

BAB I

PENDAHULUAN

A. Latar Belakang

Pelayanan kesehatan harus terus beradaptasi dengan perkembangan zaman, termasuk dalam penerapan sistem digital. Transformasi digital menjadi bagian penting dalam mempercepat ketersediaan data penyakit yang akurat. Fasilitas Kesehatan Tingkat Pertama (FKTP) yaitu puskesmas berperan dalam pelaksanaan transformasi tersebut melalui pengelolaan data yang terintegrasi (Permenkes RI No. 19, 2024). Satu diantara komponen penerapan transformasi digital di puskesmas adalah penggunaan Sistem Informasi Manajemen Puskesmas (SIMPUS) sebagai sarana pencatatan dan pelaporan elektronik. Melalui SIMPUS, data dapat terdokumentasi secara digital dan menjadi dasar penting untuk pengambilan keputusan (Permenkes RI No. 31, 2019).

Pengambilan keputusan dilakukan berdasarkan pengolahan data yang telah terdokumentasi digital. Oleh sebab itu, Perekam Medis dan Informasi Kesehatan (PMIK) memiliki peran penting dalam menjamin ketepatan data. PMIK dituntut untuk memiliki kemampuan dalam klasifikasi klinis dan kode diagnosis sesuai dengan *International Classification of Diseases and Related Health Problems, 10th Revision* (ICD-10). Akurasi dan ketepatan kode diagnosis menjadi salah satu aspek penting untuk memastikan data yang digunakan telah sesuai standar ICD-10, sehingga dapat digunakan dalam pembuatan suatu keputusan (Kepmenkes RI No. 312, 2020).

Temuan dilapangan terkait kondisi keakuratan masih belum optimal. Penelitian-penelitian di Indonesia melaporkan tingkat akurasi kode ICD-10 yang masih rendah dan bervariasi. Penelitian menemukan ketidakakuratan kode, dilatarbelakangi oleh pemberian kode yang dilakukan dokter tanpa pelatihan memadai untuk menetapkan diagnosis (Nahor, Putri and Simanjuntak, 2024). Penelitian selanjutnya menunjukkan ketidakakuratan kode diagnosis dan perlunya peningkatan kompetensi petugas dalam kode penyakit kronik secara spesifik (Ornela *et al.*, 2025). Selain itu, ketidaktepatan kode diagnosis ditemukan pada digit keempat dan kelima (Putri, Kusumo and Wulandari, 2024).

Ketidakakuratan kode diagnosis ditemukan juga pada penelitian lain dan menunjukkan pola permasalahan yang sama. Penelitian lain menunjukkan ketidakakuratan kode diagnosis (Nurmawati *et al.*, 2024). Ketidakakuratan kode diagnosis dilatarbelakangi oleh belum optimalnya pemanfaatan ICD-10 (Fitriana, Widowati and Sakinah, 2023). Kategori ketidakakuratan kode adalah pada penentuan subkategori kode diagnosis, dikarenakan sistem yang belum dilengkapi fitur deteksi kesalahan maupun validasi khusus untuk PMIK (Munandziroh *et al.*, 2024). Ketidakakuratan juga diperparah dengan tidak dilakukannya SOP pemberian kode di lapangan, serta belum adanya validasi akhir PMIK (Alvionita and Renata, 2025). Secara keseluruhan, penelitian-penelitian tersebut menegaskan bahwa ketidakakuratan kode diagnosis masih menjadi permasalahan pada pelayanan kesehatan.

Permasalahan tersebut tidak hanya berdampak pada ketepatan pencatatan rekam medis, tetapi juga menimbulkan konsekuensi yang lebih luas terhadap sistem kesehatan secara keseluruhan. Ketidakakuratan kode diagnosis memengaruhi laporan morbiditas yang digunakan sebagai dasar penetapan pola penyakit dan perencanaan program kesehatan di tingkat pelayanan primer. Penelitian menunjukkan bahwa ketidakakuratan berdampak langsung pada data morbiditas dalam pengambilan keputusan program kesehatan. (Nurmawati et al., 2024).

