Journal of the American Medical Informatics Association, 2026, 1–13
https://doi.org/10.1093/jamia/ocag094
Review

Use of SNOMED CT, 2020-2025: literature review

Eunsuk Chang, MD, MPH, PhD[1,*] and Javed Mostafa, PhD[2]

[1]Republic of Korea Air Force Aerospace Medicine Research Center, 6 Ssangsugwangi-gil, Sangdang-gu, Cheongju-si, Chungcheongbuk-do, 28187, Republic of Korea
[2]Faculty of Information, University of Toronto, 140 St George St, Toronto, ON, M5S 3G6, Canada
*Corresponding Author: Eunsuk Chang, MD, MPH, PhD, Republic of Korea Air Force Aerospace Medicine Research Center, 6 Ssangsugwangi-gil, Sangdang-gu, Cheongju-si, Chungcheongbuk-do, 28187, Republic of Korea (chang.eunsuk.dr@gmail.com)

Abstract

Objective: **To examine the evolving role and application of SNOMED CT (SCT) during 2020-2025, a period marked by the COVID-19 pandemic and accelerated adoption of artificial intelligence (AI) in healthcare.**

Materials and methods: **We searched PubMed and Embase for articles published from October 2020 to June 2025. Included articles were classified into focus categories by implementation maturity (Theoretical, Predevelopment, Implementation, Evaluation/Commodity, and Non-operational) and usage categories. We compared trends with our previous review covering January 2015-September 2020.**

Results: **Following exclusion criteria, 651 articles were included for final review. The United States ($n = 188$) and United Kingdom ($n = 92$) were the largest contributors. COVID-19 emerged as the second most-investigated domain. The 2020-2025 period witnessed a dramatic shift toward mature implementation stages: Implementation (26.4%) and Evaluation/Commodity (32.0%) categories expanded, while Theoretical and Predevelopment categories decreased. SCT was increasingly used for knowledge graph construction, machine learning model validation, and patient data retrieval from registries. The most prominent use case involved retrieving patient data from national and commercial registries ($n = 186$).**

Discussion: **SCT's role evolved from a clinical terminology to a machine-interpretable biomedical knowledge base, supporting explainable AI. However, coding consistency remains challenging, and evidence demonstrating improved patient outcomes is lacking.**

Conclusion: **SCT use has matured significantly, with widespread implementation in data repositories and emerging applications in AI. Future research must demonstrate clinical and operational benefits to ensure continued adoption.**

Key words: SNOMED CT; systematized nomenclature of medicine; terminology as topic; vocabulary, controlled.

Introduction

The period from 2020 to 2025 marked a transformative era in healthcare data management, characterized by unprecedented demands for global health data sharing, accelerated adoption of artificial intelligence (AI) in clinical settings, and intensified efforts toward semantic interoperability across heterogeneous healthcare systems.[1] At the center of these developments stands SNOMED CT (SCT), the world's most comprehensive multilingual clinical healthcare terminology.[2] Since our previous review on the use of SCT,[3] several converging developments have transformed how SCT is leveraged in healthcare systems worldwide.

The early 2020s witnessed maturation in the integration of SCT with emerging health data standards and common data models (CDMs). The collaboration between Health Level 7 (HL7) International and Observational Health Data Sciences and Informatics created a unified framework to integrate Fast Healthcare Interoperability Resources (FHIR) with the Observational Medical Outcomes Partnership CDM.[4] SNOMED International's collaboration with HL7 International to use FHIR-based terminology services with its products represented an architectural evolution designed to support real-time interoperability needs.[5] Additionally, major national and international initiatives adopted SCT as a foundational standard. The development and release of the International Patient Summary Terminology, harmonized with HL7 standards and made freely available to support cross-border patient care, exemplified efforts to reduce barriers to global health data exchange.[6]

SCT also expanded its influence on the healthcare domain based on collaborations that aligned with other terminology systems in the early 2020s. The ongoing effort to develop a Logical Observation Identifiers Names and Codes Extension within

Received: December 30, 2025. **Revised:** May 5, 2026. **Accepted:** May 26, 2026

the SCT framework facilitated the use of both standards in a single system, removing redundant mappings that hinder effective healthcare information sharing.[7] In addition, the COVID-19 pandemic, which emerged in early 2020, created a global need for standardized clinical data collection, exchange, and analysis across international borders.[8] This pandemic-driven use case revealed both the strengths and limitations of existing clinical terminology infrastructures, prompting accelerated development of interoperability solutions and harmonization initiatives.[9]

SCT is no longer a peripheral vocabulary standard but the semantic backbone that determines whether health data can be meaningfully shared, queried, and reused across institutions and borders. Understanding how SCT is being adopted, where it succeeds, and where gaps persist is therefore essential knowledge for anyone shaping the future of biomedical informatics. For healthcare informaticians and standards developers, identifying emerging patterns in SCT implementation is essential for prioritizing future development efforts and avoiding duplicative work. For researchers, understanding which applications of SCT have matured versus which remain exploratory can guide methodological choices and highlight gaps in the evidence base. For health system leaders and implementers, knowledge of successful SCT use cases across different domains and regions can inform strategic planning and risk assessment for terminology adoption initiatives. For clinicians, awareness of how SCT encodes diagnoses, findings, and procedures across real-world care settings is increasingly relevant as SCT-coded data underpin the clinical decision support tools, patient registries, and AI-driven diagnostic aids that are becoming part of everyday practice.

This literature review addresses this knowledge gap by examining the evolving role and application of SCT during 2020-2025, a period that witnessed both unprecedented challenges and advances in healthcare informatics. Building on our previous review of 2015-2020 SCT usage,[3] this work provides actionable insights for the readership by: (1) quantifying evolving trends in the use of SCT by usage categories, domains, and regions during the 2020-2025 period compared to the 2015-2020 period (specifically, October 2020-June 2025 and January 2015-September 2020 respectively),[3] and (2) highlighting salient use cases by categories that emerged during the early 2020's, with particular attention to novel applications that represent potential areas for future innovation or standardization efforts.

Materials and methods

Study selection and screening

We followed the overall review methodology established in our previous review, which was based on the approach developed by Lee et al.[10] The same query terms were applied to PubMed and Embase searches, with publication dates restricted to "2020-09-25 to 2025-06-30" (Supplementary Appendix A, available as supplementary material *Journal of the American Medical Informatics Association* online). To comprehensively capture the broad spectrum of SCT applications, we employed lenient inclusion criteria, encompassing all papers retrieved by the queries unless they met the following exclusion criteria: (1) SCT was not mentioned in the text body; (2) the full-text was not written in English; or (3) the article was published only as an abstract. We intentionally excluded abstract-only papers due to their limited information about SCT usage. We also excluded articles that did not explicitly mention SCT use, even when SCT may have been used indirectly. For example, researchers may reference the All of Us Research Hub as a data source, where diseases are coded using SCT and other terminology systems, without explicitly mentioning SCT as a means of data retrieval. To gain comprehensive perspectives on SCT applications, we included non-peer-reviewed articles, such as preprints published on bioRxiv, provided they followed sound scientific methods and contained sufficient explanation of SCT usage.

Categorization by usage criteria

We applied the SCT use categorization framework developed in our previous review.[3] Included papers were classified into one of five mutually exclusive focus categories—"Theoretical," "Predevelopment," "Implementation," and "Evaluation/Commodity," arranged in order of implementation maturity, plus an "Non-operational" category (renamed from "Indeterminate" in the previous review). Definitions for each focus category are described in Table 1.[3] Within each focus category, we established subordinate usage categories based on our previous review framework, with modifications to reflect the current spectrum of SCT applications. Table 2 presents the operational definitions of each usage category, which evolved iteratively during the classification process to ensure comprehensive coverage of all included articles.

