

# Comparison of the international terminologies SNOMED CT, LOINC and ICD-11: Applicability in a guideline-conform pathology vocabulary for urothelial carcinoma in Germany

**Franziska Van Wüllen**

University of Applied Sciences Niederrhein

**Teja Falk Radke** (✉ [tfradke@gmail.com](mailto:tfradke@gmail.com))

Hochschule Niederrhein <https://orcid.org/0000-0002-1839-156X>

**Elisabeth Pantazoglou**

University of Applied Sciences Niederrhein

**Gunter Haroske**

Professional Association of German Pathologists

**Sylvia Thun**

Berlin Institute of Health

---

## Research article

**Keywords:** SNOMED CT, LOINC, ICD-11, Terminologies, Urothelial carcinoma, Interoperability

**Posted Date:** July 17th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.11634/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Objective** With almost 30,000 new cases per year, urothelial carcinomas account for a significant proportion of cancer cases in Germany. Respective guidelines serve to help pathologists evaluate tumor material according to international classification standards, but to ensure interoperability, further regulations are required. Therefore, the study presented in this work aimed at improving the informational situation by evaluating the applicability of the international terminologies in the scope of urothelial carcinoma in Germany. **Methods** Based on the S1-guideline "Urothelkarzinom", a collection of terms recommended for a pathology vocabulary was compiled and mapped to SNOMED CT (Systematic Nomenclature of Medical Terms), LOINC (Logical Observation Identifiers Names and Codes) and ICD-11 (International Classification of Diseases 11th Revision), respectively. **Results** Of the 168 terms required, 163 (97.02%) could be mapped to SNOMED CT, 66 (39.29%) to LOINC and 70 (41.67%) to ICD-11. However, considering the equivalence of each coding and restricting the mapping accordingly resulted in significantly lower coverage. When aiming at absolute equivalence, even combining all three terminologies resulted in only 138 (82.14%) terms being mappable adequately. **Discussion** Results prove that currently even combining established terminologies does not cover the terms required for a standardized documentation of urothelial findings completely. They also highlight the importance of SNOMED CT, as within this study it provided the largest proportion of mappable terms. Results also clearly demonstrated that especially SNOMED CT and LOINC require extensive knowledge on the respective terminology itself as well as on the respective medical field to ensure reliable mappings.

## Background

Cancers of the ureter and the urinary bladder are among the most common cancer-related diseases[1], being highly associated to smoking or exposure to certain agents at work[2]. In Germany they account for nearly 30,000 new cases per year and accordingly, specific guidelines address the challenge of providing guidance in order to describe a specific tumor, its location, etc. unambiguously[3]. However, with regard to the progressing digitalization and international data exchange, further requirements, especially a conjointly used vocabulary, need to be postulated to promote and ensure interoperability between the different systems used by the various stakeholders. This would not only reduce the risk of misunderstandings due to transmission errors (such as harbored by illegible handwriting or low-quality facsimiles), but also allow for automation of conversion steps (e.g. when transferring data from the lab into the hospital information system and/or into the electronic medical record), resulting in time and effort saved.

The work presented in this study focused on examining the applicability of selected international terminologies as possible means to convert the S1-guideline "Urothelkarzinom" (urothelial carcinoma) into an interoperable electronic form[4]. In detail, the following terminologies were evaluated with regard to mappability of required terms as well as their respective equivalence:

- SNOMED CT (Systematic Nomenclature of Medical Terms – Clinical Terms)

- LOINC (Logical Observation Identifiers Names and Codes)
- ICD (International Classification of Diseases)

These terminologies were chosen since they are the ones most commonly used on an international level, although, as depicted in table 1, they differ in their intended utilization. For example, the ICD classification originally derived from the need to summarize mortality-related diseases for statistical purposes while the focus of LOINC is on the exchange of clinical observations. While ICD, as of June 2019, is still used officially in form of the 10th revision (ICD-10), the updated and fundamentally improved 11th revision (ICD-11) will replace it within the next years and was accordingly used in this study.

### **Table 1. Overview on the terminologies evaluated within the study.**

*Values for available concepts were taken directly from the respective official websites (as of June 2019).*

## **Methods**

### Analysis of the S1-guideline “Urothelkarzinom”

This guideline provides instructions on pathological-anatomical diagnostics of tumors of the renal pelvis, the ureter and the urinary bladder. It is authored by the Bundesverband Deutscher Pathologen e.V. (Professional Association of German Pathologists) and distributed free of charge as download at [www.pathologie.de](http://www.pathologie.de). For the study presented in this work, relevant terms were identified, collected in a respective spreadsheet using MS Excel 2016 and grouped into two categories:

- Concepts: Representing clinical questions.
- Values: Representing possible answers.

