

Original Paper

Real-World Applicability of Multiple International Terminology Systems in Oncology: Mapping Study of Clinical Terms for Cancer Diagnosis and Treatment

Boyun Jang^{1,2*}, MPH; Eunhye Park^{1,2*}, MPH; Sangwon Lee¹, MS; Eunbi Ju^{1,2}, BS; Yewon Jang^{1,2}, BS; In Hae Kwak¹, BS; Kui Son Choi^{1,3}, PhD; Hyun-Joo Kong¹, MS; Hyo Soung Cha^{1,3}, PhD; Hyun-Jin Kim^{1,3}, PhD; Heejung Chae^{1,4}, MD, PhD

¹Cancer Data Center, National Cancer Control Institute, National Cancer Center, Goyang, South Korea

²Department of Medical Information Management, National Cancer Center, Goyang, South Korea

³Department of Public Health and Artificial Intelligence, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, South Korea

⁴Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, South Korea

*these authors contributed equally

Corresponding Author:

Heejung Chae, MD, PhD
Department of Internal Medicine
Seoul National University Bundang Hospital
82 Gumi-ro 173beon-gil
Seongnam 13620
South Korea
Phone: +82-31-787-7863
Email: hchae21@gmail.com

Abstract

Background: With the rapid expansion of health care data, multiple international standardized terminology systems have been developed and refined to ensure semantic interoperability, which is a crucial factor in facilitating collaborative use.

Objective: This study aimed to evaluate the real-world use and applicability of standardized terminology systems by mapping clinical terms from routine cancer care and to propose guidelines that optimize mapping.

Methods: Clinical terms generated during cancer care at the National Cancer Center Korea were extracted and mapped to multiple target terminology systems: the Korean Standard Classification of Diseases; *International Classification of Diseases, 9th Revision, Clinical Modification*; Logical Observation Identifiers Names and Codes; RxNorm; Anatomical Therapeutic Chemical classification; and Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). A multidisciplinary working group of health care information managers, data experts, and oncologists verified the mapping outcomes, calculated the cross-verification discordance rate, and compared the coverage across systems.

Results: A total of 3472 source terms were identified across 7 domains: chief complaint, diagnosis, laboratory test, imaging test, surgery/procedure, medication, and radiation therapy. With postcoordination and predefined mapping guidelines, SNOMED CT achieved high overall mapping rates (98.3%-100%) across domains. Applying postcoordination as needed substantially improved exact semantic match rates, which exceeded 90% in most domains. However, more than 40% of source terms were broadly mapped in the imaging test and radiation therapy domains because procedural details such as imaging technique and radiation field configuration could not be represented in SNOMED CT. The cross-verification discordance rate between mappers, defined as the proportion of source terms with disagreement, remained low (0.3%-1.3%). RxNorm and Logical Observation Identifiers Names and Codes showed complementary strengths within their specialized domains, adding detail or specificity where SNOMED CT was less complete.

Conclusions: Standardized terminology systems are feasible and effective in encoding clinical terms for cancer care. SNOMED CT was the most comprehensive, and predefined mapping guidelines provide a structured framework for mapping and consensus resolution. Domain-specific systems should be used alongside SNOMED CT, and all systems should be updated regularly to reflect evolving clinical practice.

Keywords: clinical terms; health care terminology; mapping; oncology; Systematized Nomenclature of Medicine Clinical Terms; SNOMED CT; terminology standard

Introduction

With advances in IT and precision medicine, health care data volumes have expanded exponentially [1]. The growth of multicenter research and federated analyses has increased the complexity of data exchange. Unlike other big data, health care data originates from diverse sources, many of which are unstructured, creating interoperability challenges when integrating real-world data for clinical research [2]. Effective exchange requires 4 levels of interoperability: foundational, structural, semantic, and organizational. Semantic interoperability, enabling different systems to interpret information with the same meaning, is critical for complex, large-scale health care data [3,4].

Standardized terminology systems that define and organize medical terms have been developed to achieve semantic interoperability. In the 1940s, the World Health Organization (WHO) introduced the *International Classification of Diseases (ICD)* for morbidity and mortality statistics, the current version of which is the *ICD-11*. With advances in oncology, more specialized systems have been established, such as the *International Classification of Diseases for Oncology (ICD-O)*, which codes neoplasms by tumor site (topography) and histology (morphology) [5]. Additional domain-specific systems include the *ICD-10-Procedure Coding System (ICD-10-PCS)* for surgical procedures and treatments, Logical Observation Identifiers Names and Codes (LOINC) for laboratory tests, and RxNorm for medications [6-8]. Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), originating from the College of American Pathologists' Systematized Nomenclature of Pathology and later merged with the United Kingdom's Read Codes, is the most widely adopted standardized medical terminology system [9].

Despite substantial progress in terminology development, gaps remain in real-world clinical application. High-quality research requires minimizing information loss during terminology mapping. However, the application of multiple standardized terminology systems to original real-world clinical terms containing detailed, unaltered information has not been widely examined, and potential intermapper discrepancies have not been fully evaluated. Moreover, few studies have addressed the unique complexities of

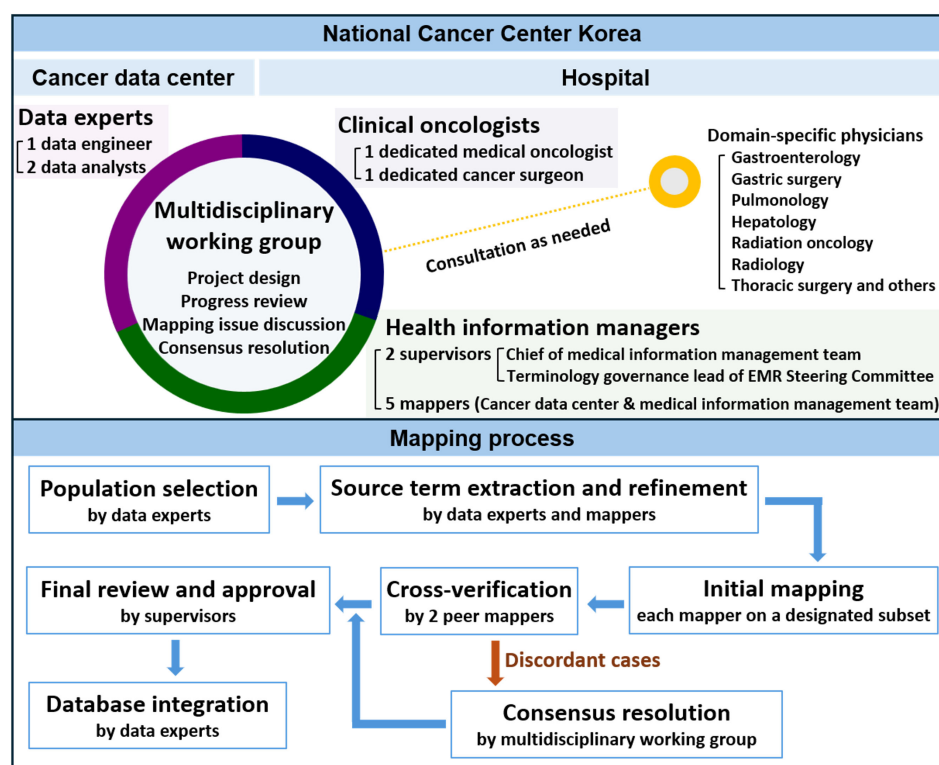
patients with cancer, who undergo more diverse and intricate diagnostic and treatment processes than general patients. In this study, a multidisciplinary team mapped clinical terms from routine cancer care to multiple standardized terminology systems to assess their applicability, identify challenges, and propose guidelines for optimizing real-world clinical terminology mapping.