Each paper was assigned to only one usage category within a single parent focus category. When an article could potentially be classified into multiple usage categories, we selected the category representing the most mature implementation stage. For example, we classified a study by Jones et al.[11] into the

Table 1. Definitions of each focus category (adopted from Lee et al.[10]).

Focus category	Definition
Theoretical	Describes SNOMED CT in terms of a terminology system but not at the stage of implementation in clinical or operational environments
Predevelopment	Refers to SNOMED CT being assessed or evaluated as to whether it can be used in full-scale implementation as a terminology standard
Implementation	Refers to SNOMED CT being used in project or operational settings
Evaluation/ commodity	Evaluates SNOMED CT's effects or impacts on operational settings or demonstrates that its function has shifted from encoding data to using data recorded in SNOMED CT codes
Non-operational	SNOMED CT being used as an example of a terminology system without any further detail on its use or implementation, is referenced in a letter by a reader, editor or author, or is included in a survey or review.

Table 2. Operational definitions of usage categories by implementation maturity used for classifying included articles.

Focus Category	Usage Category	Definition
Theoretical	Terminology analysis	The validity of SNOMED CT content, including structural, lexical, and semantic components, was examined through manual analysis or automated methods.
Predevelopment	Translation	From SNOMED CT, the description component was translated into local languages to address regional needs.
	Prospective content coverage	SNOMED CT coverage of domain terms, data elements, other terminology systems, or biomedical texts was examined and measured. This category focuses on assessing coverage rather than creating mappings.
	Mapping or comparing to other terminology systems	SNOMED CT was mapped to or from interface terminologies or reference terminology systems, or new terminologies were created integrating SNOMED CT as a foundational model. Variable codes or data elements from clinical studies or systems were categorized under "Prospective content coverage" instead.
	Design considerations	Challenges in implementing SNOMED CT as a terminology system were addressed, with suggestions to improve its design and components for better health data standardization. These studies are process-oriented, whereas "Prospective content coverage" and "Compare or map to other terminology systems" focus on reporting results.
	Interoperability standards for health information systems	SNOMED CT was incorporated as a terminology standard to enhance interoperability in health systems, including electronic health record systems and common data models, before full real-world implementation. Studies showing SNOMED CT in fully implemented real-world systems were categorized under "Implementation of SNOMED CT."
Implementation	Encoding methodologies	Studies proposed novel methods for utilizing SNOMED CT's graph structure as a knowledge base. This category was newly established to address emerging needs for using SNOMED CT in knowledge representation. Papers using off-the-shelf encoding methods were classified under "Used for classifying or coding."
	Used for classifying or coding	SNOMED CT concepts were extracted from sources to normalize and standardize clinical ideas using existing off-the-shelf concept extraction tools. Unlike the studies in "Encoding methodologies," no novel encoding methods were proposed in the studies included here. Extraction was done primarily for research purposes; real-world applications were categorized under "Implementation of SNOMED CT."
	Implementation of SNOMED CT	SNOMED CT was implemented within operational systems to provide real-world users with services such as clinical decision support systems, data registries, or terminology servers.
Evaluation/ Commodity	Gold standard for model validation	SNOMED CT was used as a gold-standard benchmark to evaluate the performance of algorithms, models, frameworks, or other terminology systems.
	Assess code use and activity	Investigators examined how SNOMED CT codes were used to record and represent diagnoses and other conditions in real-world clinical or public health practice.
	Retrieve or analyze patient data	Patient data coded with SNOMED CT were retrieved from data warehouses or registries to conduct observational studies.
	Prove utility	Studies demonstrated that using SNOMED CT in clinical settings improved operational or clinical outcomes.
Non-operational	SNOMED CT mentioned as an example	SNOMED CT was mentioned in the article but was not the primary focus of an experiment or survey.
	Other	SNOMED CT use cases were described in literature reviews, surveys, correspondence, or editorials.

"Interoperability standards for health information systems" usage category. Although data elements were first mapped to SCT in this study (corresponding to the "Design considerations" usage category, representing a less mature stage), we prioritized the more mature stage represented by the subsequent health information system implementation.

The initial 100 randomly selected articles were classified independently by two authors to establish detailed classification policies and resolve potential discrepancies. After establishing the basic classification framework, the first author (EC) classified the remaining articles according to the agreed-upon policy. When novel use cases emerged, the two authors discussed them and reached consensus on usage category assignment.

Categorization by countries and clinical domains

Understanding the geographic and clinical distribution of SCT use provides essential context for interpreting adoption patterns and identifying implementation barriers or facilitators. Tracking these patterns over time—as we do by comparing the 2020-2025 period to the pre-COVID era—enables us to assess how COVID-19 reshaped the landscape of SCT use and whether SCT is achieving its goal of becoming a global standard for interoperability, or whether adoption remains siloed in specific contexts.

We defined the region of SCT use for each paper by carefully reading the manuscript and determining in which country SCT was deployed if SCT was used in an operational setting. Studies were categorized as "international" if the research was conducted at a multi-national level, such as collaborative projects involving data harmonization across multiple countries. If SCT was utilized in theoretical or experimental settings without a clear geographic deployment context, we determined the country based on the first author's institutional affiliation.

We classified articles into one or more clinical domains (e.g., cardiovascular, musculoskeletal, infectious disorders) when experiments were performed in clinical settings or utilized datasets representing specific clinical domains. To address the substantial increase in COVID-19 research during the early 2020s pandemic, we separated "COVID-19" from the general "Infectious disease" domain. Articles could be assigned to multiple clinical domains, while others might be assigned to none if SCT was used in purely technical contexts.

Region and domain classifications were not performed for articles in the "Non-operational" focus category to exclude cases from non-clinical or non-operational settings.

Results

The literature search conducted on July 6, 2025, in PubMed and Embase retrieved 1,139 articles after deduplication. Following the exclusion process, 651 articles from these two bibliographic databases were included in the final review (Figure 1).

Countries and clinical domains

Excluding articles categorized as "Non-operational," the United States was the largest contributor of SCT-related publications with 188 papers, followed by the United Kingdom with 92 papers. Figure 2 illustrates changes in the composition of the five largest contributing countries between January 2015-September 2020 and October 2020-June 2025, with data from the earlier period obtained from our previous review.[3]

Table 3 presents the top 10 clinical domains, allowing double-counting when papers addressed multiple domains. Alimentary or hepatobiliary disorders ($n=52$) emerged as the largest contributor in the 2020-2025 period, rising from fifth place ($n=18$) during the 2015-2020 period. COVID-19 ranked second, a new domain not observed in the pre-2020 era. Although the absolute number of articles in the cancer domain increased from 37 to 42, this domain dropped from first place in the 2015-2020 period to third in the 2020-2025 period.

Evolution of SNOMED CT implementation maturity

During 2020-2025, the majority of use cases were allocated to the "Implementation" and "Evaluation/Commodity" focus categories, representing 172 (26.4%) and 208 (32.0%) papers, respectively (Figure 3). To provide a comprehensive view of how SCT implementation maturity evolved, we compared the distributions of focus categories between the January 2015-September 2020 and October 2020-June 2025 periods, excluding the "Non-operational" category. Compared with 369 articles published during 2015-2020, the 2020-2025 period witnessed expanded use cases at more mature implementation stages (i.e., "Implementation" and "Evaluation/Commodity") while both the number and proportion of papers at less mature stages (i.e., "Theoretical" and "Predevelopment") decreased (Figure 3).

Usage categories

This section describes notable trends in selected usage categories. Table 4 presents selected subcategories within each usage category. Complete usage category classifications and detailed use case descriptions, as well as countries and clinical domains, for all included articles are available in Supplementary Appendix B, available as supplementary material *Journal of the American Medical Informatics Association* online.