In this context, terms were considered as relevant when required for the grading determined by the WHO[5], the TNM-classification or to specify the location of the tumor[6], give information on previous therapies, etc..

### Mapping

Prior to mapping, all terms were translated into English and an online research was performed for each one to confirm that the translation was valid, i.e. standardly used in clinical publications or reports. In case that an expression was unique to German language, the most appropriate translation was used according to the best of the authors’ knowledge and belief.

Mapping to SNOMED CT, ICD-11 and LOINC then was performed online using the browser provided by SNOMED International, the World Health Organization International and the Regenstrief Institute, respectively[7–9].

Initial mappings were performed by three of the authors, all with different background regarding terminologies/standards and expertise in pathology/medicine. Author A had knowledge on using all three terminologies, but no significant expertise related to pathology, author B is an expert in pathology with basic knowledge on terminologies and author C is an expert in terminologies with laboratory-medical background. These mappings were used to assess the accessibility of each terminology, i.e. whether the authors identified identical codes or differed (the respective assessment considered only terms where at least one of the authors proposed a code).

Lastly, the authors consented on the final mapping, which then was used for further analysis.

### Equivalence evaluation and inter-rater reliability

In brief, to each mapping a number between 0 and 4 was assigned, according to equivalence and considering determinants specified in the standard ISO/TS 21564 provided by the International Organization of Standardization. The respective classification was as follows:

- 0 : Exact semantic matching (code equals term)
- 1 : Complete overlap of the semantic domain (code covers term, but also more)
- 2 : Incomplete overlap of the semantic domain (code partially covers term)
- 3 : Rather a comparison than overlap of the semantic domain (code represents similar domain or term)
- 4 : No overlap of the semantic domain (no appropriate code found)

An example for an exact match (ISO 0) would be the term “Klinische T-Kategorie”, translated as “clinical t-category”, for which both, SNOMED CT and LOINC, provided equivalent codes (“399504009 | cT category (observable entity) |” and “21905-5 Primary tumor.clinical [Class] Cancer”, respectively).

For ISO 1, an example would be “Vorangegangene endoluminale Chemotherapie” (previous endoluminal chemotherapy). Here, LOINC provided the code “81167-9 Cancer treatment –preoperative |”, which has a broader scope than the original term.

Furthermore, for “Urothelkarzinom” (urothelial carcinoma), the only applicable code found in LOINC was “66125-6 Urinary bladder Pathology biopsy report”. Since this only covers a part of the original term, it was considered as ISO 2.

Finally, for “Andere Angaben zum Tumortyp” (Other Information on tumor type) a single, partially suitable code was found in LOINC: “52535-2 Other useful information”. As this only represents a comparable concept, it accordingly was classified as ISO 3.

Consequently, the lower the average equivalence figure, the better the general usefulness in appropriate clinical environments.

To ensure validity of the ISO-rating, three of the authors performed equivalence evaluations independently and Fleiss' Kappa, as measurement of the inter-rater correlations, was calculated for each terminology (note that only appropriate codes, i.e. ISO 0-3, were considered). Afterwards, terms were balloted for a definitive ISO-classification if necessary, i.e. if evaluations varied between the raters.

## Results

### Mapping

Evaluation of the S1-guideline "Urothelkarzinom" revealed 72 terms that were used for concepts, with an additional 96 terms for values. This total of 168 terms then was assessed for mappability to appropriate codes in SNOMED CT, LOINC and ICD-11, respectively.

The initial, independent mappings revealed that results varied with the respective author performing the mapping. As depicted in table 2 (part A), in general author A (knowledge on standards) mapped significantly less terms than authors B (expert medical background) or author C (expert knowledge on standards plus laboratory-medical background). As an explanation, it can be postulated that in general the authors decided to rather omit potentially unfitting mappings if in doubt, clearly stressing out the need for appropriate background knowledge. These results also reflect that, especially with regard to concepts, SNOMED CT and LOINC offer more potentially applicable codes, but are consequently more complex, as demonstrated by the respective lower uniformity. In contrast, all three authors identified the same codes for the same concepts when using ICD-11, albeit significantly less than with the other two terminologies.