Methods

Working Group

The Cancer Data Center of the National Cancer Center Korea (NCCCK), South Korea's National Cancer Data Center, establishes the data infrastructure and utilization environment required to promote nationwide cancer data research. As part of these efforts, a clinical term-to-standard terminology mapping project was initiated. A multidisciplinary working group of 3 data experts, 7 health information managers, and 2 clinical oncologists was convened. Data experts, who are responsible for building and maintaining the National Cancer Data Center research database and NCCCK Clinical Research Data Warehouse (CRDW), conducted source term extraction, integration of the mapping outputs into the database, and quality control to ensure data integrity and operational suitability for standardization and reuse. Five health information managers, with expertise in electronic medical record (EMR) data management, multicenter data standardization, and standard terminology mapping, served as primary mappers. Two senior health information managers (the Chief of Medical Information Management Team and the Terminology Governance Lead of the EMR Steering Committee) supervised the overall mapping process and performed final review and approval of all mapping outputs. The 2 dedicated clinical oncologists, 1 medical oncologist and 1 cancer surgeon, contributed domain expertise in cancer care to review whether the clinical meaning of the source terms was preserved during mapping. When specialized clinical interpretation was required, additional consultation was obtained from relevant physicians within the institution. The working group held regular monthly meetings and biweekly written discussions for project design, progress review, mapping issue discussion, and consensus resolution (Figure 1).

Figure 1. Structure of the multidisciplinary working group and workflow of the standard terminology mapping process. The National Cancer Center Korea established a multidisciplinary working group comprising data experts, certified health information managers, and clinical oncologists, with domain-specific physicians consulted as needed. Each step of the mapping process was carried out by the designated members of the working group, as indicated. EMR: electronic medical record.



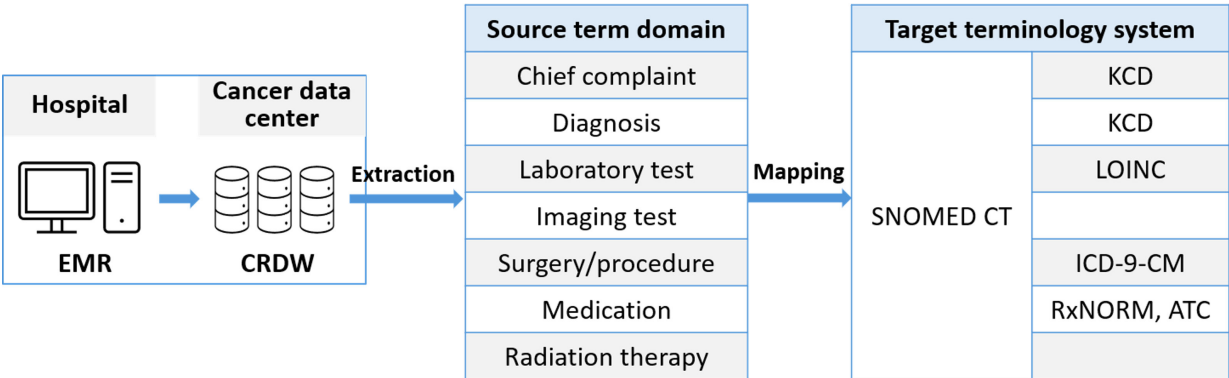
Study Population and Source Terms

Candidate source terms were identified from clinical data in the CRDW, which collects comprehensive information from the EMR system and houses deidentified data for research. Patients diagnosed with one of 6 primary cancers (stomach, breast, liver, colorectal, pancreatic, or lung) at the NCKK between 2010 and 2023 were included. Inclusion was based on *ICD-O-3* codes: C16 (stomach), C50 (breast), C18/C19/C20 (colorectal), C22 (liver), C25 (pancreas), and C33/C34 (lung). Patients with intrahepatic bile duct cancer (C22.1) and islet cell tumors of the pancreas (C25.4) were excluded. Clinical data were extracted and classified into 7 domains: chief complaint, diagnosis, laboratory test, imaging test, surgery/procedure, medication, and radiation therapy. The source terms underwent minimal refinement to preserve their original clinical meaning: correction of typographical errors and spacing, expansion of abbreviations, removal of duplicates, and splitting of multivalued terms. Electronic data interchange (EDI) reimbursement codes for Korean National Health Insurance (NHI) claims and their effective dates were extracted from the CRDW for reference during mapping.

Mapping

Each source term was mapped to relevant standardized terminology systems according to its clinical domain. Domain-specific mappings included the Korean Standard Classification of Diseases (KCD), LOINC, *International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM)* Volume 3, RxNorm, and the Anatomical Therapeutic Chemical (ATC); all terms, regardless of domain, were mapped to SNOMED CT (Figure 2). Although newer procedure classifications such as *ICD-10-PCS* have been developed, *ICD-9-CM* Volume 3 has been used by many countries and remains in use across several Asian countries because of its established role within national coding infrastructures and the operational complexity of migration [10-14]. In South Korea, *ICD-9-CM* continues to serve as the procedure classification reference for national health statistics, and clinical terms are generated and managed on this basis [15,16]. Therefore, *ICD-9-CM* Volume 3 was selected for the surgery/procedure domain to reflect the source data context.

Figure 2. Mapping of clinical source terms to domain-specific standard terminology systems. Clinical terms were extracted from the clinical research data warehouse (CRDW), which contains deidentified data sourced from the hospital’s electronic medical records (EMR) system. Source terms from 7 clinical domains—chief complaint, diagnosis, laboratory test, imaging test, surgery/procedure, medication, and radiation therapy—were mapped to appropriate standardized terminology systems. ATC: Anatomical Therapeutic Chemical; ICD-9-CM: *International Classification of Diseases, 9th Revision, Clinical Modification*; KCD: Korean Standard Classification of Diseases; LOINC: Logical Observation Identifiers Names and Codes; RxNorm: normalized naming system for generic and branded drugs.



All mappings followed the commonly recognized rules for each terminology system. For SNOMED CT, mappings adhered to the SNOMED CT Editorial Guide, compositional grammar, and concept model principles [17,18]. For postcoordination, focus concepts, attributes, and attribute values were selected in accordance with the permitted attributes and value ranges defined in the Machine Readable Concept Model (MRCM) [19]. The working group established and shared predefined guidelines with all mappers to specify mapping priorities and ensure consistency when multiple MRCM-compliant mapping options were available within the SNOMED CT framework. SNOMED CT mapping was categorized as “exact” when the target concept matched the source term in meaning, “broad” when the target concept was broader than the source, “narrow” when the target concept was narrower than the source, and “not-mapped” when no suitable target existed. Given the potential loss of semantic specificity and reduced fidelity to the original clinical meaning, narrow mappings were not counted as successful mappings, and grouped with “not-mapped” category as “narrow/not-mapped” in the final mapping performance analysis. Information that the SNOMED CT Editorial Guide identifies as outside the scope of the SNOMED CT International Edition—such as brand names in the medicinal product [20] and administrative or reimbursement-related procedure descriptions, including procedure counts or procedure order [21,22]—was not classified as a broad mapping when absent from the target concept. In contrast, the clinical meaning of the source term itself was considered essential to the mapping and had to be represented in the target concept for the mapping to be classified as exact.

Following initial mapping by 5 mappers each assigned to a designated subset of source terms, 2 peer mappers cross-reviewed each other’s mapping results and recorded agreement or disagreement. Any disagreement was escalated to the multidisciplinary working group for consensus resolution. Clinical oncologists verified that the clinical meaning of each source term was accurately preserved, while data experts and supervising health information managers assessed whether the mapping was operationally appropriate

and not overly specific in a way that would compromise standardization and reuse. Final decisions were made by prioritizing (1) accurate preservation of the original clinical meaning and (2) consistent application of predefined mapping principles. Cross-verification results were summarized using the cross-verification discordance rate, defined as the proportion of source terms for which disagreement was identified during cross-verification.

Ethical Considerations

This study analyzed a terminology dataset of clinical source terms extracted from the institutional deidentified data warehouse and their mapping results, rather than individual-level clinical records. The unit of analysis was the terminology concept rather than the individual patient, and the analytic dataset contained no identifiable personal information or patient-level linkage. Accordingly, this study was granted exemption by the Institutional Review Board of NCKK (No. 2026-0204), with a waiver of informed consent. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Results

Identification of Source Terms

Among the 468,193 patients registered in the CRDW by 2023, 66,380 who were diagnosed with primary stomach, breast, liver, colorectal, pancreatic, or lung cancer between 2010 and 2023 were included. Following extraction, source terms underwent a refinement process, including typographical correction, expansion of abbreviations, duplicate removal, and splitting of multivalued terms (Supplementary Material S1 in Multimedia Appendix 1). No semantic modification or simplification was allowed. A total of 3472 clinical terms were finalized as source terms (Supplementary Material S2 in Multimedia Appendix 1). For the chief complaint domain, the 200 most frequent terms were extracted to prioritize the terms most relevant to routine clinical use. For the medication domain, information on drugs related to cancer prevention

and treatment was extracted. All available clinical terms were extracted for the remaining domains.