Predevelopment: Mapping or comparing to other terminology systems ($n=47$)

The predominant mapping target was International Classification of Diseases (ICD) codes ($n=11$), followed by nursing terminologies such as the International Classification for Nursing Practice ($n=3$). SCT also served as a foundational model for newly developed terminologies or ontologies, incorporating its content and semantics into broader terminology networks ($n=12$), a continuing trend from the 2013-2020 era. However, unlike the 2013-2020 period when mapping tasks were primarily directed toward other formal ontologies or terminologies, the 2020-2025 period saw an increased number of use cases where interface terminologies were mapped to SCT ($n=7$). This shift highlights SCT's evolving role as a standard representation layer for clinical concepts originating from diverse local sources.

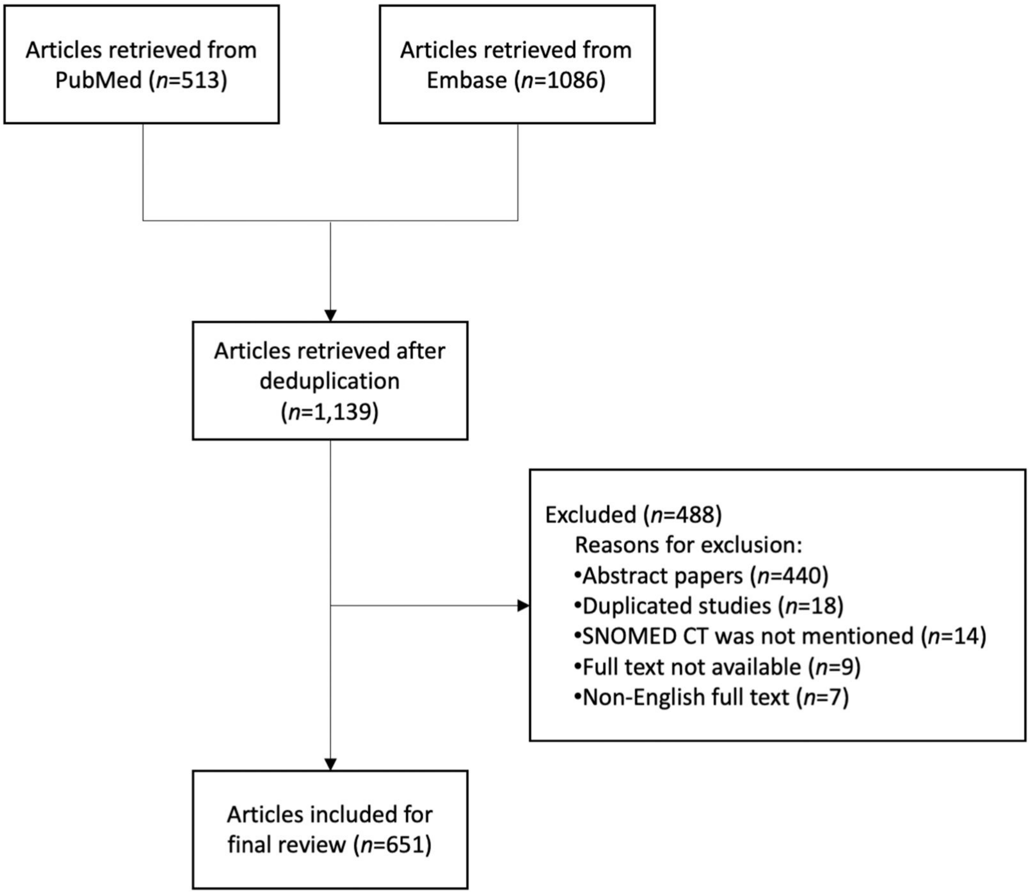

Figure 1. Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the study selection process from PubMed and Embase.

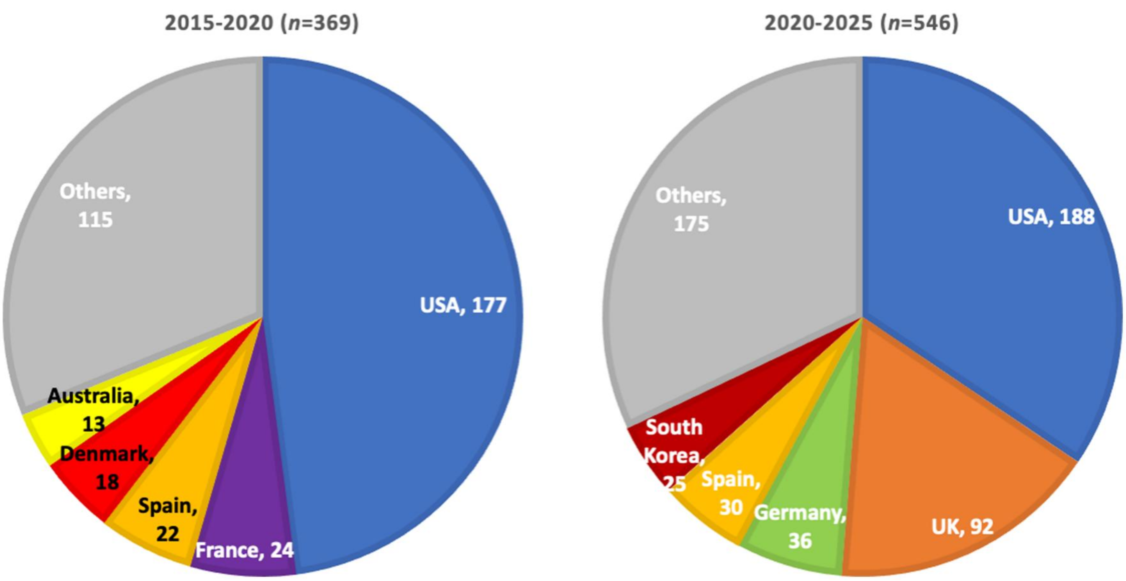

Figure 2. Geographic distribution of the five largest contributing countries comparing two periods: January 2015-September 2020 and October 2020-June 2025, excluding articles categorized into “Non-operational.” Data from the earlier period were obtained from the authors’ previous review.[3]

Table 3. Top 10 clinical domains by number of SNOMED CT use cases, excluding the "Non-operational" category (papers addressing multiple domains were counted in each relevant domain).

Domain	*n*
Alimentary or hepatobiliary disorders	52
COVID-19	47
Cancers	42
Mental disorders	26
Cardiovascular disorders	25
Primary/preventive care	21
Pulmonary disorders	19
Neurologic disorders	17
General surgery (including surgical approach to breast, thyroid, or other disorders)	17
Skin disorders	16