### **Table 2. Mapping results for the S1 guideline "Urothelkarzinom".**

*A) Initially, terms for concepts and values were mapped to SNOMED CT, LOINC and ICD-11 by three individual raters with different qualification levels/background. Results are given as absolute number of codes (mappings) identified by the respective author while uniformity, given as percentage, refers to whether all authors identified identical codes or differed regarding each term, thus representing reproducibility of mappings by different coders. B) Subsequently, the authors consented on a final mapping. Results are given as absolute number as well as percentage of mappable terms.*

Prior to further assessment, terms that were not mapped identically were discussed and consented on, resulting in a final mapping. As also depicted in table 2 (part B), the final mappability varied for each terminology, ranging from 13.89% to 95.83% regarding terms for concepts and from 19.79% to 97.92% regarding terms for values.

The data also clearly demonstrate that SNOMED CT had the highest compatibility in any aspect (with a coverage of 97.02% for the total terms), followed by ICD-11 (with 41.67% coverage), while LOINC had the lowest (with 39.29% coverage).

### Equivalence evaluation and inter-rater reliability

While for most terms a corresponding code could be found in at least one of the three terminologies assessed, it had to be evaluated whether these codes corresponded to the original meaning. Therefore, an equivalence assessment was performed in accordance with ISO/TS 21564. As depicted in table 3, respective results demonstrated that, when only considering truly equivalent codes (ISO 0), amounts of appropriate mappings decreased significantly. E.g., for SNOMED CT instead of 163 (97.02%) only 127 (75.60%) of the codes still were applicable.

### **Table 3. Equivalence evaluation results for the S1 guideline “Urothelkarzinom”.**

*Mapped terms found in SNOMED CT, LOINC and ICD-11 were evaluated with regard to equivalence and assigned numbers ranging from 0 to 4, according to ISO/TS 21564. Average equivalence was calculated for all mapped terms (i.e. ISO 0 to 3). Fleiss’ Kappa then was assessed for the ratings of each terminology, representing the reliability of the respective average equivalence.*

Regarding average equivalence and Fleiss’ Kappa, as means to measure the overall applicability of a terminology as well as its reliability, respectively, significant differences were observed. Here, SNOMED CT and LOINC both demonstrated a better average equivalence regarding the mapped terms than ICD-11 (with 0.3374 and 0.4848 vs. 1.1143) while, vice versa, the overall reliability (Fleiss’ Kappa) was higher for the later. Note that the high reliability for ICD-11 correlates directly to its lower equivalence, as “obviously” non-equivalent mappings (ISO 1-3) are more likely to be evaluated identically as such by all raters.

### Coverage

The results from the mapping already indicate that none of the terminologies alone is sufficient to cover all terms needed. Therefore, coverage by different combinations was assessed. This analysis also includes the different degrees with regard to equivalence, although it has to be noted that only a rather strict equivalence (i.e. ISO 0-1) is feasible and legit in clinical use.

### **Table 4. Coverage of terms required by combinations of SNOMED CT with other terminologies.**

*Results are given as absolute numbers as well as percentage for each combination and for different ranges of equivalence level.*

As displayed in table 4, since SNOMED CT already cover most of those terms mappable by LOINC and ICD-11, even the combination of all three terminologies only results in a minor improvement with 137 (81.55%) terms at high equivalence. (In brief, LOINC can provide further 10 and ICD-11 only 1 terms, respectively.) It is also notable that combining LOINC and ICD-11 provides mapping only for 65 (38.68%) of the required ISO-0 terms, which is significantly less than SNOMED CT alone (127 terms or 75.60%, respectively; see table 2B). When lower equivalence is accepted (i.e. ISO 0-3), the contribution by LOINC and ICD-11 gets even less with only LOINC adding 1 single term to 163 terms covered by SNOMED CT.

## **Discussion**

As indicated by the results presented in this study, as of now, none of the terminologies assessed is sufficient to cover definitely all terms needed in the S1-guideline “Urothelkarzinom” on its own. Even with low-equivalent expressions (ISO 0-3) accepted, codes only for 164 (97.62%) of the 168 terms required could be mapped when combining SNOMED CT, LOINC and ICD-11. Though, for unambiguous clinical use, restriction to at least ISO 0-1 (as required in the ISO/TS 21654) or even solely ISO 0 would be mandatory, resulting in a coverage of only 157 (93.45%) or 138 (82.14%) suitable terms, respectively.

Among the terminologies applied in this study, SNOMED CT clearly is the most promising one, as it alone covers already 127 terms (75.60%) at ISO-0 equivalence. (Interestingly, this is nearly identical to the coverage previously reported for its application in histopathological finding)[10]. In comparison, the combination of LOINC and ICD-11 without SNOMED CT resulted in only 118 (70.24%) terms even for low equivalence, and 65 (38.69%) for absolute equivalence.

