

ANALISIS AKURASI KODE DIAGNOSIS SECARA MANUAL DAN *ARTIFICIAL INTELLIGENCE* DI PUSKESMAS UMBULHARJO I

Rezqa Amaliza¹⁾, Nita Budiyantri²⁾, Syarah Mazaya Fitriana³⁾,
Abdul Hadi Kadarusno⁴⁾, Mutiara Pertiwi⁵⁾, Rahmah Nindyakinanti⁶⁾

^{1),2),3),4)}*Prodi DIII Rekam medis dan Informasi Kesehatan Poltekkes Kemenkes Yogyakarta*

^{5),6)}*Puskesmas Umbulharjo I*

Email : rezqamaliza@gmail.com

ABSTRAK

Ketidakakuratan kode diagnosis ICD-10 di puskesmas berdampak pada kualitas data morbiditas dan pelaporan penyakit, dengan 72,77% kesalahan kode. *Artificial Intelligence (AI)* terbukti mampu meningkatkan akurasi kodefikasi. Penelitian ini menganalisis perbedaan akurasi kode diagnosis ICD-10 antara metode manual dan *AI* serta identifikasi kelompok kesalahan pemberian kode di Puskesmas Umbulharjo I. *AI* yang digunakan dalam menetapkan kode adalah *Claude*, hasil penetapan tersebut dilanjutkan dengan mengidentifikasi kelompok kesalahan kode. Pendekatan kuantitatif deskriptif dengan desain *cross sectional* menggunakan sampel 398 rekam medis rawat jalan tahun 2025 melalui metode *simple random sampling*. Hasil kode manual dan kode *AI* divalidasi oleh *expert judgment* sesuai ICD-10 Volume II 2010. Hasil penelitian pemberian kode manual oleh Puskesmas Umbulharjo I diperoleh keakuratan sejumlah 93 rekam medis (23,37%), sedangkan pemberian kode menggunakan *AI* diperoleh keakuratan sejumlah 345 rekam medis (86,68%). Hasil identifikasi kesalahan kode manual didominasi kelompok kesalahan subkategori sejumlah 291 rekam medis (95,41%) dan kelompok kesalahan kategori 14 rekam medis (4,59%), sedangkan kesalahan kode menggunakan *AI* meliputi kelompok kesalahan subkategori 53 rekam medis (100%). Hasil *penelitian* mengindikasikan kesenjangan akurasi signifikan antara kode manual dan *AI*, sehingga mempertegas potensi *AI* sebagai alat bantu kodefikasi dengan verifikasi akhir oleh PMIK.

Kata Kunci: Diagnosis, Pemberian kode, *Artificial Intelligence*, *Claude*, Kesalahan

ABSTRACT

Inaccurate ICD-10 diagnosis coding in primary healthcare centers (puskesmas) negatively impacts morbidity data quality and disease reporting, with an error rate of 72.77% (Munandziroh et al., 2024). Artificial Intelligence (AI) has been shown to improve coding accuracy. This study analyzes the difference in ICD-10 diagnosis coding accuracy between manual and AI-based methods, and identifies error categories in Puskesmas Umbulharjo I. Claude was used as the AI coding tool, followed by error group identification. A descriptive quantitative approach with a cross-sectional design was applied, using 398 outpatient medical records from 2025 selected through simple random sampling. Both manual and AI-generated codes were validated by expert judgment based on ICD-10 Volume II 2010. Results showed manual coding accuracy at 93 records (23.37%), while AI coding achieved 345 records (86.68%). Manual coding errors were predominantly subcategory errors (291 records, 95.41%) and category errors (14 records, 4.59%), whereas AI coding errors consisted entirely of subcategory errors (53 records, 100%).

These findings indicate a significant accuracy gap between manual and AI coding, reinforcing AI's potential as a coding support tool with final verification by medical record professionals (PMIK).

Keywords: Diagnosis, Coding, Accuracy, Artificial Intelligence, Claude, Error.