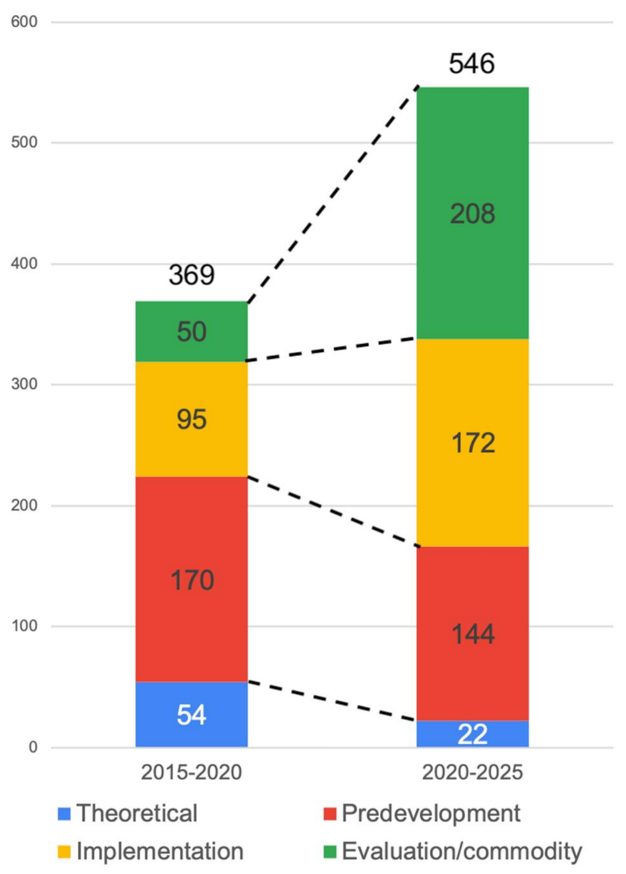

Figure 3. Distribution of SNOMED CT use by implementation maturity across four focus categories, comparing January 2015-September 2020 with October 2020-June 2025 (excluding the "Non-operational" category). Data from the earlier period were obtained from the authors' previous review.[3]

Mapping between SCT and other terminologies was performed using manual mapping tables or code rulebooks ($n = 12$), algorithmic approaches with attention-based models,[12,13] semantic similarity measurement,[14] off-the-shelf mapping tools,[15,16] or hybrid methods.[17–19] SCT expressions served as an intermediary bridge when mapping between terminology systems[20–22] or mapping free-text sources to terminologies.[23] Three studies used SCT's post-coordinated expressions to represent complex concept meanings.[20,24,25]

Predevelopment: Design considerations ($n = 26$)

SCT was employed to represent biomedical concepts in a standardized manner, either independently or in conjunction with other terminology systems, to enhance semantic interoperability. Compared to the pre-2020 era, the 2020-2025 period showed substantially increased focus on SCT's role in facilitating data harmonization, with studies applying its design principles to capture the semantic meaning of data elements from heterogeneous sources ($n = 11$) and multiple terminology systems ($n = 3$). This growing research emphasis on SCT's design paradigms for harmonization reflects its expanding application across multiple platforms and domains, as well as the increasing need for interdisciplinary collaboration.

Efforts were made to integrate formal ontology principles into SCT's logical frameworks. For example, Ceusters and Pengput introduced ontology aspects from Basic Formal Ontology to establish axioms compatible with upper-level ontologies.[26] In another study, they evaluated the feasibility of designing a first-order logic framework capable of translating SCT's terminological perspective on patient data.[27]

Predevelopment: Interoperability standard for health information systems ($n = 40$)

This category represents when data elements and response values in clinical information systems were mapped to SCT. Recognizing that SCT has increasingly been incorporated into a broader scope of health information systems beyond electronic health records (EHRs) during the 2020-2025 period,[28–32] we expanded this usage category to include any type of healthcare-related information system. Twenty-one studies demonstrated SCT incorporation into FHIR Resources/Profiles or Clinical Document Architecture. Fourteen studies applied SCT to standardize health data into CDMs across multiple populations and sites for integrated research. SCT was also incorporated into templates for standardized reporting, such as an SCT-based template for reporting Cobb angle measurements to facilitate efficient secondary use.[33]

Implementation: Encoding methodologies ($n = 40$)

This category represents when studies investigated incorporation of SCT components—relations and descriptions—into knowledge base models. We established this use case as a distinct category to reflect SCT's increased use as a graph-based biomedical knowledge database for enhancing machine interpretability and validity of healthcare AI systems during the 2020-2025 period.

Table 4. Distribution of papers by usage categories and selected subcategories (2020-2025).

Focus category	Usage category	Selected subcategory
Theoretical (*n* = 22)	Terminology analysis (*n* = 22)	Structural/hierarchical (*n* = 10) Lexical/linguistic (*n* = 6) Semantic (*n* = 2) Abstraction network (*n* = 2)
Predevelopment (*n* = 144)	Translation (*n* = 1)	German (*n* = 1)
	Prospective content coverage (*n* = 30)	Data elements (*n* = 10) Domain terms (*n* = 8) Interface codes or terminologies (*n* = 7) Medical corpora (*n* = 4)
	Mapping or comparing to other terminology systems (*n* = 47)	Interface terminologies (*n* = 12) Newly developed ontologies (*n* = 12) International Classification of Diseases (*n* = 10) Nursing terminologies (*n* = 3)
	Design considerations (*n* = 26)	The role and use of interface terminologies in conjunction with SNOMED CT to facilitate data capture (*n* = 14) Process and challenges related to the development of subsets (*n* = 8) Logical frameworks (*n* = 2) Version control, management and migration (*n* = 2)
	Interoperability standards for health information systems (*n* = 40)	Fast Healthcare Interoperability Resources and Clinical Document Architecture (*n* = 21) Common data models (*n* = 14) Electronic health records frameworks (*n* = 2)
Implementation (*n* = 172)	Encoding methodologies (*n* = 40)	Using graph feature of SNOMED CT (*n* = 28) SNOMED CT as lexical dictionaries (*n* = 11)
	Used for classifying or coding (*n* = 76)	Free-text clinical notes/narratives (*n* = 38) Data elements/codes (*n* = 10) Diagnostic criteria (*n* = 4) Consumer-generated contents (*n* = 2) Nursing records (*n* = 2) Pathology reports (*n* = 2) Radiology reports (*n* = 2)
	Implementation of SNOMED CT (*n* = 56)	Data repositories/registries (*n* = 26) Query/terminology services (*n* = 19) Clinical decision support system (*n* = 6)
Evaluation/Commodity (*n* = 208)	Gold standard for model validation (*n* = 5)	Natural language processing models (*n* = 3)
	Assess code use and activity (*n* = 15)	Critically evaluate the quality of coding (*n* = 8) Code use survey (*n* = 7)
	Retrieve or analyze patient data (*n* = 186)	National registries/cohort (*n* = 94) Commercial registries (*n* = 52) Hospital or clinical data warehouse data (*n* = 36)
	Prove utility (*n* = 2)	No subcategories
Non-operational (*n* = 105)	SNOMED CT mentioned as an example (*n* = 23)	No subcategories
	Others (*n* = 82)	Letters/correspondence (*n* = 37) Literature reviews/surveys (*n* = 33) Editorials (*n* = 12)

To represent machine-interpretable knowledge about relationships among biomedical entities, researchers utilized subject-predicate-object triples defined by the SCT relation component (*n* = 28). Studies proposed novel methods for identifying biomedical entities from free-text sources (*n* = 4), projecting recognized entity relations into embedding space (*n* = 3), or constructing knowledge graphs using identified subject-predicate-object triples defined in SCT (*n* = 11). The embedding information was used to predict missing relations in SCT (*n* = 2) or for terminology binding (*n* = 3). Semantic enrichment of data elements was achieved by aggregating variables according to their ontological common ancestry.[34] Two studies incorporated multiple biomedical ontologies, including SCT, into unified knowledge graphs to represent medical phenotype networks.[35,36]

SCT description information was incorporated into natural language processing (NLP) pipelines to provide standardized semantic lexicons of biomedical vocabularies ($n = 11$), primarily for medical entity recognition tasks ($n = 4$). SCT components, such as semantic tags, provided validation frameworks for named-entity recognition or classification tasks.[37] In another study, descriptive temporal relationship features from SCT were used to classify temporal interval properties between pairs of medically relevant events in EHRs.[38]

Implementation: used for classifying or coding ($n = 76$)

SCT concepts were used for biomedical concept identification from unstructured data using off-the-shelf tools such as EHRead ($n = 10$), MedCAT ($n = 7$), cTAKES ($n = 3$), or MetaMap ($n = 3$). During the 2020-2025 period, authors increasingly adopted transformer-based architectures such as BERT or its variants ($n = 9$) and large language models (LLMs) ($n = 2$) for extracting biomedical entities ($n = 22$), moving away from traditional machine-learning models such as support vector machines or non-transformer neural networks that had dominated the pre-2020 era. Despite these automated advances, manual extraction of SCT concepts from unstructured data sources was still performed in 13 studies to leverage domain expertise and ensure validity of the results.