Predefined Mapping Guidelines

While all mappings followed the official guidance of each terminology system, the working group established predefined guidelines incorporating oncology-specific considerations to better preserve the original clinical context of source terms and improve consistency across mappers when multiple clinically plausible and terminology-compliant mapping options were available. The predefined guidelines were as follows.

To map chief complaint and diagnosis terms to KCD, if the source term is semantically identical to a subterm of the KCD target term, it should be mapped to that subterm; otherwise, it should be mapped to the main term, which is the primary identifier for the KCD code. In the surgery/procedure domain, if a single-source term contains multiple surgeries or procedures, each surgery or procedure should be mapped individually. The Korean Standard Terminology of Medicine can be used as a supplementary reference for mapping to the ICD-9-CM [23]. Although the EDI codes for NHI claims may

be considered when available, priority should be given to the meaning of the source term from the CRDW.

For SNOMED CT, each source term should typically be mapped to the relevant hierarchy; however, if a chief complaint indicates the purpose of admission (eg, “for endoscopic variceal ligation”), it should be mapped to the procedure hierarchy. In the medication domain, information outside the scope of SNOMED CT, such as brand name or clinical indication (eg, transcatheter arterial chemoembolization), should be excluded. When the focus concept lacks dose form or strength information, postcoordination should be used. In the radiation therapy domain, if the meaning of a source term is ambiguous, it should be clarified by referencing the EDI codes for NHI claims before mapping. Narrow mappings that might distort the original meaning of the source term should be avoided, and exact mappings should be prioritized over broad mappings. When multiple postcoordinated representations are possible within SNOMED CT Editorial Guide and MRCM-defined attribute constraints, focus concepts, attributes, and attribute values should be selected according to the priority rules summarized in Table 1.

Table 1. Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) postcoordination priorities by clinical domains.

Domain and focus concept	Attribute-value (semantic tag)
Chief complaint	
Symptom	
Select the focus concept corresponding to the main symptom	Postcoordinate “Site (body structure),” if applicable
Reason for admission	
Select “304566005 Admission for treatment (procedure) ” as the focus concept	Apply postcoordination in the following priority order: ① Treatment name (procedure, regimen/therapy) ② Causative procedure (procedure)
Select “76464004 Hospital admission, for observation (procedure) ” as the focus concept	Postcoordinate “Test name (procedure),” if applicable
Complications	
Select the focus concept corresponding to the main symptom	Postcoordinate “Causative procedure or surgery (procedure),” if applicable
Diagnosis	
Select the focus concept that best captures the following information, in priority order: ① Behavior + morphology + site ② Morphology + site ③ Behavior + site ④ Behavior + morphology ⑤ Morphology ⑥ Behavior	Postcoordinate additional information not captured by the focus concept in the following priority order: ① Detailed site or laterality (body structure) ② Detailed morphology (morphologic abnormality) ③ Other information (disorder, observable entity, qualifier value, procedure)
Laboratory test	
Select the focus concept that best captures the following information, in priority order: ① Test name + sample type ② Test name	Postcoordinate additional information not captured by the focus concept in the following priority order: ① Test substance (substance, cell) ② (Detailed) sample type (specimen) ③ Test method (procedure) ④ Quantitative/qualitative (qualifier value) ⑤ Other information (qualifier value, physical object, person)

Domain and focus concept	Attribute-value (semantic tag)
Imaging test	
In-house imaging study	
Select the focus concept that best captures the following information, in priority order:	Postcoordinate additional information not captured by the focus concept in the following priority order:
① Modality + site + position	① (Detailed) site or laterality (body structure)
② Modality + site + injected agents	② Injected agents (substance, procedure)
③ Modality + site	③ Test method (procedure)
④ Modality + injected agents	④ Purpose of test (procedure)
⑤ Modality	⑤ Using device (physical object, physical force)
	⑥ Other information (disorder, finding, qualifier value, physical force, person)
Review of outside imaging	
Select “28191001 Consultation and report by radiologist (procedure) ” as the focus concept	Apply postcoordination in the following priority order:
	① Modality + site + injected agents (procedure)
	② Modality + site (procedure)
	③ (Detailed) site (body structure)
Surgery/procedure	
No additional priority rules	No additional priority rules
Medication	
Select the focus concept that best captures the following information, in priority order:	Postcoordinate additional information not captured by the focus concept in the following priority order:
① Ingredient + dose form + strength	① Detailed dose form (dose form)
② Ingredient + dose form	② Detailed dose form unit (unit of presentation)
	③ Numerator value (qualifier value)
	④ Numerator unit (qualifier value)
	⑤ Denominator value (qualifier value)
	⑥ Denominator unit (qualifier value)
Radiation therapy	
No additional priority rules needed	No additional priority rules needed

All mappings followed the SNOMED CT Editorial Guide and MRCM. The priorities listed below were applied additionally, only when multiple valid mapping options existed within these rules.

SNOMED CT concepts are displayed in the format “concept identifier |term|.” The numeric identifier represents the SNOMED CT concept ID, and the term enclosed by vertical bars (“|” “|”) is the human-readable SNOMED CT term [18]. Numbered items (①, ②, ③, etc) indicate the priority order used in this study for selecting focus concepts or postcoordination attributes when a source term contained multiple mappable elements.

Mapping Results

Table 2 provides examples of how source terms were mapped to the corresponding target terminology systems in each clinical domain and how the SNOMED CT postcoordination priorities described in Table 1 were applied. Supplementary Material S3 in Multimedia Appendix 1 provides further details on the selection of SNOMED CT postcoordination attributes. For instance, in the case of “Carcinosarcoma, connective and soft tissue, unspecified” in the diagnosis domain, no single target term representing both “① behavior

+ morphology + site” and “② morphology + site” could be identified. Therefore, the focus concepts “448274000 |Malignant neoplasm of connective tissue (disorder)|” and “269469005 |Malignant neoplasm of soft tissue (disorder)|” were used to capture the “③ behavior (malignant) and the sites (connective and soft tissue)” in a 1:N relationship. The remaining information, morphology, was captured using a postcoordinated expression: “116676008 |Associated morphology(attribute)| = 63264007 |Carcinosarcoma(morphologic abnormality)|.” In the laboratory test domain, following the priority order of “① Test name + sample type,” no single concept capturing both elements was available for “Anti-HCV, immunoassay.” The second priority, test name, was mapped to the focus concept “64411004 |Hepatitis C antibody measurement (procedure)|.” The remaining details were represented through postcoordination by specifying the “② (Detailed) sample type (specimen),” and “③ Test method (procedure)” as follows: “116686009 |Has specimen(attribute)| = 119364003 |Serum specimen (specimen)|” and “370129005 |Measurement method (attribute)| = 104417008 |Immunoassay, qualitative, multiple step method (procedure)|,” respectively.

Table 2. Examples of mappings for each domain.

Domain: clinical source terms	Target
Chief complaint: Back pain	
KCD ^a	M54.99 lDorsalgia, unspecified, site unspecifiedl
SNOMED CT ^b	161891005 lBackache (finding)l
(1:1 Exact)	
Diagnosis: carcinosarcoma, connective and soft tissue, unspecified	
KCD ^c	C49.9 lMalignant neoplasm of connective and soft tissue, unspecifiedl
SNOMED CT	① 448274000 lMalignant neoplasm of connective tissue (disorder)l : { 116676008 lAssociated morphology (attribute)l = 63264007 lCarcinosarcoma (morphologic abnormality)l } ② 269469005 lMalignant neoplasm of soft tissue (disorder)l : { 116676008 lAssociated morphology (attribute)l = 63264007 lCarcinosarcoma (morphologic abnormality)l }
(1:N Exact)	
Laboratory test: Anti-HCV ^d (test name: hepatitis test and serology; sample name: serum; test unit: X) (EDI ^e : precision immunoassay-hepatitis C antibody)	
LOINC ^f	13955-0 lHepatitis C virus Ab [Presence] in Serum or Plasma by Immunoassayl
SNOMED CT	64411004 lHepatitis C antibody measurement (procedure)l : { 116686009 lHas specimen (attribute)l = 119364003 lSerum specimen (specimen)l, 370129005 lMeasurement method (attribute)l = 104417008 lImmunoassay, qualitative, multiple step method (procedure)l }
(1:1 Exact)	
Imaging test: CT ^g dynamic liver (contrast)	
SNOMED CT	438591004 lComputed tomography triple phase of liver (procedure)l
(1:1 Exact)	
Surgery/ procedure: Laparoscopic pelvic lymphadenectomy	
ICD-9-CM ^h	40.3 lRegional lymph node excisionl
SNOMED CT	14059008 lPelvic lymphadenectomy (procedure)l : { 424226004 lUsing device (attribute)l = 86174004 lLaparoscope, device (physical object)l }
(1:1 Exact)	
Medication: Adriamycin RDF ⁱ 10 mg (doxorubicin)	
RxNorm	1790097 l5 ML doxorubicin hydrochloride 2 mg/mL Injectionl
ATC ^j	L01DB01 lDoxorubicinl
SNOMED CT	1209144002 lProduct containing precisely doxorubicin hydrochloride 2 mg/mL conventional release solution for infusion and/or injectionl
(1:1 Exact)	
Radiation therapy: Intensity modulated radiation therapy [per session] (EDI: IMRT-VMAT)	
SNOMED CT	1156530009 lVolumetric modulated arc therapy (procedure)l
(1:1 Exact)	