PENDAHULUAN

Pelayanan kesehatan dituntut untuk terus beradaptasi seiring perkembangan zaman, termasuk melalui penerapan sistem digital di fasilitas kesehatan tingkat pertama. Salah satu wujud transformasi digital di puskesmas adalah penggunaan Sistem Informasi Manajemen Puskesmas (SIMPUS) sebagai sarana pencatatan dan pelaporan untuk mendukung pengambilan keputusan (Permenkes RI No. 31, 2019). Dalam konteks ini, Perekam Medis dan Informasi Kesehatan (PMIK) memegang peran strategis dalam menjamin ketepatan data, khususnya melalui kemampuan klasifikasi klinis dan pemberian kode diagnosis sesuai *International Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)*. Akurasi kode diagnosis merupakan aspek krusial agar data yang dihasilkan memenuhi standar dan dapat digunakan sebagai dasar pengambilan keputusan yang valid (Kepmenkes RI No. 312, 2020).

Kondisi di lapangan menunjukkan tingkat keakuratan kode diagnosis masih jauh dari optimal. Berbagai penelitian di Indonesia melaporkan akurasi kode ICD-10 yang rendah dan bervariasi, yang disebabkan oleh beberapa faktor, antara lain penetapan diagnosis oleh dokter tanpa pelatihan pemberian kode (Nahor et al., 2024), sistem yang belum dilengkapi fitur validasi khusus untuk PMIK (Munandziroh et al., 2024), serta

tidak dijalankannya prosedur operasional standar pemberian kode dan absennya validasi akhir oleh PMIK (Alvionita & Renata, 2025). Temuan-temuan tersebut secara kolektif menegaskan bahwa ketidakakuratan kode diagnosis masih menjadi permasalahan dalam pelayanan kesehatan di Indonesia.

Kesenjangan antara standar ketepatan kode dan realita lapangan mendorong perlunya terobosan inovatif. Pemanfaatan *Artificial Intelligence (AI)* dinilai berpotensi meningkatkan akurasi kode, efisiensi waktu, deteksi kesalahan secara otomatis, serta kualitas data kesehatan secara keseluruhan (Mousavi Baigi & al., 2025). Sejalan dengan hal tersebut, hasil pengujian terhadap empat jenis *AI* menunjukkan bahwa *Claude* memperoleh kinerja paling unggul dengan skor *relevance* sebesar 3,6 dan *completeness* sebesar 3,4, melampaui *ChatGPT* (2,3; 2,5), *Bard* (2,5; 2,2), dan *Perplexity* (3,0; 3,0) (Uppalapati & Nag, 2024). Temuan ini memberikan landasan kuat untuk mengeksplorasi penerapan *AI*, khususnya *Claude*, sebagai alat bantu meningkatkan akurasi pemberian kode diagnosis di puskesmas.

Validasi empiris penerapan *AI-assisted coding* di Taiwan terhadap 2.632 rekam medis menunjukkan bahwa *AI* mampu meningkatkan konsistensi kode sekaligus menurunkan waktu kerja (Lu & al., 2025). Lebih lanjut, *Claude*

mampu menghasilkan ringkasan medis dengan akurasi klinis mencapai 90 poin dan mendekati akurasi tenaga kesehatan profesional (Jin et al., 2024). Pada perbandingan beberapa *AI*, *Claude* bahkan mencapai akurasi tertinggi sebesar 58,8%, melampaui akurasi manusia yang hanya sebesar 49,4% (Kaczmarczyk & al., 2024). Meskipun demikian, hasil pemberian kode oleh *AI* tetap harus divalidasi 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 penelitian mengenai pemanfaatan *AI* umumnya masih dilakukan di luar negeri atau rumah sakit, sehingga belum mencerminkan kondisi pelayanan primer di Indonesia.

Puskesmas Umbulharjo I dipilih sebagai lokasi penelitian karena pada tahun 2024 Kecamatan Umbulharjo memiliki jumlah penduduk tertinggi di Kota Yogyakarta, yaitu 72.200 jiwa (BPS Yogyakarta, 2025). Studi pendahuluan yang dilakukan di Puskesmas Umbulharjo I pada bulan November 2025 untuk rekam medis rawat jalan bulan September 2025 menunjukkan bahwa dari 60 sampel rekam medis, kode diagnosis manual memiliki tingkat ketidakakuratan sebanyak 35 rekam medis (57,38%), dengan 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 (Budiyanti, 2025). Berdasarkan

