eISSN: 2229-3213 pISSN: 2229-3205
The Journal of Applied Science and Social Science
Volume 16,Issue 06 June, 2026

SEMANTIC MODELING OF DISEASE AND TREATMENT IN AI-BASED MEDICAL SYSTEMS

Gazieva Oydin Djamoldinovna
Lecturer at the Department of
Language Teaching in Medicine
Andijan Branch of Kokand University
Email: turayev.ali01@mail.ru
https://doi.org/10.5281/zenodo.20611948

Abstract. Artificial intelligence in healthcare is often discussed as an algorithmic problem, yet its reliability depends just as strongly on how clinical meaning is represented. This mini-review examines how semantic modeling structures disease and treatment information for AI-based medical systems. A focused narrative synthesis of eight peer-reviewed publications published between 2004 and 2024 was conducted across four domains: terminology integration, logic-based clinical vocabularies, FHIR-based semantic structures, and EHR-derived knowledge graphs. The reviewed literature suggests that dependable medical AI requires more than data aggregation. It requires normalized concepts, typed relations, contextual qualifiers, and traceable provenance. From this synthesis, a four-layer schema emerged: concept normalization, relation modeling, clinical context, and governance. Systems built on these layers are more likely to support transportable prediction, explainable reasoning, and safer clinical deployment. Future work should extend semantic models beyond diagnosis coding toward richer treatment representations that capture intent, sequence, dosage, response, and uncertainty.

Keywords: semantic modeling; disease representation; treatment representation; clinical AI; knowledge graphs; interoperability.

Introduction

Artificial intelligence in medicine is frequently framed as a question of model performance, but in clinical environments the deeper problem is often semantic consistency. The same disease may appear in one system as a billing code, in another as a problem-list entry, and elsewhere as free text; treatments can likewise refer to a prescribed drug, a completed procedure, a planned intervention, or a discontinued regimen. When those distinctions are flattened, AI systems learn from noisy proxies instead of stable clinical meaning. This matters because machine learning is already being used to support diagnosis, treatment, prognosis, workflow optimization, and triage across modern care settings.[1]

Semantic modeling offers a way to reduce that ambiguity. Terminology systems such as the Unified Medical Language System (UMLS) and SNOMED CT, together with interoperable structures such as HL7 FHIR, make it possible to encode diseases and treatments as computable concepts linked by explicit relations rather than as isolated strings or local database fields.[2] The purpose of this article is to examine how such semantic modeling supports AI-based medical systems and to identify the minimum semantic components required for representing disease and treatment in a clinically useful way.

Methods

This article uses a focused narrative mini-review designed to fit an IMRAD format while remaining concise. Eight peer-reviewed English-language publications from 2004 to 2024 were selected because

1 Alvin Rajkomar, Jeffrey Dean, and Isaac Kohane, “Machine Learning in Medicine,” *New England Journal of Medicine* 380, no. 14 (2019): 1347-1358, https://doi.org/10.1056/NEJMra1814259.

2 Olivier Bodenreider, “The Unified Medical Language System (UMLS): integrating biomedical terminology,” *Nucleic Acids Research* 32, suppl_1 (2004): D267-D270, https://doi.org/10.1093/nar/gkh061; R. Cornet and N. de Keizer, “Forty years of SNOMED: a literature review,” *BMC Medical Informatics and Decision Making* 8, suppl_1 (2008): S2, https://doi.org/10.1186/1472-6947-8-S1-S2; Parinaz Tabari et al., “State-of-the-Art Fast Healthcare Interoperability Resources (FHIR)-Based Data Model and Structure Implementations: Systematic Scoping Review,” *JMIR Medical Informatics* 12 (2024): e58445, https://doi.org/10.2196/58445.

they represent influential points in the semantic stack used by medical AI: terminology integration, ontology-based clinical vocabularies, formal semantic validation of FHIR profiles, knowledge-graph learning from electronic medical records, and deployment architectures for interoperable AI. The goal was not exhaustive retrieval but conceptual coverage.

Each source was read against four analytical dimensions: **(1)** concept normalization, **(2)** relation modeling, **(3)** clinical context, and **(4)** governance or provenance. These dimensions were then compared across the corpus to identify recurring design requirements for semantic models of disease and treatment. No patient-level data were collected, and no statistical meta-analysis was attempted; the result is therefore an interpretive synthesis rather than an evidence-aggregation study.

Results

The reviewed literature converged on one broad conclusion: semantic modeling becomes clinically valuable when it moves from simple labeling toward structured representation.

At the first level, **concept normalization** remains foundational. Bodenreider shows that UMLS functions as an integrating layer across biomedical vocabularies, while Cornet and de Keizer describe SNOMED's evolution into a logic-based clinical terminology.[3] For disease modeling, this means that synonymous expressions can be mapped to one concept rather than treated as separate events. For treatment modeling, it means that ingredients, drug classes, procedures, and therapeutic intents can be distinguished instead of collapsed into generic "care" variables. In practice, normalization reduces variation caused by local coding habits and makes cross-site AI training more plausible.

At the second level, **relation modeling** determines whether a system can do more than count concepts. Solbrig and colleagues demonstrate that FHIR profiles can be modeled and validated using semantic web Shape Expressions, while Tabari and colleagues' systematic scoping review shows that FHIR-based data models are increasingly central to interoperable clinical data use, even though implementation choices remain heterogeneous.[4] In semantic terms, disease and treatment should not be stored as disconnected items. They should be represented through typed relations such as disease-has manifestation, disease-treated by medication, procedure-contraindicated in condition, or therapy-associated with outcome. Once relations are explicit, AI systems can reason over care pathways rather than merely over columns.