Six studies implemented digital phenotyping by mapping data elements to SCT and classifying them into predefined group entities for computational use of patient information from EHRs. While most studies identified pre-coordinated SCT concepts, two studies utilized post-coordinated concept sets.[39,40]

SCT concepts were identified from unstructured sources for secondary purposes. Food domain corpora were annotated with SCT concepts, contributing to the development of NLP methods for extracting food information from textual data.[41,42] Lexicons of family history-related entities were constructed to provide standardized sets of reusable and interpretable linguistic patterns for NLP.[43] Xin et al.[44] performed topic modeling on social media posts using SCT's semantic type information.

Implementation: implementation of SNOMED CT ($n = 56$)

Twenty-six of 56 articles in this category presented data repositories or registries. Query and terminology systems, including annotation mapping tools ($n = 4$), post-coordination or expression constraint language builders ($n = 4$), terminology servers ($n = 3$), and search engines ($n = 3$), contributed significantly to this use case ($n = 19$). Six studies used SCT to provide domain-specific features for clinical decision support systems (CDSSs), such as question-answering agents[45] and drug interaction checkers.[46] Beyond CDSSs, consumer applications such as personal health record applications[47] or bibliographic databases[48] were developed using vocabulary mappings incorporating SCT—a use case not identified in our 2013-2020 review.

Evaluation/commodity: gold standard for model validation ($n = 5$)

This newly identified usage category emerged with the growing need for algorithm and machine learning model validation in the 2020s. Data annotated with SCT were used as gold-standard benchmarks to evaluate the performance of classification,[49,50] machine translation,[51] and information retrieval models.[52]

Evaluation/commodity: assess code use and activity ($n = 15$)

This usage category represents a new use case that emerged after 2020 as SCT adoption in EHR systems became widespread. While auditing of SCT in the pre-2020 era primarily targeted structural validity of the terminology itself, yielding 44 papers classified into the "Terminology auditing" usage category in the previous review, the 2020-2025 period shifted focus to analyzing patterns in how desirably SCT codes were utilized in practical settings among clinicians or coders. Seven studies investigated how SCT was used in clinical activities such as patient referral ($n = 3$) and outpatient encounters ($n = 2$). Eight studies assessed the quality of SCT coding in EHR systems ($n = 6$), reporting systems ($n = 1$), and cohort data ($n = 1$).

Evaluation/commodity: retrieve or analyze data ($n = 186$)

In this most prominent usage category, 94 studies retrieved SCT codes for retrospective epidemiologic research from national or public registries, including All of Us ($n = 23$), UK Clinical Practice Research Datalink (CPRD) ($n = 12$), Danish Pathology Databank ($n = 9$), and Epidemiology Strengthened by histoPathology Reports in Sweden (ESPRESSO) ($n = 8$). The substantial contribution of the All of Us and CPRD registries to secondary data analysis represents a notable development not observed in our 2013-2020 review. In the United States, SCT-coded patient data managed by commercial federated networks were actively utilized, with researchers retrieving data from IBM® Explorys (Cleveland, OH) ($n = 36$) or TriNetX (Cambridge, MA) ($n = 5$).

Beyond clinical or pathologic diagnoses, SCT codes were used to identify procedures ($n = 12$), encounter reasons ($n = 3$), medication use ($n = 3$), anatomical sites,[53] smoking status,[54] employment status,[55] environmental hazards,[56] and special education-qualifying health outcomes.[57] SCT codes also supported non-epidemiologic studies, with commercial databases such as IBM MarketScan® providing insurance claims data coded with SCT for healthcare service research ($n = 3$).

Evaluation/commodity: prove utility ($n = 2$)

Two studies were identified in this category. Golburean et al.[58] qualitatively examined the impact of transitioning from ICD-10 to SCT on physicians' documentation practices. Roberts et al.[59]

demonstrated that SCT codes with higher specificity explained more variance in predicting length of stay for patients with pneumonia compared to well-established clinical indicators.

SNOMED CT use trends from technical literature

Given the increased volume of use cases where SCT was integrated into real-world applications identified in the current review, along with review articles on SCT applications in machine learning and knowledge graph technologies for enhanced standardization of represented entities,[60–67] we conducted a supplementary post-hoc analyses to examine SCT use in technical contexts. We searched ACM Digital Library and IEEE Xplore using the same publication date range as the primary PubMed and Embase search. Anticipating that SCT would have limited recognition as a resource for the technical development in these databases, we applied a simple "snomed" query without further query refinement to capture the broadest possible landscape of SCT in technical domains.

We retrieved 148 articles from ACM Digital Library and 23 from IEEE Xplore. After excluding duplicates ($n=1$), abstract-only publications ($n=23$), articles that did not mention SCT ($n=5$), articles also indexed in PubMed or Embase ($n=3$), and non-English articles ($n=2$), we analyzed 137 eligible articles. The complete list of articles from these databases is available in Supplementary Appendix C, available as supplementary material *Journal of the American Medical Informatics Association* online. While a substantial portion of retrieved articles were classified into the "Non-operational" focus category ($n=51$, 37.2%), where SCT was briefly mentioned as an example of baseline knowledge graph models or interoperability standards, the majority of retrieved articles comprised the "Implementation" focus category ($n=54$, 39.4%). Approximately half of these implementation studies ($n=28$) were in the "Encoding methodologies" usage category, where SCT was used for entity recognition and knowledge graph construction.

Discussion

In this review, we observed a fundamental evolution in how SCT is utilized: whereas earlier use was predominantly characterized by codification and storage of clinical information, the post-2020 era (2020-2025) witnessed SCT's emergence as a graph-based biomedical knowledge database capable of enhancing semantic inference in AI applications in a prominent way. Crucially, this evolution extended SCT's user base beyond clinicians and health system administrators to include software developers, data scientists, and other technical professionals in the computer science and engineering communities in a deeper way—a demographic minimally represented in our previous review. Beyond modifying the scope of the review, in comparison to the scope utilized in the previous review article, we also expanded the literature base to include established computer science and engineering databases and included research evidence from the latter communities. The approach allowed us to comprehensively document and draw conclusions about the full scope of SCT's evolving role across both clinical and technical domains.

Prospects

Compared with the previous five-year period, 2020-2025 witnessed a dramatic shift toward more mature stages of SCT implementation. An increasing number of studies used SCT to retrieve patient data from registries, enabled by its widespread adoption as a standardized terminology for representing, storing, and sharing clinical information among healthcare institutions.

The 2020-2025 period experienced several converging developments that transformed how SCT was leveraged in healthcare systems worldwide. The COVID-19 pandemic, which emerged in early 2020, created an urgent global need for standardized clinical data management across international borders. COVID-19-related studies emerged as the second most-investigated domain in our review. SNOMED International's response—publishing coronavirus-related concepts in January 2020 and issuing interim releases to support pandemic surveillance—demonstrated the terminology's agility and critical role in global health emergencies.[68] This pandemic-driven use case revealed both the strengths and limitations of existing clinical terminology infrastructures, prompting accelerated development of interoperability solutions and harmonization initiatives.