temuan tersebut, penelitian ini bertujuan menganalisis perbedaan akurasi kode diagnosis ICD-10 dengan metode manual dan *AI* menggunakan *Claude*, serta mengidentifikasi kelompok kesalahan kode diagnosis berdasarkan validasi *expert judgment* di Puskesmas Umbulharjo I Tahun 2025.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan desain *cross sectional*. Populasi penelitian adalah seluruh Rekam Medis Elektronik (RME) rawat jalan Puskesmas Umbulharjo I periode 1 Januari-31 Desember 2025 yang berjumlah 68.995 rekam medis. Sampel berjumlah 398 rekam medis yang diperoleh melalui *simple random sampling*.

Variabel penelitian meliputi akurasi kode diagnosis ICD-10 (metode manual dan metode *AI*) serta kelompok kesalahan kode (kategori dan subkategori). Kode manual diperoleh langsung dari SIMPUS Puskesmas Umbulharjo I. Kode *AI* diperoleh dengan memasukkan data *file* berisikan diagnosis, anamnesis, dan kode puskesmas ke *Claude*, disertai *prompt*:

“Anda adalah asisten kodifikasi medis profesional yang ahli dalam sistem klasifikasi ICD-10 tahun 2010. Tugas Anda adalah melakukan kode diagnosis medis dengan standar kualitas tinggi sesuai pedoman ICD-10 tahun 2010 dengan instruksi kode:

- a. Bacalah dengan teliti setiap baris data yang berisi informasi pasien meliputi nomor pasien, jenis kelamin, usia, dan diagnosis.
- b. Lakukan kode dengan memberikan kode ICD-10 paling spesifik. Jika diagnosis tidak memiliki keterangan lanjut, maka gunakan kode *unspecified* yaitu poin 9.
- c. Lakukan pengodean kombinasi, jika beberapa diagnosis memiliki kode yang bisa dijadikan 1 kode.
- d. Lengkapi kode blok S dan T dengan kode *external cause*.
- e. Sajikan hasil kode dalam bentuk tabel seperti *file* sebelumnya dan kamu bisa menambahkan kolom kode *Claude*.
- f. Periksa kembali setiap kode untuk memastikan kesesuaian antara diagnosis dan kode ICD-10 yang dipilih, serta tidak ada kesalahan dalam penulisan kode ICD-10.

Lakukan kode pada seluruh data dari *file* yang diberikan dengan mengikuti instruksi kode. Pastikan hasil kode konsisten, akurat, dan dapat dipertanggungjawabkan secara administratif.”

Kode standar ditetapkan peneliti melalui *expert judgment* dan digunakan sebagai acuan perbandingan keakuratan kedua metode. Pengumpulan data dilakukan melalui observasi partisipasi pasif untuk mengamati

penggunaan SIMPUS dan SOP pemberian kode yang berlaku di puskesmas. Selain itu, peneliti juga menghitung persentase keakuratan dan kelompok kesalahan.

HASIL PENELITIAN

1. Akurasi kode ICD-10 antara Metode Manual oleh Puskesmas Umbulharjo I dan Metode AI oleh Claude

Hasil observasi menunjukkan bahwa alur pemberian kode di Puskesmas Umbulharjo I tidak sesuai dengan SOP yang ditetapkan. Dokter langsung menginputkan kode ke SIMPUS tanpa menelusuri ICD-10 *Volume III* dan memverifikasinya melalui *Volume I*. SIMPUS yang digunakan tidak dilengkapi fitur peringatan kode tidak spesifik maupun fitur validasi untuk PMIK, serta sering mengalami gangguan *loading*.

Analisis keakuratan terhadap 398 rekam medis menggunakan kode ICD-10 disajikan pada Tabel 1.

Tabel 1. Analisis Keakuratan Kode ICD-10 pada 2 Metode di Puskesmas Umbulharjo I

| Analisis Keakuratan | Kode Manual | Kode AI |
|---------------------|------------------|------------------|
| | Jumlah (%) | Jumlah (%) |
| Akurat | 93 (23,37) | 345 (86,68) |
| Tidak Akurat | 305 (76,63) | 53 (13,32) |
| Total | 398 (100) | 398 (100) |

Tabel 1 menunjukkan kode manual hanya akurat pada 93 rekam medis (23,37%), sementara 305 rekam medis (76,63%) tidak akurat. Sebaliknya, kode AI yang dihasilkan Claude menunjukkan 345 rekam medis akurat (86,68%) dan hanya 53 rekam medis tidak akurat (13,32%).