Fenomena yang telah dipaparkan menunjukkan adanya kesenjangan antara standar ketepatan kode dan kondisi nyata di lapangan. Oleh sebab itu, diperlukan terobosan untuk mengurangi kesalahan kode diagnosis. Salah satunya adalah pemanfaatan *Artificial Intelligence (AI)* untuk peningkatan akurasi dan konsistensi kode, efisiensi waktu, deteksi kesalahan otomatis, kemampuan pembelajaran berkelanjutan, dukungan terhadap pengambilan keputusan klinis, serta peningkatan kualitas data kesehatan secara keseluruhan (Baigi et al., 2025). Potensi penggunaan *AI* tidak hanya bersifat konseptual, tetapi juga telah dibuktikan melalui implementasi empiris. Validasi penerapan *AI-assisted coding* di Taiwan terhadap 2.632 rekam medis menunjukkan bahwa *AI* mampu meningkatkan konsistensi kode sekaligus menurunkan waktu kerja (Lu et al., 2024). Penelitian ini menunjukkan bahwa *AI* tidak hanya relevan sebagai alat bantu, tetapi juga efektif dalam meningkatkan kinerja kode diagnosis. Keberhasilan implementasi tersebut mendorong penelitian yang mengevaluasi *AI* dalam pengambilan keputusan medis.

Sejalan dengan hal tersebut, 4 jenis *AI* telah diuji dan *Claude* menunjukkan kinerja paling stabil dan unggul. Temuan memperlihatkan bahwa *Claude* memperoleh skor *relevance* sebesar 3,6 dan *completeness* sebesar 3,4, nilai yang lebih tinggi dibandingkan *ChatGPT* (2,3; 2,5), *Bard* (2,5; 2,2), dan *Perplexity* (3,0; 3,0) (Uppalapati and Nag, 2024). Selain itu, *Claude* mampu menghasilkan ringkasan medis dengan akurasi klinis mencapai 90 poin, dan mendekati akurasi tenaga kesehatan profesional (Jin *et al.*, 2024). Bahkan, penelitian tentang perbandingan beberapa *AI*, menunjukkan bahwa *Claude* mampu mencapai akurasi tertinggi 58,8% dan melebihi akurasi manusia yaitu sebesar 49,4% (Kaczmarczyk *et al.*, 2024). Berbagai temuan tersebut telah menunjukkan potensi *Claude* dalam meningkatkan akurasi kode diagnosis, meskipun hasilnya tetap harus dilakukan validasi oleh *expert judgment* (D'Urso *et al.*, 2025).

Mayoritas penelitian sebelumnya di Indonesia berfokus pada penilaian tingkat akurasi kode diagnosis, dan belum menguraikan secara rinci kelompok kesalahan dalam kode diagnosis. Sementara itu, penelitian mengenai pemanfaatan *AI* umumnya masih dilakukan di luar negeri, pada rumah sakit rujukan, atau dalam konteks penyusunan ringkasan medis, sehingga belum mencerminkan kondisi pelayanan primer di Indonesia. Keterbatasan tersebut menunjukkan bahwa *AI* belum banyak diuji dalam kode diagnosis berbasis ICD-10 pada puskesmas.

Puskesmas Umbulharjo I dipilih sebagai lokasi penelitian karena pada tahun 2024 Kecamatan Umbulharjo memiliki jumlah penduduk tertinggi di

Kota Yogyakarta, yaitu 72,20 jiwa (BPS Yogyakarta, 2025). Studi pendahuluan yang dilakukan di Puskesmas Umbulharjo I pada bulan November 2025 untuk rekam medis rawat jalan bulan September 2025. Hasil analisis dari 60 sampel rekam medis menunjukkan bahwa kode diagnosis manual memiliki tingkat ketidakakuratan sebanyak 35 rekam medis (57,38%) dan seluruh kesalahan berada pada subkategori. Sementara itu, kode *AI* menunjukkan akurasi yang lebih tinggi dengan ketidakakuratan hanya sebanyak 4 rekam medis (6,56%) dan seluruhnya memiliki kelompok kesalahan subkategori. Keterbatasan akses terhadap pedoman kodifikasi maupun perangkat digital dapat menjadi salah satu penyebab perbedaan hasil kode (Budiyantri, 2025). Oleh sebab itu, penelitian akan berfokus pada penggunaan *Claude* dalam pencarian kode sehingga dapat dibandingkan dan diidentifikasi kelompok kesalahan kode diagnosis.