The current review revealed that the United Kingdom substantially increased its contribution to SCT research in the 2020-2025 period. With only 10 publications identified during 2015-2020, the UK's increased output to 92 during 2020-2025 is likely attributable to nationwide clinical data repositories that utilize SCT as a standard terminology, exemplified by the CPRD. Data repositories coded with SCT have facilitated epidemiologic studies due to the terminology's robust capacity to represent clinical concepts, offering substantially greater detail, flexibility, and retrieval capabilities than the ICD.[69] The proliferation of data warehouses utilizing SCT continues to facilitate real-world data research—the structured, semantically rich data provided by SCT enables deeper insights into patient outcomes and supports the development of CDSSs.[70]

Analysis of articles indexed by ACM Digital Library and IEEE Xplore revealed that NLP and technical communities increasingly employ SCT as a biomedical knowledge base for processing large volumes of data. SCT's ontological features can address the black box nature of machine learning algorithms and LLMs by enhancing their explainability and interpretability.[71] SCT also contributes to reducing data noise and improving the reliability of predictive models by standardizing inputs across clinical, demographic, and behavioral data.[72] As regulatory frameworks increasingly require explainable AI, particularly in sensitive domains such as healthcare, SCT is well-positioned to meet these requirements while maintaining high performance and providing clear explanations.

Challenges

The emergence of articles in the "Assess code use and activity" category contrasted with the decrease in "Terminology analysis" articles from 36 in 2015-2020 to 22 in 2020-2025, suggesting research focus has shifted from auditing SCT's logical structure to examining how clinicians appropriately use SCT in practice. As noted in previous literature, comprehensive guidance for selecting optimal SCT concepts for specific clinical scenarios remains limited, and both novice and experienced users have developed individualized approaches and tools, which may lead to coding variation.[73] Lack of coding consistency is a significant concern for data standards advocates, as it can undermine the fundamental goals of implementing standardized terminologies.

Although research using SCT-coded real-world data has potentially contributed to advances in biomedical science and public health, no studies in this review demonstrated that SCT use directly improved patient outcomes. Only two studies examined non-clinical, operational outcomes, such as changes in documentation practices and the predictive power of SCT for clinical outcome variance. The absence of evidence demonstrating improved patient outcomes represents a significant barrier to user acceptance of SCT in their workflows, compounded by usability challenges, insufficient training, and inadequate integration within existing electronic health record systems.[74] When integrated into health information systems, SCT has the potential to deliver tangible benefits, including improved data sharing, enhanced CDSSs, more efficient secondary data analysis for quality improvement and research, and ultimately, improved patient outcomes. Future research must provide more robust evidence for the clinical and operational benefits of SCT use to ensure its utility and adaptability among users.

Limitations of the current review

To maintain scientific rigor in our review methodology, we excluded abstract-only papers due to their limited detailed descriptions of SCT usage. Had we included use cases published as abstracts, the number of papers in the "Evaluation/Commodity" category would have expanded further, given the substantial volume of retrospective observational studies that rely on SCT codes for patient data retrieval.

A substantial collection of SCT use cases likely appears in gray literature, including textual and video formats, such as presentations at SNOMED CT Expo held by SNOMED International. Although we could not systematically analyze use cases presented at SNOMED CT Expo due to the absence of textual conference abstracts, interested readers may find valuable use cases in the recorded sessions available on YouTube.

Conclusion

By reviewing SCT use in the early 2020s and comparing it with the previous five-year period, we gained valuable perspectives on both the prospects and challenges for SCT implementation in the late 2020s. While SCT has played an essential role in collecting, exchanging, and analyzing biomedical data, the early 2020s marked a new era in which SCT emerged as a machine-interpretable biomedical knowledge base, owing to its comprehensiveness, structured ontological design, and widespread integration into EHRs. The ubiquitous use of SCT across the biomedical domain—from theoretical information science to clinical practice—will continue to expand, supporting neurosymbolic approaches that enhance the explainability and interpretability of AI. To realize this potential, more comprehensive evidence demonstrating the impact of SCT use on operational and clinical outcomes must be documented.

Author contributions

Eunsuk Chang (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft), Javed Mostafa (Conceptualization, Methodology, Resources, Supervision, Validation, Writing—review & editing)

Supplementary material

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

Funding

None declared.

Conflicts of interest

None declared.

Data availability

Data supporting the findings of this study are available in Supplementary Materials B and C, which provide detailed information for all included articles: title, authors, publication date, journal, countries, clinical domains, focus categories, usage categories, subtasks, and operational tasks. These resources are provided to support reproducibility and enable further research on SNOMED CT use cases.

References

1. World Health Organization. *Global Strategy on Digital Health 2020-2025*. World Health Organization; 2021.
2. National Library of Medicine. SNOMED CT. Health data standards and terminologies: a tutorial. May 16, 2025. Accessed February 17, 2026. https://www.nlm.nih.gov/oet/ed/healthdatastandards/02-510.html
3. Chang E, Mostafa J. The use of SNOMED CT, 2013-2020: a literature review. *J Am Med Inform Assoc*. 2021;28:2017-2026. https://doi.org/10.1093/jamia/ocab084
4. Vorisek CN, Lehne M, Klopfenstein SAI, et al. Fast healthcare interoperability resources (FHIR) for interoperability in health research: systematic review. *JMIR Med Inform*. 2022; 10:e35724. https://doi.org/10.2196/35724
5. SNOMED International. HL7 International and SNOMED International continue long-standing partnership. 2021. Accessed November 16, 2025. https://www.snomed.org/news/hl7-international-and-snomed-international-continue-long-standing-partnership
6. Schippers D, Stegwee RA. Exploring the clinical value of the International Patient Summary - a systematic review.

In: Abramson D, Vassar S, Jin C, eds. *Proceedings of the 2024 Australasian Computer Science Week*. ACM; 2024:78-86. https://doi.org/10.1145/3641142.3641160