The main reason for SNOMED CT offering a higher coverage and equivalence is likely due to its flexibility. By applying post-coordination, codes can often be precisely defined to exactly match a specific term. This, however, requires detailed knowledge on the hierarchy and attributes allowed (especially the extensions used in post-coordination or semantic tags)[11,12], as well as rigorous quality control, as the results from the initial mapping stress out. Especially when coding terms for more complicated areas, such as the pathology vocabulary used in this study, it is mandatory that the responsible coder has sufficient knowledge not only on the terminologies and their requirements, but also on the medical terms and their specific, detailed meanings and context. For example, when coding a tumor relapse, the concept “Recurrent tumor” benefits from respective extensions to be unambiguously defined as stand-alone entry, referring to a specific tumor in the patient’s history. Without post-coordination, the general context is required in order to interpret the concept correctly. This might be negligible for complete electronic reports (since all data are represented in the according context), but is crucial for electronic evaluation or search in clinical studies (e.g. when searching for data on relapse of a specific tumor type).

While these results also strongly emphasize the current problems associated with the progressing digitization and the resulting need for interoperability, it has to be mentioned that SNOMED CT and LOINC are under constant development[13]. In addition, since 2012, there are constant efforts regarding the harmonization between these two terminologies. Thus, coverage and subsequently applicability might improve in future but, of course, the responsible organizations are also required to adapt or develop national guidelines in accordance to international standards whenever possible (e.g. by requesting codes for national extensions).

In addition, be reminded that the focus of this study was solely on urothelial carcinomas. Therefore, it does not necessarily reflect applicability to reports of other cancers, although similar results can be expected. On the side, the results also confirm that mappability does not necessarily correlate with semantic equivalence. While in this setting SNOMED CT had both, the highest mappability as well as the best equivalence, there was a notable difference in between LOINC and ICD-11. While both provided comparable amounts of applicable codes (66 and 70, respectively), for those codes found LOINC had a

significantly better average equivalence (0.4848) than ICD-11 (1.1143), which also was comparable to that of SNOMED CT (0.3374).

## Conclusions

In brief, the results of this study can be summarized as follows:

- Combining SNOMED CT, LOINC and ICD-11 is a feasible approach to compile a guideline-conform pathology vocabulary for urothelial carcinomas based on international standards in Germany.
- While SNOMED CT alone already provides most of the terms needed, supplementation with codes from LOINC/ICD-11 is still required.
- Solely combining LOINC and ICD-11 covers less than 50% of the terms needed and as such is regarded not feasible.
- Basic knowledge on the standards alone is not sufficient for complex medical domains. The personnel responsible for mapping also needs experience in post-coordination as well as knowledge of the specific medical field in order to choose the correct and most equivalent codes.

## Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors' contributions



This paper was jointly conceived and written by the authors, with FVW and TFR having contributed equally.

Acknowledgements

Not applicable.

## References

1. Burger M, Catto JWF, Dalbagni G, et al. Epidemiology and risk factors of urothelial bladder cancer. *Eur Urol* 2013;63(2):234–41.
2. Freedman ND, Silverman DT, Hollenbeck AR, et al. Association between smoking and risk of bladder cancer among men and women. *JAMA* 2011;306(7):737–45.
3. Barnes B, ed. Bericht zum Krebsgeschehen in Deutschland 2016. Berlin: Robert Koch-Institut 2016.
4. Helpap B, Knüchel-Clarke R, Hartmann A. Anleitung zur pathologisch-anatomischen Diagnostik von Tumoren des Nierenbeckens, des Ureters und der Harnblase. Berlin: Bundesverband Deutscher Pathologen e. V. & Deutsche Gesellschaft für Pathologie 2012.
5. Moch H, Humphrey PA, Ulbright TM. WHO classification of tumours of the urinary system and male genital organs, 4th edn. 2016.
6. Brierley JD, Gospodarowicz MK, Wittekind C, eds. TNM Classification of Malignant Tumours, 8th edn. New York: John Wiley & Sons 2016.
7. SNOMED International. SNOMED CT browser. Available at: <https://browser.ihtsdotools.org> Accessed June 12, 2019.
8. World Health Organization. International statistical classification of diseases and related health problems (11th Revision) 2018. Available at: <https://icd.who.int/browse11/l-m/en> Accessed June 12, 2019.
9. Regenstrief Institute. Search LOINC. Available at: <https://search.loinc.org/searchLOINC/> Accessed June 12, 2019.
10. Campbell WS, Campbell JR, West WW, et al. Semantic analysis of SNOMED CT for a post-coordinated database of histopathology findings. *J Am Med Inform Assoc* 2014;21(5):885–92.
11. Rector A, Iannone L. Lexically suggest, logically define: quality assurance of the use of qualifiers and expected results of post-coordination in SNOMED CT. *J Biomed Inform* 2012;45(2):199–209.