B. Rumusan Masalah

Studi pendahuluan di Puskesmas Umbulharjo I menunjukkan adanya ketidakakuratan kode manual oleh puskesmas sebanyak 35 rekam medis (57,38%) dan kode *AI* oleh *Claude* sebanyak 4 rekam medis (6,56%). Dengan demikian, bagaimana perbedaan akurasi kode diagnosis secara manual dan *AI* serta identifikasi kelompok kesalahan kode diagnosis ICD-10 di Puskesmas Umbulharjo I Tahun 2025?

C. Tujuan Penelitian

1. Tujuan Umum

Penelitian bertujuan menganalisis perbedaan akurasi kode diagnosis ICD-10 dengan metode manual dan *AI*, serta mengidentifikasi kelompok kesalahan kode diagnosis berdasarkan validasi *expert judgment*.

2. Tujuan Khusus:

- a. Menganalisis perbedaan akurasi kode ICD-10 antara metode manual oleh Puskesmas Umbulharjo I dan metode *AI* oleh *Claude* berdasarkan validasi *expert judgment*.
- b. Mengidentifikasi kelompok kesalahan kode diagnosis dari kode tidak akurat pada metode manual oleh Puskesmas Umbulharjo I dan metode *AI* oleh *Claude*.

D. Ruang Lingkup

1. Ruang Lingkup Tempat

Puskesmas Umbulharjo I, Jalan Veteran 43, Kecamatan Umbulharjo, Kota Yogyakarta, Provinsi Daerah Istimewa Yogyakarta, 55163.

2. Ruang Lingkup Waktu

Pengumpulan data selama 2 bulan, yaitu bulan Februari-Maret 2026. Data yang dianalisis adalah rekam medis pasien rawat jalan tahun 2025.

3. Ruang Lingkup Materi:

- a. Analisis akurasi kode diagnosis berdasarkan ICD-10 antara metode manual oleh Puskesmas dan metode *AI* oleh *Claude*.
- b. Identifikasi kelompok kesalahan kode (kategori dan subkategori) dari kode tidak akurat pada masing-masing metode.
- c. Penelitian bersifat kuantitatif, dengan peneliti berperan dalam pemberian kode, penginputan diagnosis ke *Claude*, pengolahan data, identifikasi kelompok kesalahan, dan interpretasi data.

E. Manfaat Penelitian

Hasil penelitian yang akan dilakukan ini diharapkan dapat memberikan manfaat bagi berbagai pihak, yakni:

1. Manfaat Teoritis

Penelitian ini diharapkan mampu memberikan kontribusi pada pengembangan ilmu rekam medis dan informasi kesehatan, khususnya terkait akurasi kode diagnosis berdasarkan ICD-10 serta potensi penerapan teknologi *AI* di pelayanan kesehatan primer.

2. Manfaat Praktis

a. Bagi Puskesmas Umbulharjo I Yogyakarta:

- 1) Memberikan analisis mengenai tingkat akurasi kode diagnosis yang dilakukan secara manual dan dengan bantuan *AI*.
- 2) Menyediakan informasi yang dapat digunakan sebagai bahan pertimbangan dalam pengembangan sistem kode diagnosis.

b. Bagi Peneliti

Mengembangkan kemampuan peneliti dalam bidang rekam medis dan informasi kesehatan, terutama kode menggunakan ICD-10.

F. Keaslian Penelitian

Penelitian terkait perbandingan akurasi kode diagnosis antara metode manual dan *AI* di Indonesia belum pernah dilakukan sebelumnya. Adapun penelitian yang menjadi referensi acuan, diantaranya:

Tabel 1. Keaslian Penelitian

No.	Penulis dan Tahun	Judul	Metode Penelitian	Variabel Penelitian	Hasil	Persamaan	Perbedaan
1.	Yasir Abdelgadir, Charat Thongpra-yoon, Jing Miao Tahun 2024.	<i>AI integration in nephrology: evaluating ChatGPT for accurate ICD-10 documentation and coding.</i>	Kuantitatif	Akurasi ChatGPT.	Penelitian dilakukan pada rekam medis kasus nefrologi dan akurasi AI 87,6% dalam kode diagnosis nefrologi.	Penelitian ini memiliki persamaan mengevaluasi akurasi AI dalam kode ICD-10, menggunakan AI, dan validasi expert judgment.	Penelitian ini memiliki perbedaan karena spesifik bidang nefrologi, dilakukan di rumah sakit luar negeri, dan menggunakan ChatGPT sedangkan penelitian ini menggunakan Claude.