7. Regenstrief Institute, Inc, SNOMED International. Introducing the LOINC Ontology: A LOINC and SNOMED CT Interoperability Solution. 2025. Accessed November 16, 2025. https://loincsnomed.org/#:∼:text=Introducing%20the%20LOINC%20Ontology:%20A, RESOURCES
8. International Health Terminology Standards Development Organisation. SNOMED CT COVID-19 Coding Guide. 2025. Accessed November 15, 2025. https://docs.snomed.org/implementation-guides/snomed-ct-guide-for-covid-19
9. Jani BD, Pell JP, McGagh D, et al. Recording COVID-19 consultations: review of symptoms, risk factors, and proposed SNOMED CT terms. *Br J Gen Pract Open*. 2020;4:bjgpopen20X101125. https://doi.org/10.3399/bjgpopen20X101125
10. Lee D, de Keizer N, Lau F, Cornet R. Literature review of SNOMED CT use. *J Am Med Inform Assoc*. 2014;21:e11-9-e19. https://doi.org/10.1136/amiajnl-2013-001636
11. Jones D, Shao J, Wallis H, et al. Towards a newborn screening common data model: the Utah newborn screening data model. *Int J Neonatal Screen*. 2021;7. https://doi.org/10.3390/ijns7040070
12. Chen Y, Hu D, Li M, Duan H, Lu X. Automatic SNOMED CT coding of Chinese clinical terms via attention-based semantic matching. *Int J Med Inform*. 2022;159:104676. https://doi.org/10.1016/j.ijmedinf.2021.104676
13. Huh S. Comparative analysis of ChatGPT-4 for automated mapping of local medical terminologies to SNOMED CT. *Stud Health Technol Inform*. 2025;327:813-817. https://doi.org/10.3233/SHTI250472
14. Hong N, Zhang Y, Ren Y, et al. Cross-language terminology mapping between ICD-10-CN and SNOMED-CT. *Stud Health Technol Inform*. 2022;290:42-46. https://doi.org/10.3233/SHTI220028
15. Stroganov O, Fedarovich A, Wong E, et al. Mapping of UK Biobank clinical codes: challenges and possible solutions. *PLoS One*. 2022;17:e0275816. https://doi.org/10.1371/journal.pone.0275816
16. van Damme P, Fernández-Breis JT, Benis N, Miñarro-Gimenez JA, de Keizer NF, Cornet R. Performance assessment of ontology matching systems for FAIR data. *J Biomed Semantics*. 2022;13:19. https://doi.org/10.1186/s13326-022-00273-5
17. Block LJ, Wong ST, Handfield S, Hart R, Currie LM. Comparison of terminology mapping methods for nursing wound care knowledge representation. *Int J Med Inform*. 2021;153:104539. https://doi.org/10.1016/j.ijmedinf.2021.104539
18. Wang P, Hu Y, Bai S, Zou S. Matching biomedical ontologies: construction of matching clues and systematic evaluation of different combinations of matchers. *JMIR Med Inform*. 2021;9:e28212. https://doi.org/10.2196/28212
19. Yoon D, Han C, Kim DW, et al. Redefining health care data interoperability: empirical exploration of large language models in information exchange. *J Med Internet Res.* 2024; 26:e56614. https://doi.org/10.2196/56614
20. Ohlsen T, Kruse V, Krupar R, Banach A, Ingenerf J, Drenkhahn C. Mapping of ICD-O tuples to OncoTree codes using SNOMED CT post-coordination. *Stud Health Technol Inform*. 2022;294:307-311. https://doi.org/10.3233/SHTI220464
21. Xu J, Fung KW, Bodenreider O. Sequential mapping - a novel approach to map from ICD-10-CM to ICD-11. *Stud Health Technol Inform*. 2022;290:96-100. https://doi.org/10.3233/SHTI220039
22. Güngör B, Deppenwiese N, Mang JM, Toddenroth D. Analysis of the representation of frequent clinical attributes in the unified medical language system. *Stud Health Technol Inform*. 2022;299:217-222. https://doi.org/10.3233/SHTI220987
23. Isaradech N, Khumrin P. Auto-mapping clinical documents to ICD-10 using SNOMED-CT. *AMIA Jt Summits Transl Sci Proc*. 2021;2021:296-304.
24. Haffer N, Thun S. Postcoordination of LOINC codes in SNOMED CT. *Stud Health Technol Inform*. 2021;278:19-26. https://doi.org/10.3233/SHTI210045
25. Fung KW, Xu J, Brear H, et al. Promoting interoperability between SNOMED CT and ICD-11: lessons learned from the pilot project mapping between SNOMED CT and the ICD-11 Foundation. *J Am Med Inform Assoc*. 2024;31:1631-1637. https://doi.org/10.1093/jamia/ocae143
26. Ceusters W, Pengput A. Axiomatizing SNOMED CT disorders: should there be room for interpretation? *Form Ontol Inf Syst*. 2023;377:140-154. https://doi.org/10.3233/FAIA231124
27. Pengput A, Ceusters W. Setting the scene to link SNOMED CT to realism-based ontologies. *Stud Health Technol Inform*. 2024;310:84-88. https://doi.org/10.3233/SHTI230932
28. Shivers J, Amlung J, Ratanaprayul N, Rhodes B, Biondich P. Enhancing narrative clinical guidance with computer-readable artifacts: authoring FHIR implementation guides based on WHO recommendations. *J Biomed Inform*. 2021;122:103891. https://doi.org/10.1016/j.jbi.2021.103891
29. Kazemi-Arpanahi H, Shanbehzadeh M, Mirbagheri E, Baradaran A. Data integration in cardiac electrophysiology ablation toward achieving proper interoperability in health information systems. *J Educ Health Promot*. 2020;9:262. https://doi.org/10.4103/jehp.jehp_751_19
30. Shanbehzadeh M, Kazemi-Arpanahi H, Valipour AA, Zahedi A. Notifiable diseases interoperable framework toward improving Iran public health surveillance system: Lessons learned from COVID-19 pandemic. *J Educ Health Promot*. 2021;10:179. https://doi.org/10.4103/jehp.jehp_1082_20
31. Bathelt F, Reinecke I, Peng Y, et al. Opportunities of digital infrastructures for disease management-exemplified on COVID-19-related change in diagnosis counts for diabetes-related eye diseases. *Nutrients*. 2022;14:2016. https://doi.org/10.3390/nu14102016
32. Rinaldi E, Drenkhahn C, Gebel B, et al. Towards interoperability in infection control: a standard data model for microbiology. *Sci Data*. 2023;10:654. https://doi.org/10.1038/s41597-023-02560-x
33. Shahid A, Kim J, Byon SS, Hong S, Lee I, Lee B-D. An end-to-end pipeline for automated scoliosis diagnosis with standardized clinical reporting using SNOMED CT. *Sci Rep*. 2025;15:17274. https://doi.org/10.1038/s41598-025-01952-w
34. Toner TM, Pancholi R, Miller P, Forster T, Coleman HG, Overton IM. Strategies and techniques for quality control and semantic enrichment with multimodal data: a case study in