^aKCD: Korean Standard Classification of Diseases.^bSNOMED CT: Systematized Nomenclature of Medicine Clinical Terms.^cKCD: Korean Standard Classification of Diseases.^dHepatitis C virus.^eEDI: EDI codes for NHI claims.^fLOINC: Logical Observation Identifiers Names and Codes.^gCT: computed tomography.^hICD-9-CM: *International Classification of Diseases, 9th Revision, Clinical Modification*.ⁱRDF: ready to dilute formulation.^jATC: Anatomical Therapeutic Chemical.

After mapping all the source terms, we examined how each term was represented across multiple target terminology systems to identify system-specific characteristics (Table 3). In the chief complaint and diagnosis domains, KCD could not fully represent certain terms, such as purposes of admission (eg, “for transplantation work-up”) or morphological detail in “adenocarcinoma of sigmoid colon,” whereas SNOMED CT captured these terms via postcoordination. In the surgery/procedure domain, behavioral characteristics (eg, malignancy) could not be included in ICD-9-CM procedure codes, whereas SNOMED CT allowed behavior to be represented as attributes in addition to the procedure focus. In the laboratory test domain, LOINC supports the specification of the unit of measure (eg, mg/dL for phosphate) through recommended units, whereas SNOMED CT lacks this level of unit

specificity. Conversely, LOINC could not capture additional contextual information, such as “(em)” denoting an emergency procedure, whereas SNOMED CT could via post-coordination. In the medication domain, RxNorm generally represented detailed medication information most completely; for example, it captured ingredient, dose form, strength, and brand name for “Avastin 400 mg injection (bevacizumab).” In contrast, SNOMED CT represented ingredient, dose form, and strength but typically did not include brand name and often expressed strength per unit (eg, 25 mg per 1 mL) rather than total dose. The ATC classification included only ingredient-level information. Although RxNorm was usually the most complete, exceptions existed; for example, for “Mitomycin-C Kyowa,” SNOMED CT provided more complete information than RxNorm.

Table 3. Comparison of mapping results according to the target terminology system.

Domain: clinical source terms	Target
Chief complaint: For transplantation work-up	
KCD ^a	Z04.8 Examination and observation for other specified reasons
SNOMED CT ^b	
(1:1 Exact)	76464004 Hospital admission, for observation (procedure) : { 363702006 Has focus (attribute) = 1303222006 Pretransplant evaluation of recipient (procedure) }
Diagnosis: Adenocarcinoma of sigmoid colon	
KCD ^c	C18.7 Malignant neoplasm of sigmoid colon
SNOMED CT	301756000 Adenocarcinoma of sigmoid colon (disorder)
(1:1 Exact)	
Laboratory test: Phosphorus (em)	
LOINC ^d	2777-1 Phosphate [Mass/volume] in Serum or Plasmal
SNOMED CT	8364005 Phosphorus measurement (procedure) : { 116686009 Has specimen (attribute) = 119364003 Serum specimen (specimen) , 260870009 Priority (attribute) = 25876001 Emergency (qualifier value) }
(1:1 Exact)	
Surgery/ procedure: Excision of retroperitoneal tumor (malignant)	
ICD-9-CM ^e	54.4 Excision or destruction of peritoneal tissue
SNOMED CT	8951004 Excision of lesion of retroperitoneum (procedure) : { 363700003 Direct morphology (attribute) = 1240414004 Malignant neoplasm (morphologic abnormality) }
(1:1 Exact)	
Medication:	
Avastin 400 mg inj	
(bevacizumab)	
RxNorm	1657074 16 ML bevacizumab 25 MG/ML Injection [Avastin]
ATC ^f	L01FG01 Bevacizumab
SNOMED CT	781788003 Product containing precisely bevacizumab 25 milligram/1 milliliter conventional release solution for injection (clinical drug)
(1:1 Exact)	
Medication:	
Mitomycin-c kyowa	
10 mg inj (mitomycin)	
RxNorm	1165372 Mitomycin Injectable Product
ATC	L01DC03 Mitomycin
SNOMED CT	326815001 Product containing precisely mitomycin 10 milligram/1 vial powder for conventional release solution for injection (clinical drug)
(1:1 Exact)	

^aKCD: Korean Standard Classification of Diseases.^bSNOMED CT: Systematized Nomenclature of Medicine Clinical Terms.^cKCD: Korean Standard Classification of Diseases.^dLOINC: Logical Observation Identifiers Names and Codes.^eICD-9-CM: *International Classification of Diseases, 9th Revision, Clinical Modification*.^fATC: Anatomical Therapeutic Chemical.

Cross-Verification and Consensus Resolution

The initial mapping performed by the assigned mapper was additionally reviewed by 2 peer mappers, who recorded their agreement or disagreement with each mapping result. A mapping was considered discordant if either reviewer disagreed with the initial mapping result, and such cases were referred to the multidisciplinary working group for consensus resolution. There were no major discrepancies in the chief complaint, imaging test, or medication domains; the remaining domains showed cross-verification discordance rates of 0.3% to 1.3%. In the diagnosis domain,

the cross-verification discordance rate was 0.3%, primarily involving source terms containing more than 1 anatomical site. This was resolved by adding a rule specifying that if multiple sites in a source term belong to the same parent concept, the term should be mapped to that parent concept rather than mapped 1:N to each child concept. For example, “Neoplasm of uncertain or unknown behavior of liver, gallbladder, and bile ducts” should be mapped to “liver and/or biliary passages,” not separately to “liver,” “gallbladder,” and “bile duct.” The surgery/procedure domain showed a 1.0% cross-verification discordance rate, largely from inconsistent handling of parenthetical details; this was addressed by

mapping to target terms that incorporate the parenthetical meaning whenever possible (eg, retaining “malignant” in targets for “excision of retroperitoneal tumor (malignant)”). A 1.3% cross-verification discordance rate in the radiation therapy domain was mostly attributable to limited familiarity with the SNOMED CT relationship groups. Even when the meaning of a source term is not entirely reflected in the fully specified name of a focus concept, relationship group attributes may convey the necessary details, negating additional postcoordination. For example, “external irradiation–high energy radiotherapy” should map to the SNOMED CT concept “high energy beam therapy,” which includes a device attribute in its relationship group that represents external beam delivery, eliminating the need to encode “external irradiation” and “high energy.” Through this process of cross-verification and consensus discussion within the multidisciplinary working group, final mappings were confirmed; as examples, the mapping results for the 20% most frequent terms per each clinical domain are provided in [Multimedia Appendix 2](#).

