UNIVERSITAS
GADJAH MADA
**Pengembangan Sistem Rekomendasi Klinis untuk Penyakit Jantung Koroner Kronis Menggunakan Knowledge
Graph Berbasis Rekam Medis FHIR, Panduan Klinis, dan Literatur Biomedis**
Bagaskara Waskita Ahmad, Teguh Bharata Adji, S.T., M.T., M.Eng., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Penyakit jantung koroner kronis (CCD) berkontribusi sebagai beban terbesar pada penyakit kardiovaskular secara global, namun pelaksanaan sesuai panduan klinis pada praktik rutin masih rendah: hanya 1,0% dari 4.548 pasien pada studi INTERASPIRE yang mencapai pelaksanaan yang optimal. Sistem rekomendasi klinis dapat mengatasinya, tetapi harus dilakukan penyatuan rekam medis, panduan klinis, dan literatur biomedis secara sekaligus ke dalam satu struktur yang dapat dikueri, menjembatani terminologi klinis yang heterogen, serta menghasilkan keluaran dalam bahasa alami yang *faithfulness*-nya dapat diverifikasi. Penelitian ini mengembangkan sistem rekomendasi klinis CCD yang bersifat FHIR-native dan berbasis *knowledge graph* untuk menjawab ketiganya.

Rekam medis sintetis sejumlah 5.490 pasien dibangkitkan menggunakan Synthea, diekspor sebagai *bundle* HL7 FHIR R4, dimuat ke basis data graf Neo4j, lalu disatukan dengan panduan AHA/ACC CCD 2023 dan 4.801 abstrak PubMed menggunakan lapisan konsep bersama berbasis UMLS. Karena observasi numerik ikut dimuat sebagai data graf, rekomendasi yang disyaratkan pada nilai terukur—misalnya fraksi ejeksi ventrikel kiri atau indeks massa tubuh—dapat digerbangi menurut nilai pasien yang sebenarnya, yang tidak dapat dilakukan jika berbasis kondisi penyakitnya saja. Sistem juga menghasilkan *grounded briefing* yang faktanya ditelusuri secara deterministik dari graf dan model bahasa besar hanya merumuskan kalimatnya saja, lalu diperiksa secara otomatis. Evaluasi yang mencakup 876 pasien CCD dilakukan terhadap standar rujukan otomatis yang independen dari sistem.

Graf yang terbentuk memuat 1.804.379 simpul dan 8.796.425 relasi. Ablasi menunjukkan *observation gate* dapat menaikkan F_1 pada keselarasan terhadap panduan klinis dari 0,743 menjadi 0,810 melalui kenaikan presisi dari 0,612 menjadi 0,684 dengan *recall* yang praktis tidak berubah dan seluruh perolehan itu terpusat pada 15 dari 68 rekomendasi yang memang bergerbang. Standar rujukan otomatisnya sepakat dengan pelabelan manual terhadap 30 pasien pada 99,5% sel dengan Cohen's $\kappa = 0{,}991$ (interval kepercayaan 95% 0,984—0,996). Seluruh 30 *briefing* yang dibangkitkan lolos pemeriksaan *faithfulness* otomatis, tanpa satu pun rekomendasi maupun sitasi yang dikarang, poin wajib yang terlewat, ataupun klaim antikoagulan yang berlebih. Telaah oleh dokter spesialis jantung menilai 77,7% dari 247 rekomendasi bagi 15 pasien tepat, dengan penilaian belum tepat terpusat pada terapi antitrombotik.

Pemuatan observasi numerik sebagai data kelas satu karenanya menaikkan presisi keselarasan terhadap panduan klinis secara terukur. FHIR menyediakan observasi itu secara paling langsung, tetapi ekspor tabular pun dapat mereproduksi gerbang yang sama. Penelusuran yang deterministik menjadikan *faithfulness* pada tingkat sitasi dapat diputuskan secara mekanis, bukan aproksimatif seperti pada *retrieval augmented generation* (RAG), tetapi tidak menjamin bahwa setiap kalimat setia secara makna. Keterbatasannya pun dapat diukur: ekspor FHIR tidak memuat tanggal berhenti pengobatan sehingga logika panduan yang bergantung pada durasi hanya sebagian yang dapat diterapkan pada saat penelusuran dan seluruh angkanya berasal dari rekam medis sintetis.

Kata kunci : sistem rekomendasi klinis, *knowledge graph*, HL7 FHIR, penyakit jantung koroner kronis, keselarasan panduan

UNIVERSITAS
GADJAH MADA
**Pengembangan Sistem Rekomendasi Klinis untuk Penyakit Jantung Koroner Kronis Menggunakan Knowledge
Graph Berbasis Rekam Medis FHIR, Panduan Klinis, dan Literatur Biomedis**
Bagaskara Waskita Ahmad, Teguh Bharata Adji, S.T., M.T., M.Eng., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Chronic coronary disease (CCD) carries the largest single share of the global cardiovascular burden, yet guideline adherence in routine practice remains poor: of the 4,548 coronary patients in the INTERASPIRE study, only 1.0% met its definition of optimal adherence. A recommendation system can close that gap only if it solves three problems at once—fusing the patient record, the clinical guideline and the biomedical literature into one queryable structure, bridging the terminologies in which each is written, and generating natural-language output whose faithfulness can actually be checked against the source record. This thesis develops a FHIR-native, knowledge-graph-based recommendation system for CCD that addresses all three.

Synthetic records for 5,490 patients were generated with Synthea, exported as HL7 FHIR R4 bundles, loaded into a Neo4j property graph, and fused with the 2023 AHA/ACC CCD guideline and 4,801 PubMed abstracts through a UMLS concept layer. Because numeric observations are loaded as graph data rather than discarded, recommendations conditioned on measured values—ejection fraction or body-mass index, for instance—are gated on what the patient was actually measured at, which a condition-only pipeline cannot do. The system also produces a grounded briefing in which every fact is retrieved deterministically from the graph and the language model does nothing but phrase it, after which an automated check verifies the result. Evaluation covered 876 CCD patients, scored against a reference standard generated independently of the system.

The resulting graph holds 1,804,379 nodes and 8,796,425 relationships. An ablation shows that observation gating raises guideline concordance F_1 from 0.743 to 0.810, through precision rising from 0.612 to 0.684 with recall essentially unchanged, and the whole gain is confined to the 15 of 68 recommendations that are gated at all. The automated reference standard agrees with manual labelling of 30 patients on 99.5% of cells, Cohen's $\kappa = 0.991$ (95% CI 0.984–0.996, resampling patients). All 30 generated briefings passed the automated faithfulness check, with no fabricated recommendation or citation, no omitted required item, and no anticoagulation over-claim. A cardiologist judged 77.7% of 247 recommendations for 15 patients appropriate, with the judgements of inappropriateness concentrated on antithrombotic therapy.

Loading numeric observations as first-class data therefore raises the precision of guideline concordance measurably. FHIR supplies those observations most directly, but a tabular export reproduces the same gate. Deterministic retrieval makes the citation-level faithfulness of generated clinical text decidable rather than estimated, but it does not guarantee that every sentence is semantically faithful. The limits are measured too: the FHIR export carries no medication stop dates, so duration-dependent guideline logic can only be partly recovered, and every figure above comes from synthetic records.

Keywords : clinical recommendation system, knowledge graph, HL7 FHIR, chronic coronary disease, guideline concordance