Claude memberikan kode diagnosis tambahan dan melakukan kombinasi kode sesuai kaidah ICD-10 Tahun 2010, yang ditunjukkan pada tabel 2 dan 3.

Tabel 2. Penambahan Kode *External Cause* pada RME di Puskesmas Umbulharjo I

| Diagnosis | Kode Manual | Kode AI | Kode ICD-10 |
|---|-------------|-----------------|-----------------|
| <i>Injury of unspecified body region</i> | T14 | T14.9 V23.49 | T14.9 V23.49 |
| <i>Superficial injuries involving multiple body regions</i> | T00 | T00.9 V23.49 | T00.9 V23.49 |
| <i>Contusion of other parts of wrist and hand</i> | S60.2 | S60.2 V29.49 | S60.2 V29.49 |

Tabel 2 menunjukkan 3 rekam medis cedera yang tidak dilengkapi kode *external cause* oleh pemberian kode secara manual. *Claude* secara otomatis menambahkan kode *external cause* untuk kendaraan bermotor (V23.49 dan V29.49) yang tidak dikode pada kode manual, sehingga ketiga rekam medis tersebut dinyatakan akurat pada kode AI dan tidak akurat pada kode manual.

Tabel 3. Penggabungan Kode Kombinasi pada RME di Puskesmas Umbulharjo I

| Diagnosis | Kode Manual | Kode AI | Kode ICD-10 |
|--|-------------|---------|-------------|
| <i>Myalgia</i> | M79.1 | M79.13 | M79.13 |
| <i>Essential primary hypertension</i> | I10 I11 | I11.9 | I11.9 |
| <i>Hypertensive heart disease without congestive heart failure</i> | | | |

| Diagnosis | Kode Manual | Kode AI | Kode ICD-10 |
|--|---------------------|----------------|----------------|
| <i>Essential primary hypertension</i> | I10 B35.6 I64 | B35.6 I64 | B35.6 I64 |
| <i>Tinea cruris</i> | | | |
| <i>Stroke, not specified as haemorrhage or infaction</i> | | | |
| <i>Essential primary hypertension</i> | I10 I20.0 | I20.0 | I20.0 |
| <i>Unstable angina</i> | | | |
| <i>Essential primary hypertension</i> | I10 I50.0 | I11.0 | I11.0 |
| <i>Congestive heart failure</i> | | | |
| <i>Chronic kidney disease, stage 3</i> | N18.3 I10 E11 | I12.0 E11.8 | I12.0 E11.8 |
| <i>Essential primary hypertension</i> | | | |
| <i>Non-insulin-dependent diabetes mellitus</i> | | | |
| <i>Essential primary hypertension</i> | I10 N18 | I12.0 | I12.0 |
| <i>Chronic kidney disease</i> | | | |
| <i>Atherosclerotic cardiovascular disease</i> | I25 I10 I20 | I25.8 | I25.8 |
| <i>Essential primary hypertension</i> | | | |
| <i>Angina pectoris</i> | | | |

Tabel 3 menunjukkan *Claude* secara otomatis melakukan kombinasi pada kode-kode diagnosis tersebut, sehingga ketujuh rekam medis tersebut dinyatakan akurat pada kode AI dan tidak akurat pada kode manual.

2. Kelompok Kesalahan Kode Diagnosis pada Metode Manual dan AI Rekam Medis Rawat Jalan Puskesmas Umbulharjo I Tahun 2025

Dari 305 kode tidak akurat pada metode manual dan 53 kode tidak akurat pada metode AI, disimpulkan kelompok kesalahan yang terdiri atas kesalahan kategori (kode tiga digit tidak sesuai diagnosis, termasuk pengabaian kode kombinasi dan *external cause*) dan kesalahan subkategori (kategori benar, tapi digit keempat atau kelima keliru atau tidak digunakan). Analisis kelompok kesalahan disajikan pada Tabel 4.