Mapping Rates by Standardized Terminology System

Overall mapping rates were high across most domain-system pairs, with the chief complaint domain showing the lowest rates at 97.7% for KCD and 98.3% for SNOMED CT (Supplementary Material S5 in [Multimedia Appendix 1](#)). The narrow/not-mapped terms in the chief complaint domain were primarily attributable to ambiguity or insufficient specificity in the source terms, such as “oozing,” “prolapse,” “malfunction,” and “for admission,” for which a semantically appropriate SNOMED CT target could not be reliably selected. In the diagnosis domain, KCD achieved a 100% mapping rate, whereas SNOMED CT achieved 99.9% because a small number of source terms that could only

be narrowly mapped (eg, malignant neoplasm of overlapping lesion of other and ill-defined sites) were treated as narrow/not mapped. Outside these two domains, all source terms were successfully mapped to SNOMED CT. In the medication domain, 1.2% of the terms were not mapped to RxNorm owing to the absence of codes for drugs approved only in specific countries (eg, TS-1). All source terms were successfully mapped to other systems (LOINC, *ICD-9-CM* Volume 3, and ATC) within their respective domains.

The detailed SNOMED CT mapping results are presented in [Table 4](#). Among the exact mappings, precoordination predominated in chief complaint (64.7%), diagnosis (79.1%), imaging test (58.9%), surgery/procedure (63.1%), medication (90.1%), and radiation therapy (70.7%), whereas postcoordination was more frequent in laboratory tests (80.2%). In laboratory tests, postcoordination captures additional information such as quantitative/qualitative details and test methods beyond the focus concept. Broad mappings occurred most frequently in the imaging test (43.3%) and radiation therapy domains (40.7%). In the imaging test domain, broad mappings arose mainly because information on view and projection (eg, pelvic inlet and outlet), patient position (eg, decubitus, supine, and erect), and other acquisition techniques (eg, 3D reconstruction and low-dose protocols) could not be represented in the target concepts. In the radiation therapy domain, source terms in the treatment prescription subdomain showed a high broad mapping rate (61.3%), while source terms in the treatment site subdomain achieved a 100% exact mapping rate. Specifically, most broad mappings resulted from incomplete representation of the number of ports (eg, 3 ports) and field arrangement (eg, parallel-opposed 2 fields), whereas information on the preparation processes (eg, simulation and treatment planning), modalities (eg, tomotherapy), and radiation types (eg, proton therapy) was precisely represented in the target concepts.

Table 4. Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) mapping results^a.

Mapping type ^b	Chief complaint (%)	Diagnosis (%)	Laboratory test (%)	Imaging test (%)	Surgery/procedure (%)	Medication (%)	Radiation therapy (%) ^c
Exact	97.1	97.8	98.3	56.7	89.3	100	59.3
PRE	64.7	79.1	19.8	58.9	63.1	90.1	70.7
POST	34.1	20.2	75.9	39.7	30.6	9.9	29.3
PRE + POST	1.2	0.7	4.3	1.4	6.2	0.0	0.0
Broad ^d	1.2	2.2	1.7	43.3	10.7	0.0	40.7
PRE	0.0	91.2	0.0	89.6	63.8	0.0	83.8
POST	100	8.8	100	10.4	30.4	0.0	16.2
PRE + POST	0.0	0.0	0.0	0.0	5.8	0.0	0.0
Narrow/not-mapped ^e	1.7	0.1	0.0	0.0	0.0	0.0	0.0

^aPRE, pre-coordinated concept (when an existing SNOMED CT concept fully represented the source term); POST, post-coordinated expression (when a single pre-coordinated concept did not fully capture clinically meaningful details).

^bAbsence of information outside the scope of the SNOMED CT International Edition (eg, brand names, reimbursement-related information) was not classified as a broad mapping.

^cBy subdomain, mapping results were as follows: Treatment prescription, Exact 38.7% (PRE 32.6%, POST 67.4%), Broad 61.3% (PRE 83.8%, POST 16.2%), and Narrow/not-mapped 0.0%; Treatment site, Exact 100% (PRE 100%, POST 0.0%), Broad 0.0%, and Narrow/not-mapped 0.0%.

^dExamples of broad mapping for each clinical domain are provided in Supplementary Material S6 in [Multimedia Appendix 1](#).

^eAll narrow/not-mapped cases with detailed rationale are summarized in Supplementary Material S7 in [Multimedia Appendix 1](#). These cases were attributable to ambiguity or insufficient specificity of the source terms, which prevented reliable selection of a semantically appropriate SNOMED CT target, rather than to the absence of relevant concepts in SNOMED CT.

Discussion

We mapped 3472 clinical source terms from real-world records of patients with 6 cancers to multiple standardized terminology systems. Several studies have mapped clinical terms to standardized systems; however, few have documented detailed, real-world mapping experiences. For example, a South Korean study introduced a 9-step process to map local EMR terms to SNOMED CT [4], and a Hungarian group mapped 4661 national procedural codes from 2 university hospitals to SNOMED CT [24]. Although these studies covered a wide range of clinical terms and offered valuable insights, they limited their source terms to specific domains (eg, procedures), restricted their targets to SNOMED CT, and reported ~60% exact mapping rates because postcoordination was not used. In contrast, we focused on patients with cancer, whose care is particularly complex, to better understand the standardization of clinical terms in oncology, where multicenter research demand is high. We included terms across the continuum of cancer care and mapped them not only to SNOMED CT but also to other standardized systems, depending on the clinical domain. We also conducted exhaustive cross-verification for all mappings, whereas previous studies omitted or only partially applied cross-verification.

In our study, SNOMED CT achieved high overall mapping rates, ranging from 98.3% to 100% across all domains, confirming its broad applicability as one of the most comprehensive standardized terminology systems. Applying postcoordination as needed substantially improved exact semantic match rates, which exceeded 90% in most domains. Prior studies have noted lower exact mapping rates in certain domains [25–28]. For instance, Sung [25] et al reported that only 1574/4,111 (38.2%) EDI codes for NHI claims mapped in the laboratory test domain, and Kang et al [28] found a 398/2500 (31.9%) match for therapeutic/surgical procedures; Jung [27] et al showed 11,453/22,610 (50.6%) for medications. However, these rates reflect a reliance on precoordinated concepts. After applying postcoordination, Sung et al [25] and Kang et al [28] improved to 90.4% and 63.6%, respectively. Similarly, our precoordinated match in laboratory tests was 19.8%, rising to 98.3% with postcoordination. SNOMED CT postcoordination also captured details not represented by LOINC attributes, underscoring its role in fully representing the original clinical data and enabling precise use across domains [29,30].

However, even when postcoordination was used, exact semantic match rates remained relatively low in the imaging test and radiation therapy domains, with broad mappings accounting for more than 40% of source terms in each domain. In the imaging test domain, imaging adjustments related to view or projection, patient position, and

technique were not permitted in postcoordination, although such information was considered clinically important for the correct acquisition and interpretation of images [31]. Similarly, in the radiation therapy domain, the number of ports and field arrangement could not be represented in SNOMED CT International Edition, despite their relevance to treatment configuration and dosimetric evaluation. Although these limitations reflect the scope and modeling principle of the SNOMED CT International Edition, which does not include certain protocol- or configuration-level details to maintain standardized representation of clinical procedures, our findings suggest that this approach may reduce the semantic fidelity of complex imaging and radiation therapy data that are crucial to cancer care and research. To improve semantic representation, frequently used details could be represented through SNOMED CT national extensions and organized through implementation-specific reference sets. In addition, acquisition parameters and radiation therapy protocol details could be captured using complementary domain-specific terminologies and standards such as Radiology Lexicon and Digital Imaging and Communications in Medicine, including its radiotherapy extension (Digital Imaging and Communications in Medicine – Radiation Therapy) [32–34]. Such a combined approach may help preserve detailed procedural information and improve semantic interoperability in real-world oncology data.

Despite the breadth of SNOMED CT, the complementary use of domain-specific systems is necessary to capture the full characteristics of source terms. For example, although SNOMED CT aligns with the Identification of Medicinal Products, an international standard for dose forms and clinically appropriate strengths, it has limitations with respect to brand names, packaging, and country-specific details (eg, drug codes and reimbursement/insurance status) [35–37]. In contrast, RxNorm represents drug information in greater detail, including multi-ingredient combinations, brand names, exact strengths, and national drug codes. In our study, with precoordination alone, coverage for ingredients, dose form, and strength was same for SNOMED CT (90.1%) and RxNorm (90.1%), but RxNorm more precisely represented the dosage by explicitly encoding the total dose. Although postcoordination can improve SNOMED CT coverage, encoding dosage through postcoordination increases structural complexity, which is a practical limitation. Such country-specific details can also be addressed through a national extension of SNOMED CT, which allows member countries to add concepts and reference sets to meet national requirements while retaining interoperability with the international core [38]. In the United Kingdom, for example, the NHS dictionary of medicines and devices, implemented as the SNOMED CT UK Drug Extension, represents brand names, packaging, and reimbursement-related information

using SNOMED CT-compatible identifiers [39]. As South Korea joined SNOMED International in 2020 and is currently developing a national extension, such administrative or policy-related details could also be represented in addition to internationally standardized clinical concepts.

