedures as S
Exact-Q	T holds for the same (individual) procedures as S, when quantitative restrictions on S are neglected
Broader	The individual procedures denoted by S are a (still significant) subset of those denoted by T
No mapping	There is no code or expression T that allows any of the above judgement.

Results

Cardinality of map (SNOMED CT codes per OPS codes)

0	1	2	3	4
48	617	282	42	11

Quality of mapping

Broader	Exact	Exact-Q	No mapping	Revisit
610	310	32	48	0

Type of logical combination

None	AND	ADD	OR	Complex
665	178	79	56	22

Balance of effort: 234 hours (1 senior terminologist, two medical students)

Inter-coder agreement

Type of agreement	Agreement [95% CI]	
	OPS ₃₀₁ - OPS ₄₀₀	OPS ₈₀₁ - OPS ₁₀₀₀
Coders agree on at least one core SNOMED CT concept per OPS code	68% [58%; 76%]	65% [58%; 71%]
Coders agree on the same set of SNOMED CT concepts per OPS code	54% [44%; 63%]	46% [38%; 52%]
Coders agree on the same set of SNOMED CT concepts per OPS code and agree regarding the mapping score	41% [31%; 50%]	36% [30%; 43%]

Typical cases

1 Procedure with finding: finding is not represented in the map (would require complex post-coordination)

1-265.4 Electrophysiological examination of the heart, catheter-assisted: In tachycardia with narrow QRS complex or atrial tachycardia

175131000 Percutaneous transluminal electrophysiological studies on conducting system of heart

2 Procedure with device: device is not represented in the map (would require complex post-coordination)

1-266.1 Electrophysiological examination of the heart, not catheter-assisted: implanted cardioverter defibrillator (ICD)

252425004 Cardiac electrophysiology

Typical cases

3 Procedure with body part: body part is not represented in the map (would require complex post-coordination)

1-268.3 Cardiac Mapping: Right Ventricle

21032000 Cardiac mapping

4 Part of the procedure requires separate coding

1-430.1 Endoscopic biopsy of respiratory organs: bronchus
312849006

312849006 Biopsy of
bronchus

10847001
Bronchoscopy

ADD

Typical cases

5 Logical conjunction of specific procedure and anatomy-related procedure

<i>1-490.6 Biopsy without incision on skin and subcutaneous tissue: lower leg</i>	287538006 Non-surgical skin biopsy	118714000 Procedure on lower leg	AND
---	------------------------------------	----------------------------------	-----

6 Different granularity in SNOMED requires post-coordination in one case but not in another

<i>3-825 Magnetic resonance imaging of the abdomen with contrast</i>		432369004 Magnetic resonance imaging of abdomen with contrast	
--	--	---	--

<i>3-826 Magnetic resonance imaging of the musculoskeletal system with contrast agent</i>	58713006 Magnetic resonance imaging of musculoskeletal structures	51619007 Magnetic resonance imaging with contrast	AND
---	---	---	-----

Typical cases

7 Coordination needed to add the feature that a procedure is a diagnostic one:

<i>1-631 Diagnostic Esophagogastroscopy</i>	392153002 Esophagogastroscopy	103693007 Diagnostic procedure	AND
---	----------------------------------	--------------------------------------	-----

8 Missing of aggregations at the level “vessel” (regardless of whether artery or vein)

<i>3-611.x Phlebography of cervical and thoracic vessels: Other</i>	4008007 Phlebography of neck	60006002 Intrathoracic phlebography	OR
---	---------------------------------	---	----

Typical cases

9 Exclusion statements for OPS codes (cannot be expressed by SNOMED CT semantics)

*1-207.2 Video-EEG (10/20
Electrodes). Excl.: Video-EEG
during pre and intraoperative
epilepsy assessment*

252738008 Video electroencephalogram

10 Residual class “other”, i.e. logical complement (cannot be expressed by SNOMED CT semantics)

*1-273.x Right heart
catheterization: Other*

40403005 Catheterization of right heart

Typical cases

11 Explicit definition “without” (cannot be expressed by SNOMED CT semantics)

<i>1-275.0 Transarterial Left Heart Catheter Examination: Coronary angiography without further action</i>	33367005	Coronary angiography	67629009	Catheterization of left heart	AND
---	----------	----------------------	----------	-------------------------------	-----

12 Distinction between logical conjunction “AND” and addition (more than one instance in the target representation (“ADD”))

<i>1-650.2 Diagnostic Colonoscopy: Total, with ileoscopy</i>	174184006	Diagnostic endo-scopic examination on colon	235150006	Total colonoscopy	265387003	Diagnostic endo-scopic examination of ileum	(X AND Y) ADD Z
--	-----------	---	-----------	-------------------	-----------	---	--------------------

Conclusion

- 95% of top 1000 OPS codes could be mapped to SNOMED CT; only 1/3 exactly
- OPS, and probably other procedure classifications for billing often include criteria like age, number or duration of treatments, indication etc., which are not considered in the SNOMED hierarchy → broader mappings
- Complex rules and exclusions preclude exact maps
- Different degrees of pre-coordination requires simple SNOMED concept post-coordination by logical conjunction
- Low IAA known issue with SNOMED CT and other large terminologies*
- More fine-grained anatomic dissections in SNOMED CT require disjunctive target expressions
- Full post-coordination would achieve close to 100% coverage, but probably no significant increase in exact mappings

*Miñarro-Giménez JA, Cornet R, Jaulent MC, Dewenter H, Thun S, Gøeg KR, Karlsson D, Schulz S. Quantitative analysis of manual annotation of clinical text samples. *Int J Med Inform.* 2019 Mar;123:37-48.

Outlook

- Currently mapping of remaining 1125 top codes in progress
→ coverage of ~90% of all token mappings by the end of the year
- Use in practice (exact use cases to be defined) will show
 - Whether the current target representation (simple post-coordination is feasible)
 - Whether the loss of information in the “broader maps” is acceptable
- Use of map in inverse direction
 - So far irrelevant, because no SNOMED coded EHR content in Germany
 - Might be helpful to support manual mapping by creating shortlists
- Automated, fully meaning-preserving mapping would require
 - High quality SNOMED CT annotation of the EHR
 - Reconstruction of procedure codes as queries* – **Classifications ≠ Ontologies**

*On beyond Gruber: “Ontologies” in today’s biomedical information systems and the limits of OWL
Rector A, Schulz, S, Rodrigues JM, Chute CG, Solbrig H. Journal of Biomedical Informatics. In print.

Thanks for your attention

Contact:

stefan.schulz@averbis.com
philipp.daumke@averbis.com
matvey.palchuk@trinetx.com