No.	Penulis dan Tahun	Judul	Metode Penelitian	Variabel Penelitian	Hasil	Persamaan	Perbedaan
2.	Vamsi Krishna Uppalapati, Deb Sanjay Nag Tahun 2024.	<i>A Comparative Analysis of AI Models in Complex Medical Decision Making: Scenarios Evaluating ChatGPT, Claude AI, Bard, and Perplexity.</i>	Kuantitatif	Akurasi AI.	Penelitian dilakukan pada rekam medis kasus umum 2023. <i>Claude</i> unggul dengan akurasi 82,4%, lebih tinggi dari <i>ChatGPT</i> (76,2%), <i>Bard</i> (73,5%), dan <i>Perplexity</i> (70,1%).	Penelitian ini memiliki persamaan menggunakan AI untuk analisis medis, serta mengukur akurasi.	Penelitian ini memiliki perbedaan karena fokus pada pengambilan keputusan medis, bukan pengodean ICD-10, tidak membandingkan AI dengan metode manual, serta bukan di pelayanan primer.
3.	Haijiao Jin, Jinglu Guo, Qisheng Lin, Shaun Wu, Weiguo Hu, Xiaoyang Tahun 2024.	<i>Comparative Study of Claude 3.5-Sonnet and Human Physicians in Generating Discharge Summaries for Patients with Renal Insufficiency.</i>	Kuantitatif	Akurasi <i>Claude</i> dan Manusia.	<i>Claude</i> mencapai skor akurasi 90 poin dan dokter 92 poin dengan efisiensi waktu 52% lebih cepat, yaitu 30 detik dan dokter 15 menit.	Penelitian ini memiliki persamaan menggunakan <i>Claude</i> untuk analisis medis, validasi <i>expert judgment</i> , serta fokus akurasi.	Penelitian ini memiliki perbedaan karena fokus pada pengambilan keputusan medis bukan pengodean ICD-10, hanya pada pasien kasus ginjal, serta bukan di pelayanan primer.
4.	Riza Suci Ernaman Putri, Retno Kusumo, Siti Wulandari Tahun 2024.	<i>Analysis of the accuracy of ICD-10 codes in outpatient diagnoses</i>	Kuantitatif	Akurasi kode diagnosis.	Pada 200 rekam medis pasien rawat jalan Januari dan Februari 2024 ditemukan	Penelitian ini memiliki persamaan dalam menganalisis akurasi kode ICD-10 di Pus-	Penelitian ini memiliki perbedaan karena tidak membahas lebih lanjut terkait identifikasi kelom-

No.	Penulis dan Tahun	Judul	Metode Penelitian	Variabel Penelitian	Hasil	Persamaan	Perbedaan
		<i>at the Kijang Health Center.</i>			ketidakakuratan sebesar 33,3% dengan kesalahan pada digit keempat dan kelima.	kesmas, fokus pasien rawat jalan, dan adanya validasi oleh <i>expert judgment</i> .	pok kesalahan, serta tidak menggunakan <i>AI</i> sehingga tidak ada perbandingan 2 metode.
5.	Munandziroh, Andri Asmoro-wati, Cahyono Rahadiyanto, Asih Prasetyowati Tahun 2024.	<i>Analysis of the Accuracy of Diagnosis Codes Based on ICD-10 with the Application of the 4th Character at the Puskesmas of Semarang City.</i>	Kuantitatif	Akurasi kode diagnosis.	Pada 450 rekam medis pasien rawat jalan ditemukan ketidakakuratan sebesar 72,77% dengan kesalahan terbanyak pada digit keempat dan kelima.	Penelitian ini memiliki persamaan dalam menganalisis akurasi kode ICD-10 di Puskesmas, lalu fokus pasien rawat jalan.	Penelitian ini memiliki perbedaan karena tidak menggunakan <i>AI</i> , fokus keterangan kesalahan hanya pada karakter ke 4 dan 5 dan tidak ada pengelompokkan spesifik, serta tidak menggunakan validasi oleh <i>expert judgment</i> .