For laboratory tests, LOINC provides a more specialized representation than SNOMED CT via its six-axis model, namely component, property, time, system, scale, and method [40,41]. In our study, precoordinated SNOMED CT concepts alone accounted for only 19.8% of exact SNOMED CT mappings in the laboratory test domain, whereas the remaining 80.2% of exact mappings required postcoordination. Although SNOMED CT postcoordination enabled high mapping rates and captured additional contexts not represented by LOINC (eg, urgent testing status), this limited precoordinated coverage explains the widespread adoption of LOINC-based coding system for laboratory test data in many institutions. These findings support the complementary use of domain-specific terminology systems alongside SNOMED CT. Beyond using these systems in parallel, the recently released LOINC Ontology, a LOINC and SNOMED CT interoperability solution jointly developed by Regenstrief Institute and SNOMED International, may offer a more integrated approach [42,43]. By representing LOINC content within a SNOMED CT concept model, it is expected to enhance LOINC's computability and improve interoperability between LOINC-coded laboratory observations and broader SNOMED CT-based clinical concepts. Such integrated approaches, rather than the independent use of separate domain-specific terminology systems, may further enhance the semantic interoperability and reusability of real-world clinical data.

To broaden the adoption of SNOMED CT, whose strengths are detailed semantics and wide applicability, the complexity of postcoordination must be addressed [44]. Postcoordination was necessary for 9.9% to 80.2% of the exact mappings in our study. However, SNOMED CT's flexibility, supporting various precoordinated and postcoordinated expressions, can yield inconsistent code usage across mappers, challenging standardization, because multiple concepts with similar meanings may exist and several valid combinations of focus concepts, attributes, and attribute values may exist [30,45]. In this study, the use of predefined mapping priorities in addition to SNOMED CT Editorial Guide and MRCM supported high consistency in postcoordination decisions (cross-verification discordance rate, 0.3%-1.3%). To maintain consistency without sacrificing the flexibility of SNOMED CT, it is necessary to incorporate structured mapper training in both clinical domain knowledge and terminology systems, predefined priority rules on selecting focus concepts and attributes, standardized postcoordination templates or reusable expression libraries for recurrent source-term patterns, and cross-verification with regular consensus review. In addition, SNOMED CT content should be periodically updated through the review of frequently used postcoordinated expressions in routine clinical practice, which may be considered as candidates for precoordinated concepts or national extensions.

This study has the limitation in that mappings were developed at a single institution without multicenter validation. Because standard terminology mapping can be influenced by institution-specific workflows, local code management practices, and information system environments, whether the present results can be applied identically to other institutions remains uncertain. Nevertheless, interinstitutional heterogeneity in South Korea might be relatively less pronounced than in health care systems with more fragmented insurance and coding structures. South Korea operates a single-payer NHI system, and medical services are documented and managed using national-level classification systems. In addition, recent government-led initiatives have continued to promote the standardization of health data across institutions, through the establishment of multicenter common clinical data infrastructures (eg, the Data-Driven Hospital Project and the Korea-Clinical Data Utilization Network for Research Excellence), as well as national standards for health data elements terminology (eg, the Korea Core Data for Interoperability) [46,47]. Accordingly, source terms linked to nationally standardized coding systems, such as KCD, LOINC, and EDI claims codes for NHI, are less likely to show large interinstitutional variation and may therefore be more consistently applicable across other institutions. In contrast, locally maintained terms, such as institution-specific names for procedures, surgeries, laboratory tests, and imaging studies, or free-text entries in chief complaints, may yield different mapping results across institutions. Further multicenter validation should focus on these locally maintained terms to determine whether our mapping results remain reproducible across institutions and to identify which source-term categories require local adaptation before broader implementation.

Another limitation concerns the choice of procedure classification system. *ICD-9-CM* Volume 3 was selected for the surgery/procedure domain, considering the Korean clinical coding context in which *ICD-9-CM*-based procedure classification continues to serve as a reference classification for national health statistics [15]. However, this may limit the international applicability of the mapped data. Procedure and intervention classification remains highly heterogeneous across countries. Although the WHO introduced International Classification of Procedures in Medicine in 1978, it was not subsequently maintained, and many countries developed their own procedure classification systems with different structures and purposes, limiting international comparability of intervention data [48-50]. In response to the need for more harmonized frameworks, the WHO has developed the International Classification of Health Interventions to provide a common structure to which codes in national classifications can be mapped for compiling international data [51]. Further studies will be needed to extend the mapping of local clinical terms beyond *ICD-9-CM* toward emerging global standards, thereby enhancing the international interoperability of real-world clinical data.

To the best of our knowledge, this is the first large-scale investigation to compare the strengths and limitations of multiple standardized terminology systems by

mapping real-world clinical terms across a full continuum of cancer care. Our study demonstrates the feasibility of applying standardized terminology systems to real-world oncology terms. SNOMED CT is the most comprehensive and universally applicable system, whereas domain-specific systems play essential roles in capturing more specialized aspects of clinical data. Clear mapping priorities in addition to the official guidance for each terminology system, together with cross-verification and consensus review, provided a

practical framework for resolving mapping disagreement while maintaining flexibility. Regular updates of standardized terminology systems to accommodate newly emerging clinical terms, together with ongoing mapper education, are important to maintain mapping consistency. Future work may enhance the interoperability of real-world cancer data and contribute to the standardization of research data for the development of clinical decision support systems, health care information exchange, and multicenter research facilitation.

Acknowledgments

Generative AI, OpenAI ChatGPT, was used for language refinement and editing of text originally drafted by the authors. It was not used to analyze data, interpret results, generate figures, or collect references. The authors reviewed and revised all AI-assisted suggestions and take full responsibility for the final manuscript.

Funding

This work was supported by the National Cancer Center Korea [2310850-3]. The funding source had no role in the study design, data collection, data analysis, interpretation of results, manuscript preparation, or the decision to submit the manuscript for publication.

Data Availability

The dataset used in this research contains original clinical information extracted from an electronic medical record system and, therefore, cannot be made publicly available. Deidentified data may be shared by the corresponding author upon reasonable request for academic research purposes.

Authors' Contributions

Conceptualization: KSC, SL, HSC

Data curation: SL, EJ, YJ, IHK

Investigation: BJ, EP, SL, EJ, YJ

Resources: KSC, HJ Kong, HSC, HJ Kim

Writing – original draft: BJ, EP, SL, EJ, YJ

Writing – review & editing: HC

Conflicts of Interests

None declared.

Multimedia Appendix 1

Supplementary Material S1. Examples of source term refinement prior to mapping Supplementary Material S2. Number of source terms for mapping by clinical domain and cancer type Supplementary Material S3. Examples illustrating the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) postcoordination process Supplementary Material S4. Representative mapping results for the most frequent 20% of clinical terms in each domain to standard terminologies Supplementary Material S5. Mapping rates of source terms to each standard terminology system by clinical domain Supplementary Material S6. Representative examples of broad mapping in SNOMED CT for each clinical domain Supplementary Material S7. Complete list of not-mapped cases in SNOMED CT mapping and their rationale.

[[DOCX File \(Microsoft Word File\)](#), 443 KB-[Multimedia Appendix 1](#)]

Multimedia Appendix 2

Representative mapping results for the most frequent 20% of clinical terms.