At the third level, **knowledge-graph modeling** offers a practical way to operationalize those relations. Rotmensch and colleagues showed that high-quality health knowledge graphs can be learned from electronic medical records, linking diseases and symptoms directly from routine documentation.[5] L. Murali and colleagues, in a later literature study, found that EHR-based medical knowledge graphs support extraction, completion, diagnosis prediction, recommendation, and decision support, while also facing persistent challenges related to dimensionality, fusion, and updating.[6] Although many published examples emphasize diagnostic relations, the same graph logic extends naturally to treatment semantics by adding edges for dosage, sequencing, adverse effects, and response.

Across the corpus, a four-layer schema emerged for semantic modeling in AI-based medical systems. The **concept layer** supplies standardized identifiers for diseases, symptoms, drugs, procedures, and outcomes. The **relation layer** specifies how those entities connect. The **context layer** adds temporality, severity, uncertainty, route, dose, and treatment intent. The **governance layer** records provenance, versioning, and the source of each assertion. Kasparick and colleagues argue that open semantic interoperability is precisely what enables AI systems in high-acuity medical environments to combine

3 Bodenreider, "Unified Medical Language System"; Cornet and de Keizer, "Forty years of SNOMED."

4 Harold R. Solbrig et al., "Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx)," *Journal of Biomedical Informatics* 67 (2017): 90-100, https://doi.org/10.1016/j.jbi.2017.02.009; Tabari et al., "State-of-the-Art Fast Healthcare Interoperability Resources."

5 Maya Rotmensch et al., "Learning a Health Knowledge Graph from Electronic Medical Records," *Scientific Reports* 7, no. 1 (2017): 5994, https://doi.org/10.1038/s41598-017-05778-z.

6 Lino Murali et al., "Towards electronic health record-based medical knowledge graph construction, completion, and applications: A literature study," *Journal of Biomedical Informatics* 143 (2023): 104403, https://doi.org/10.1016/j.jbi.2023.104403.

patient, device, and outcome data in a usable way.[7] The implication is clear, a disease-treatment model is clinically trustworthy only when meaning, context, and lineage remain visible.

Discussion

The findings suggest that semantic modeling is not a preliminary data-cleaning step but a structural condition for safe clinical AI. A disease label without stage, severity, or diagnostic certainty is often too coarse for decision support. A treatment label without intent, dose, sequence, and observed response is equally incomplete. An antibiotic order, for example, may represent empiric therapy, targeted therapy, prophylaxis, or a canceled plan; those states are not interchangeable, and an AI system that treats them as equivalent risks producing misleading recommendations.

The review also clarifies why semantic structure matters for explainability. Black-box predictions may be tolerated in low-stakes digital services, but medicine requires traceability. When a model suggests a treatment, clinicians need to know whether the recommendation emerged from diagnosis codes alone, from longitudinal symptom patterns, from prior response, or from contraindication logic. Knowledge graphs and semantically constrained FHIR representations do not eliminate uncertainty, but they do make the chain of reasoning more inspectable. That improves portability across institutions, because the system depends less on idiosyncratic local field names and more on shared clinical meaning.[8]

This study has limitations. The corpus was deliberately small, the analysis was interpretive, and the selected sources emphasize infrastructure more than prospective bedside trials. Even so, the pattern across the literature is consistent. Effective AI-based medical systems require disease and treatment to be semantically normalized, relationally linked, contextually qualified, and provenance-tracked. Better algorithms remain important, but better meaning structures may be the more urgent prerequisite. Without them, medical AI scales ambiguity; with them, it becomes more transferable, more auditable, and more clinically credible.

References

1. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. New England Journal of Medicine. 2019;380(14):1347-1358. doi:10.1056/NEJMra1814259.
2. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Research. 2004;32(Database issue):D267-D270. doi:10.1093/nar/gkh061.
3. Cornet R, de Keizer N. Forty years of SNOMED: a literature review. BMC Medical Informatics and Decision Making. 2008;8(Suppl 1):S2. doi:10.1186/1472-6947-8-S1-S2.
4. Solbrig HR, Prud'hommeaux E, Grieve G, et al. Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx). Journal of Biomedical Informatics. 2017;67:90-100. doi:10.1016/j.jbi.2017.02.009.
5. Rotmensch M, Halpern Y, Tlimat A, Horng S, Sontag D. Learning a Health Knowledge Graph from Electronic Medical Records. Scientific Reports. 2017;7(1):5994. doi:10.1038/s41598-017-05778-z.
6. Kasparick M, Andersen B, Franke S, et al. Enabling artificial intelligence in high acuity medical environments. Minimally Invasive Therapy & Allied Technologies. 2019;28(2):120-126. doi:10.1080/13645706.2019.1599957.
7. Murali L, Gopakumar G, Viswanathan DM, Nedungadi P. Towards electronic health record-based medical knowledge graph construction, completion, and applications: A literature study. Journal of Biomedical Informatics. 2023;143:104403. doi:10.1016/j.jbi.2023.104403.
8. Tabari P, Costagliola G, De Rosa M, Boeker M. State-of-the-Art Fast Healthcare Interoperability Resources (FHIR)-Based Data Model and Structure Implementations: Systematic Scoping Review. JMIR Medical Informatics. 2024;12:e58445. doi:10.2196/58445.

7 Martin Kasparick et al., "Enabling artificial intelligence in high acuity medical environments," *Minimally Invasive Therapy & Allied Technologies* 28, no. 2 (2019): 120-126, https://doi.org/10.1080/13645706.2019.1599957.

8 Rajkomar, Dean, and Kohane, "Machine Learning in Medicine"; Kasparick et al., "Enabling artificial intelligence in high acuity medical environments."