12. Bona JP, Ceusters W. Mismatches between major subhierarchies and semantic tags in SNOMED CT. *J Biomed Inform* 2018;81:1–15.

13. Bodenreider O, Cornet R, Vreeman DJ. Recent Developments in Clinical Terminologies - SNOMED CT, LOINC, and RxNorm. *Yearb Med Inform* 2018;27(1):129–39.

## Tables

**Table 1. Overview on the terminologies evaluated within the study.**

	SNOMED CT	LOINC	ICD-11
Main application/focus	Clinical content	Medical lab values	Health status/Diagnoses
Structure	Polyhierarchy	Coded unique 6-axis terms	Polyhierarchy with optional linearization
Available terms/concepts	>335,000	>89,000	~ 55,000
Continuous development	Yes	Yes	No
Post-coordination	Yes	Partially/restricted	Yes
Charging system	Licensing (national/individual)	Free of charge	Free of charge

*Values for available concepts were taken directly from the respective official websites (as of June 2019).*

**Table 2. Mapping results for the S1 guideline “Urothelkarzinom”.**

A)

Terminology	Terms for	Author A	Author B	Author C	Total unique	Uniformity
SNOMED CT	Concepts	52	69	69	69	27,54%
	Values	85	91	94	94	75,53%
LOINC	Concepts	25	51	48	53	32,08%
	Values	23	23	19	23	82,61%
ICD-11	Concepts	10	10	10	10	100,00%
	Values	50	50	62	62	80,65%

B)

Terms for	SNOMED CT	LOINC	ICD-11
Concepts (72)	69 (95.83%)	47 (65.28%)	10 (13.89%)
Values (96)	94 (97.92%)	19 (19.79%)	60 (62.50%)
Total terms (168)	163 (97.02%)	66 (39.29%)	70 (41.67%)

A) Initially, terms for concepts and values were mapped to SNOMED CT, LOINC and ICD-11 by three individual raters with different qualification levels/background. Results are given as absolute number of codes (mappings) identified by the respective author while uniformity, given as percentage, refers to whether all authors identified identical codes or differed regarding each term, thus representing reproducibility of mappings by different coders.

B) Subsequently, the authors consented on a final mapping. Results are given as absolute number as well as percentage of mappable terms.

**Table 3. Equivalence evaluation results for the S1 guideline “Urothelkarzinom”.**

ISO/TS 21564	SNOMED CT	LOINC	ICD-11
ISO 0	127 (75.60%)	46 (27.38%)	23 (13.69%)
ISO 1	26 (15.48%)	12 (7.14%)	26 (15.48%)
ISO 2	1 (0.60%)	4 (2.38%)	11 (6.55%)
ISO 3	9 (5.36%)	4 (2.38%)	10 (5.95%)
ISO 4	5 (2.98%)	102 (60.71%)	98 (58.33%)
<b>Average equivalence (mapped)</b>	<b>0.3374</b>	<b>0.4848</b>	<b>1.1143</b>
Fleiss' Kappa (mapped)	0.6409	0.4713	0.8036

Mapped terms found in SNOMED CT, LOINC and ICD-11 were evaluated with regard to equivalence and assigned numbers ranging from 0 to 4, according to ISO/TS 21564. Average equivalence was calculated for all mapped terms (i.e. ISO 0 to 3). Fleiss' Kappa then was assessed for the ratings of each terminology, representing the reliability of the respective average equivalence.

**Table 4. Coverage of terms required by combinations of SNOMED CT with other terminologies.**

Equivalence level ISO/TS 21564	SNOMED CT	SNOMED CT	ICD-11 +LOINC	SNOMED CT
	+ LOINC	+ ICD-11		+ ICD-11 +LOINC
ISO 0	137 (81.55%)	128 (76.19%)	65 (38.68%)	138 (82.14%)
ISO 0-1	156 (92.86%)	154 (91.67%)	102 (60.71%)	157 (93.45%)
ISO 0-2	157 (93.45%)	154 (91.67%)	109 (64.88%)	157 (93.45%)
ISO 0-3	164 (97.62%)	163 (97.02%)	118 (70.24%)	164 (97.62%)

Results are given as absolute numbers as well as percentage for each combination and for different ranges of equivalence level.