[[XLSX File \(Microsoft Excel File\)](#), 146 KB-[Multimedia Appendix 2](#)]

References

1. Matengo W, Otsieno E, Wanjiru K. Big data analytics in healthcare. In: Chaari L, editor. Digital Health in Focus of Predictive, Preventive and Personalised Medicine. Springer; 2020:129-138. [doi: [10.1007/978-3-030-49815-3_15](https://doi.org/10.1007/978-3-030-49815-3_15)]
2. Amar F, April A, Abran A. Electronic health record and semantic issues using Fast Healthcare Interoperability Resources: systematic mapping review. J Med Internet Res. Jan 30, 2024;26:e45209. [doi: [10.2196/45209](https://doi.org/10.2196/45209)] [Medline: [38289660](https://pubmed.ncbi.nlm.nih.gov/38289660/)]
3. Healthcare Information and Management Systems Society (HIMSS). 2020. URL: <https://legacy.himss.org/resources/interoperability-healthcare> [Accessed 2026-08-02]

4. Sung S, Park HA, Jung H, Kang H. A SNOMED CT mapping guideline for the local terms used to document clinical findings and procedures in electronic medical records in South Korea: methodological study. *JMIR Med Inform*. Apr 18, 2023;11:e46127. [doi: [10.2196/46127](https://doi.org/10.2196/46127)] [Medline: [37071456](https://pubmed.ncbi.nlm.nih.gov/37071456/)]
5. Fritz AG, editor. *International Classification of Diseases for Oncology: ICD-O*. 3rd ed. World Health Organization; 2000. ISBN: 978-92-4-154534-1
6. Averill RF, Mullin RL, Steinbeck BA, Goldfield NI, Grant TM. Development of the ICD-10 procedure coding system (ICD-10-PCS). *Top Health Inf Manage*. Feb 2001;21(3):54-88. [Medline: [11234732](https://pubmed.ncbi.nlm.nih.gov/11234732/)]
7. McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem*. Apr 2003;49(4):624-633. [doi: [10.1373/49.4.624](https://doi.org/10.1373/49.4.624)] [Medline: [12651816](https://pubmed.ncbi.nlm.nih.gov/12651816/)]
8. Nelson SJ, Zeng K, Kilbourne J, Powell T, Moore R. Normalized names for clinical drugs: RxNorm at 6 years. *J Am Med Inform Assoc*. 2011;18(4):441-448. [doi: [10.1136/amiajnl-2011-000116](https://doi.org/10.1136/amiajnl-2011-000116)] [Medline: [21515544](https://pubmed.ncbi.nlm.nih.gov/21515544/)]
9. Stearns MQ, Price C, Spackman KA, Wang AY. SNOMED clinical terms: overview of the development process and project status. *Proc AMIA Symp*. 2001:662-666. [Medline: [11825268](https://pubmed.ncbi.nlm.nih.gov/11825268/)]
10. Zafirah SA, Nur AM, Puteh SEW, Aljunid SM. Potential loss of revenue due to errors in clinical coding during the implementation of the Malaysia diagnosis related group (MY-DRG®) Casemix system in a teaching hospital in Malaysia. *BMC Health Serv Res*. Jan 25, 2018;18(1):38. [doi: [10.1186/s12913-018-2843-1](https://doi.org/10.1186/s12913-018-2843-1)] [Medline: [29370785](https://pubmed.ncbi.nlm.nih.gov/29370785/)]
11. Opitasari C, Nurwahyuni A. The completeness and accuracy of clinical coding for diagnosis and medical procedure on the INA-CBGs claim amounts at a hospital in South Jakarta. *Health Sci J Indones*. Oct 19, 2018;9(1):14-18. [doi: [10.22435/hsji.v9i1.464](https://doi.org/10.22435/hsji.v9i1.464)]
12. Chan KKP, Ng TCC, Sze CY, et al. Validation of diagnosis codes for pleural diseases and procedure codes for relevant respiratory procedures in a healthcare database in Hong Kong: a single tertiary centre study. *Hong Kong Med J*. Feb 2026;32(1):13-22. [doi: [10.12809/hkmj2412275](https://doi.org/10.12809/hkmj2412275)] [Medline: [41612789](https://pubmed.ncbi.nlm.nih.gov/41612789/)]
13. Yu L, Lang J. Diagnosis-related Groups (DRG) pricing and payment policy in China: where are we? *Hepatobiliary Surg Nutr*. Dec 2020;9(6):771-773. [doi: [10.21037/hbsn-2020-8](https://doi.org/10.21037/hbsn-2020-8)] [Medline: [33299831](https://pubmed.ncbi.nlm.nih.gov/33299831/)]
14. Giordani B, Burgio A, Grippo F, Barone A, Eugeni E, Baglio G. The use of ICD-9-CM coding to identify COVID-19 diagnoses and determine risk factors for 30-day death rate in hospitalized patients in Italy: retrospective study. *JMIR Public Health Surveill*. Feb 23, 2024;10:e44062. [doi: [10.2196/44062](https://doi.org/10.2196/44062)] [Medline: [38393763](https://pubmed.ncbi.nlm.nih.gov/38393763/)]
15. Korean Health Information Management Association. Study on development of Korean classification of health interventions, year 7 [Report in Korean]. Statistics Korea; 2023. URL: <https://mlib1.kostat.go.kr/search/detail/CATSUZ00000099128?location=A0000001> [Accessed 2026-08-02]
16. Lee YK, Hong SO, Park SJ, et al. Data resource profile: the Korea National Hospital Discharge In-depth Injury Survey. *Epidemiol Health*. 2021;43:e2021052. [doi: [10.4178/epih.e2021052](https://doi.org/10.4178/epih.e2021052)] [Medline: [34412448](https://pubmed.ncbi.nlm.nih.gov/34412448/)]
17. SNOMED CT editorial guide. SNOMED International. 2025. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-editorial-guide> [Accessed 2026-08-02]
18. Compositional grammar specification. SNOMED International. 2025. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-compositional-grammar-specification> [Accessed 2026-08-02]
19. SNOMED CT Machine Readable Concept Model. SNOMED International. 2026. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-machine-readable-concept-model> [Accessed 2026-08-02]
20. Medicinal product. SNOMED International. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-editorial-guide/readme/authoring/domain-specific-modeling/pharmaceutical-and-biologic-product/medicinal-product> [Accessed 2026-08-02]
21. Out of scope. SNOMED International. 2025. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-editorial-guide/readme/snomed-ct-introduction/out-of-scope> [Accessed 2026-08-02]
22. Procedure modeling. SNOMED International. 2026. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-editorial-guide/readme/authoring/domain-specific-modeling/procedure/index-2> [Accessed 2026-08-02]
23. Korean Standard Terminology of Medicine (KOSTOM), version 7. Ministry of Health and Welfare, Republic of Korea; URL: <https://www.hins.or.kr> [Accessed 2026-10-12]
24. Mészáros Á, Kovács S, Héja T, Bagyura Z, Zemplényi A. Mapping Hungarian procedure codes to SNOMED CT. *BMC Med Res Methodol*. Oct 18, 2023;23(1):240. [doi: [10.1186/s12874-023-02036-x](https://doi.org/10.1186/s12874-023-02036-x)] [Medline: [37853326](https://pubmed.ncbi.nlm.nih.gov/37853326/)]
25. Sung S, Park HA, Park SK, Jung H, Kang H, Lee MS. Mapping Korean National Health Insurance claim codes for laboratory test to SNOMED CT. *Stud Health Technol Inform*. May 18, 2023;302:78-82. [doi: [10.3233/SHTI230068](https://doi.org/10.3233/SHTI230068)] [Medline: [37203613](https://pubmed.ncbi.nlm.nih.gov/37203613/)]
26. Hwang EJ, Park HA, Sohn SK, et al. Mapping Korean EDI medical procedure code to SNOMED CT. *Stud Health Technol Inform*. Aug 21, 2019;264:178-182. [doi: [10.3233/SHTI190207](https://doi.org/10.3233/SHTI190207)] [Medline: [31437909](https://pubmed.ncbi.nlm.nih.gov/31437909/)]