colorectal cancer with eHDPrep. *Gigascience*. 2022;12: giad030. https://doi.org/10.1093/gigascience/giad030
35. Talukder AK, Schriml L, Ghosh A, Biswas R, Chakrabarti P, Haas RE. Diseasomics: actionable machine interpretable disease knowledge at the point-of-care. *PLOS Digit Health*. 2022;1:e0000128. https://doi.org/10.1371/journal.pdig.0000128
36. Hu H, Kerschberg L. Improving causal Bayesian networks using expertise in authoritative medical ontologies. *ACM Trans Comput Healthcare*. 2023;4:1-32. https://doi.org/10.1145/3604561
37. Stojanov R, Popovski G, Cenikj G, Koroušić Seljak B, Eftimov T. A fine-tuned bidirectional encoder representations from transformers model for food named-entity recognition: Algorithm development and validation. *J Med Internet Res*. 2021; 23:e28229. https://doi.org/10.2196/28229
38. Sujansky W, Campbell KE. A formal model for the representation of binary temporal relations in healthcare applications and an efficient algorithm for logic-based temporal subsumption testing and pattern matching. *medRxiv*. Published online November 18, 2023. https://doi.org/10.1101/2023.11.17.23298715
39. Hwang JE, Park H-A, Shin S-Y. Mapping the Korean national health checkup questionnaire to standard terminologies. *Healthc Inform Res*. 2021;27:287-297. https://doi.org/10.4258/hir.2021.27.4.287
40. Schulz S, Martínez Costa C. Semantic representation of medical data collection forms using standards. *Stud Health Technol Inform*. 2025;327:708-712. https://doi.org/10.3233/SHTI250440
41. Cenikj G, Valenčič E, Ispirova G, et al. CafeteriaSA corpus: scientific abstracts annotated across different food semantic resources. *Database (Oxford)*. 2022;2022:baac107. https://doi.org/10.1093/database/baac107
42. Ispirova G, Cenikj G, Ogrinc M, et al. CafeteriaFCD corpus: food consumption data annotated with regard to different food semantic resources. *Foods*. 2022;11:2684. https://doi.org/10.3390/foods11172684
43. Wang L, He H, Wen A, et al. Acquisition of a Lexicon for family history information: bidirectional encoder representations from transformers-assisted sublanguage analysis. *JMIR Med Inform.* 2023;11:e48072. https://doi.org/10.2196/48072
44. Xin Y, Grabowska ME, Gangireddy S, et al. Improving topic modeling performance on social media through semantic relationships within biomedical terminology. *PLoS One*. 2025;20:e0318702. https://doi.org/10.1371/journal.pone.0318702
45. Vazrala S, Khatoon Mohammed T. RBTM: A Hybrid gradient Regression-Based transformer model for biomedical question answering. *Biomed Signal Process Control*. 2025;102: 107325. https://doi.org/10.1016/j.bspc.2024.107325
46. Thiess H, Del Fiol G, Malone DC, et al. Coordinated use of Health Level 7 standards to support clinical decision support: case study with shared decision making and drug-drug interactions. *Int J Med Inform*. 2022;162:104749. https://doi.org/10.1016/j.ijmedinf.2022.104749
47. Park HS, Kim KI, Chung H-Y, et al. A worker-centered personal health record app for workplace health promotion using national health care data sets: design and development study. *JMIR Med Inform*. 2021;9:e29184. https://doi.org/10.2196/29184
48. Shi X, Yin D, Bai Y, et al. A bibliographic dataset of health artificial intelligence research. *Health Data Sci*. 2024;4:0125. https://doi.org/10.34133/hds.0125
49. Berkowitz JS, Srinivasan A, Acitores Cortina JM, Fatapour Y, Tatonetti NP. Biomedical text normalization through generative modeling. *J Biomed Inform*. 2025;167:104850. https://doi.org/10.1016/j.jbi.2025.104850
50. Chew M, Yu C, Stojevski L, et al. A novel natural language processing tool improves colonoscopy auditing of adenoma and serrated polyp detection rates. *J Gastroenterol Hepatol*. 2025;40:1230-1237. https://doi.org/10.1111/jgh.16933
51. Schulz S, Boeker M, Prunotto A. Validation of multiple path translation for SNOMED CT localisation. *Stud Health Technol Inform*. 2022;294:961-962. https://doi.org/10.3233/SHTI220641
52. Shung D, Tsay C, Laine L, et al. Early identification of patients with acute gastrointestinal bleeding using natural language processing and decision rules. *J Gastroenterol Hepatol*. 2021;36:1590-1597. https://doi.org/10.1111/jgh.15313
53. Vadhan JD, Thoppil J, Vasquez O, et al. Primary infection site as a predictor of sepsis development in emergency department patients. *J Emerg Med*. 2024;67:e128-e137. https://doi.org/10.1016/j.jemermed.2024.01.016
54. Razavi AC, Richardson LC, Coronado F, et al. Prevalence and correlates of lipoprotein(a) testing in a diverse cohort of U.S. adults. *JACC Adv*. 2025;4:101826. https://doi.org/10.1016/j.jacadv.2025.101826
55. Finnegan A, Randles R. Where are all the veterans? A mixed methods assessment of a systematic strategy to increase veteran registration in UK primary healthcare practices. *BMJ Open*. 2023;13:e068904. https://doi.org/10.1136/bmjopen-2022-068904
56. Zhu H, Shaw V, Dokic Y, Camacho-Hubbard I, Ranario JS, Orengo IF. Association between non-melanoma skin cancer, psoriasis, and immunomodulatory therapy: a national epidemiologic study. *Arch Dermatol Res*. 2024;317:101. https://doi.org/10.1007/s00403-024-03649-y
57. Rochat R, Goodman E, Miller J, Wang W, Demmler-Harrison GJ. Healthcare outcomes and special education eligibility in children with congenital CMV. *PLoS One*. 2024;19:e0313163. https://doi.org/10.1371/journal.pone.0313163
58. Golburean O, Pedersen R, Melby L, Faxvaag A. Exploring physicians' dual perspectives on the transition from free text to structured and standardized documentation practices: interview and participant observational study. *JMIR Formativ Res.* 2025;9:e63902. https://doi.org/10.2196/63902
59. Roberts L, Lanes S, Peatman O, Assheton P. The importance of SNOMED CT concept specificity in healthcare analytics. *Health Inf Manag*. 2024;53:157-165. https://doi.org/10.1177/18333583221144662
60. Nair SS, Li C, Doijad R, Nagy P, Lehmann H, Kharrazi H. A scoping review of knowledge authoring tools used for developing computerized clinical decision support systems. *JAMIA Open*. 2021;4:ooab106. https://doi.org/10.1093/jamiaopen/ooab106
61. Denck J, Ozkirimli E, Wang K. Machine-learning-based adverse drug event prediction from observational health

data: a review. *Drug Discov Today*. 2023;28:103715. https://doi.org/10.1016/j.drudis.2023.103715
62. Tejani AS, Cook TS, Hussain M, Sippel Schmidt T, O'Donnell KP. Integrating and adopting AI in the radiology workflow: a primer for standards and integrating the healthcare enterprise (IHE) profiles. *Radiology*. 2024;311:e232653. https://doi.org/10.1148/radiol.232653
63. Chang E, Sung S. Use of SNOMED CT in large language models: scoping review. *JMIR Med Inform.* 2024; 12: e62924. https://doi.org/10.2196/62924
64. Noll R, Frischen LS, Boeker M, Storf H, Schaaf J. Machine translation of standardised medical terminology using natural language processing: a scoping review. *N Biotechnol*. 2023;77:120-129. https://doi.org/10.1016/j.nbt.2023.08.004
65. Gholipour M, Khajouei R, Amiri P, Hajesmaeel Gohari S, Ahmadian L. Extracting cancer concepts from clinical notes using natural language processing: a systematic review. *BMC Bioinformatics*. 2023;24:405. https://doi.org/10.1186/s12859-023-05480-0
66. Mitha S, Schwartz J, Hobensack M, et al. Natural language processing of nursing notes: an integrative review. *Comput Inform Nurs*. 2023;41:377-384. https://doi.org/10.1097/CIN.0000000000000967
67. Gaudet-Blavignac C, Foufi V, Bjelogrlic M, Lovis C. Use of the systematized nomenclature of medicine clinical terms (SNOMED CT) for processing free text in health care: systematic scoping review. *J Med Internet Res*. 2021;23: e24594. https://doi.org/10.2196/24594
68. SNOMED International. SNOMED CT and COVID-19. 2025. Accessed November 27, 2025. https://www.snomed.org/covid-19
69. Fung KW, Xu J, Rosenbloom ST, Campbell JR. Using SNOMED CT-encoded problems to improve ICD-10-CM coding-a randomized controlled experiment. *Int J Med Inform*. 2019;126: 19-25. https://doi.org/10.1016/j.ijmedinf.2019.03.002
70. Lieberman MI, Ricciardi TN, Masarie FE, Spackman KA. The use of SNOMED CT simplifies querying of a clinical data warehouse. *AMIA Annu Symp Proc*. 2003;2003:910.
71. Confalonieri R, Guizzardi G. On the multiple roles of ontologies in explanations for neuro-symbolic AI. *Neurosymbolic Artif Intel*. 2025;1. https://doi.org/10.3233/NAI-240754
72. Ambalavanan R, Snead RS, Marczika J, Towett G, Malioukis A, Mbogori-Kairichi M. Ontologies as the semantic bridge between artificial intelligence and healthcare. *Front Digit Health*. 2025;7:1668385. https://doi.org/10.3389/fdgth.2025.1668385
73. Andrews JE, Richesson RL, Krischer J. Variation of SNOMED CT coding of clinical research concepts among coding experts. *J Am Med Inform Assoc*. 2007;14:497-506. https://doi.org/10.1197/jamia.M2372
74. Park H-A, Yu S-J, Jung H. Strategies for adopting and implementing SNOMED CT in Korea. *Healthc Inform Res*. 2021;27: 3-10. https://doi.org/10.4258/hir.2021.27.1.3

Journal of the American Medical Informatics Association, 2026, 00, 1–13
https://doi.org/10.1093/jamia/ocag094
Review