27. Jung H, Park HA, Kang H, Lee M, Sung S, Park S. Mapping Korean National Health Insurance pharmaceutical claim codes to SNOMED CT. *Stud Health Technol Inform.* May 25, 2022;294:297-301. [doi: [10.3233/SHTI220462](https://doi.org/10.3233/SHTI220462)] [Medline: [35612080](https://pubmed.ncbi.nlm.nih.gov/35612080/)]
28. Kang H, Park HA. Mapping Korean National Health Insurance reimbursement claim codes for therapeutic and surgical procedures to SNOMED-CT to facilitate data reuse. *Stud Health Technol Inform.* Jun 6, 2022;290:101-105. [doi: [10.3233/SHTI220040](https://doi.org/10.3233/SHTI220040)] [Medline: [35672979](https://pubmed.ncbi.nlm.nih.gov/35672979/)]
29. Richesson RL, Andrews JE, Krischer JP. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. *J Am Med Inform Assoc.* 2006;13(5):536-546. [doi: [10.1197/jamia.M2093](https://doi.org/10.1197/jamia.M2093)] [Medline: [16799121](https://pubmed.ncbi.nlm.nih.gov/16799121/)]
30. Campbell JR, Xu J, Fung KW. Can SNOMED CT fulfill the vision of a compositional terminology? Analyzing the use case for problem list. *AMIA Annu Symp Proc.* 2011;2011:181-188. [Medline: [22195069](https://pubmed.ncbi.nlm.nih.gov/22195069/)]
31. Imaging adjustments for view, projection, or technique. SNOMED International. 2025. URL: <https://docs.snomed.org/snomed-ct-specifications/snomed-ct-editorial-guide/readme/authoring/domain-specific-modeling/procedure/index-1/index/imaging-adjustments-for-view-projection-or-technique> [Accessed 2026-08-02]
32. Marquis M, Bossenko I, Ross P. RadLex and SNOMED CT integration: a pilot study for standardising radiology classification. *Insights Imaging.* Mar 13, 2025;16(1):58. [doi: [10.1186/s13244-025-01935-5](https://doi.org/10.1186/s13244-025-01935-5)] [Medline: [40080225](https://pubmed.ncbi.nlm.nih.gov/40080225/)]
33. Kahn CE, Langlotz CP, Channin DS, Rubin DL. Informatics in radiology: an information model of the DICOM standard. *Radiographics.* 2011;31(1):295-304. [doi: [10.1148/rg.311105085](https://doi.org/10.1148/rg.311105085)] [Medline: [20980665](https://pubmed.ncbi.nlm.nih.gov/20980665/)]
34. Law MYY, Liu B. Informatics in radiology: DICOM-RT and its utilization in radiation therapy. *Radiographics.* 2009;29(3):655-667. [doi: [10.1148/rg.293075172](https://doi.org/10.1148/rg.293075172)] [Medline: [19270073](https://pubmed.ncbi.nlm.nih.gov/19270073/)]
35. Bodenreider O, Cornet R, Vreeman DJ. Recent developments in clinical terminologies—SNOMED CT, LOINC, and RxNorm. *Yearb Med Inform.* Aug 2018;27(1):129-139. [doi: [10.1055/s-0038-1667077](https://doi.org/10.1055/s-0038-1667077)] [Medline: [30157516](https://pubmed.ncbi.nlm.nih.gov/30157516/)]
36. Bodenreider O, James J. The new SNOMED CT International Medicinal Product model. *CEUR Workshop Proc.* Dec 2018;2285:36. [Medline: [36277122](https://pubmed.ncbi.nlm.nih.gov/36277122/)]
37. Noel Nikiema J, Bodenreider O. Comparing the representation of medicinal products in RxNorm and SNOMED CT—consequences on interoperability. *CEUR Workshop Proc.* Aug 2019;2931:F1-F6. [Medline: [36276234](https://pubmed.ncbi.nlm.nih.gov/36276234/)]
38. Extension and customization. SNOMED International. 2025. URL: <https://docs.snomed.org/snomed-ct-practical-guides/snomed-ct-starter-guide/10-extension-and-customization> [Accessed 2026-08-02]
39. Dictionary of medicines and devices (dm+d). NHS England Digital. URL: <https://digital.nhs.uk/services/terminology-and-classifications/dm-d> [Accessed 2026-08-02]
40. LOINC users' guide. LOINC. URL: <https://loinc.org/kb/users-guide> [Accessed 2026-08-02]
41. Bodenreider O. Issues in mapping LOINC laboratory tests to SNOMED CT. *AMIA Annu Symp Proc.* Nov 6, 2008:51-55. [Medline: [18999311](https://pubmed.ncbi.nlm.nih.gov/18999311/)]
42. Implementation guide for the LOINC ontology. SNOMED International. 2026. URL: <https://docs.snomed.org/implementation-guides/loinc-implementation-guide> [Accessed 2026-08-02]
43. Regenstrief Institute & SNOMED International. URL: <https://loincsnomed.org/> [Accessed 2026-08-02]
44. Benson T. Principles of Health Interoperability HL7 and SNOMED. 2nd ed. Springer; 2012. ISBN: 9781447128007
45. Miñarro-Giménez JA, Martínez-Costa C, Karlsson D, Schulz S, Gøeg KR. Qualitative analysis of manual annotations of clinical text with SNOMED CT. *PLoS One.* 2018;13(12):e0209547. [doi: [10.1371/journal.pone.0209547](https://doi.org/10.1371/journal.pone.0209547)] [Medline: [30589855](https://pubmed.ncbi.nlm.nih.gov/30589855/)]
46. Lee S, Choi YH, Kim HM, et al. The Cancer Clinical Library Database (CCLD) from the Korea-Clinical Data Utilization Network for Research Excellence (K-CURE) project. *Cancer Res Treat.* Jan 2025;57(1):19-27. [doi: [10.4143/crt.2024.218](https://doi.org/10.4143/crt.2024.218)] [Medline: [39010797](https://pubmed.ncbi.nlm.nih.gov/39010797/)]
47. Kwon A, Lee HY, Shin SY, et al. Current health data standardization project and future directions to ensure interoperability in Korea. *Healthc Inform Res.* Apr 2024;30(2):93-102. [doi: [10.4258/hir.2024.30.2.93](https://doi.org/10.4258/hir.2024.30.2.93)] [Medline: [38755100](https://pubmed.ncbi.nlm.nih.gov/38755100/)]
48. International Classification of Procedures in Medicine. World Health Organization; 1978. ISBN: 9789241541244
49. Paviot BT, Madden R, Moskal L, et al. Development of a new international classification of health interventions based on an ontology framework. *Stud Health Technol Inform.* 2011;169:749-753. [Medline: [21893847](https://pubmed.ncbi.nlm.nih.gov/21893847/)]
50. Mellinghoff SC, Bruns C, Al-Monajjed R, et al. Harmonized procedure coding system for surgical procedures and analysis of surgical site infections (SSI) of five European countries. *BMC Med Res Methodol.* Aug 12, 2022;22(1):225. [doi: [10.1186/s12874-022-01702-w](https://doi.org/10.1186/s12874-022-01702-w)] [Medline: [35962320](https://pubmed.ncbi.nlm.nih.gov/35962320/)]
51. WHO-FIC maintenance platform. International Classification of Health Interventions (ICHI). URL: <https://icd.who.int/dev11/l-ichi/en> [Accessed 2026-08-02]

Abbreviations

ATC: Anatomical Therapeutic Chemical
CRDW: Clinical Research Data Warehouse
EDI: electronic data interchange
EMR: electronic medical record
ICD: *International Classification of Diseases*
ICD-10-PCS: *International Classification of Diseases, 10th Revision, Procedure Coding System*
ICD-9-CM: *International Classification of Diseases, 9th Revision, Clinical Modification*
ICD-O: *International Classification of Diseases for Oncology*
KCD: Korean Standard Classification of Diseases
LOINC: Logical Observation Identifiers Names and Codes
MRCM: Machine Readable Concept Model
NCCK: National Cancer Center Korea
NHI: National Health Insurance
SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms
WHO: World Health Organization

Edited by Arriel Benis, Ivan Steenstra; peer-reviewed by Baby Satravada, Karen Triep, Silke D Werle; submitted 02.Apr.2026; final revised version received 15.Sep.2026; accepted 16.Sep.2026; published 27.Oct.2026

Please cite as:

Jang B, Park E, Lee S, Ju E, Jang Y, Kwak IH, Choi KS, Kong HJ, Cha HS, Kim HJ, Chae H
Real-World Applicability of Multiple International Terminology Systems in Oncology: Mapping Study of Clinical Terms for Cancer Diagnosis and Treatment
JMIR Med Inform 2026;14:e96548
URL: <https://medinform.jmir.org/2026/1/e96548>
doi: [10.2196/96548](https://doi.org/10.2196/96548)

© Boyun Jang, Eunhye Park, Sangwon Lee, Eunbi Ju, Yewon Jang, In Hae Kwak, Kui Son Choi, Hyun-Joo Kong, Hyo Soung Cha, Hyun-Jin Kim, Heejung Chae. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.