Tabel 4. Analisis Kelompok Kesalahan Kode Dua Metode di Puskesmas Umbulharjo I

| Analisis Kelompok Kesalahan | Kode Manual | Kode AI |
|-----------------------------|------------------|-----------------|
| | Jumlah (%) | Jumlah (%) |
| Kategori | 14 (4,59) | 0 (0,00) |
| Subkategori | 291 (95,41) | 53 (100) |
| Total | 305 (100) | 53 (100) |

Tabel 4 menunjukkan kode manual mengalami kelompok kesalahan kategori sejumlah 14 rekam medis (4,59%) dan kelompok kesalahan kategori sejumlah 291 rekam medis (95,41%). Sedangkan AI tidak mengalami satupun kesalahan kategori (0%), dan seluruh kesalahan berada pada kelompok kesalahan subkategori sejumlah 53 rekam medis (100%). Rincian lima besar kesalahan kode manual disajikan pada Tabel 5.

Tabel 5. Lima Besar Kesalahan Kode Manual dari RME Puskesmas Umbulharjo I

| Diagnosis | Kode Manual | Kode ICD-10 | Total RM |
|---|-------------|-------------|------------|
| <i>Non-insulin-dependent Diabetes Mellitus</i> | E11 | E11.8 | 100 |
| <i>Disorders of Lipoprotein Metabolism and Other Lipidaemias (Dislipidemia)</i> | E78 | E78.8 | 33 |
| <i>Acute Upper Respiratory Infections of Multiple and Unspecified Sites</i> | J06 | J06.9 | 13 |
| <i>Gonarthrosis [Arthrosis of Knee]</i> | M17 | M17.9 | 13 |
| <i>Disorders of Purine and Pyrimidine Metabolism</i> | E79 | E79.9 | 10 |

Tabel 5 menunjukkan lima besar kesalahan kode manual seluruhnya merupakan kesalahan subkategori. Penggunaan E11 tanpa subkategori menjadi kesalahan terbanyak dengan 100 rekam medis terdampak. Selanjutnya adalah rincian lima besar kesalahan kode AI disajikan pada Tabel 6 sebagai berikut.

Tabel 6. Lima Besar Kesalahan Kode AI dari RME di Puskesmas Umbulharjo I

| Diagnosis | Kode AI | Kode ICD-10 | Total RM |
|--|---------|-------------|-----------|
| <i>Disorders of Lipoprotein Metabolism and Other</i> | E78.9 | E78.8 | 31 |

| Diagnosis | Kode AI | Kode ICD-10 | Total RM |
|---|---------|-------------|----------|
| <i>Lipidaemias (Dislipidemia)</i> | | | |
| <i>Other Dermatitis</i> | L30.9 | L30.8 | 5 |
| <i>Myalgia</i> | M79.1x | M79.1x | 4 |
| <i>Diarrhoea and Gastroenteritis of Presumed Infectious Origin</i> | A09 | A09.0 | 4 |
| <i>Disorders of Lipoprotein Metabolism</i> | Z71.9 | Z71.8 | 2 |

Tabel 6 menunjukkan seluruh kesalahan *AI* bersifat subkategori sedangkan *Claude* menetapkan kode *unspecified* (.9), kondisi ini terjadi ketika anamnesis berisi informasi singkat, seperti pada kasus Dislipidemia yang memberikan dampak total kesalahan terbesar *AI*.

PEMBAHASAN

1. Akurasi kode ICD-10 antara Metode Manual oleh Puskesmas Umbulharjo I dan Metode *AI* oleh *Claude*

Pedoman ICD-10 *Volume* II Tahun 2010 menetapkan bahwa alur pemberian kode yang benar diawali dengan penelusuran istilah diagnosis melalui *Volume* III kemudian diverifikasi melalui *Volume* I, seluruh diagnosis yang memengaruhi perawatan harus dikodekan, dan setiap petunjuk yang ada harus diikuti (WHO, 2010). Rekam medis memuat diagnosis lengkap dan akurat beserta kodenya (Permenkes RI No. 24, 2022), dan PMIK memiliki peran utama dalam menunjang hal tersebut didukung oleh kompetensi kelima yang mencakup penggunaan ICD-10 secara tepat (Kepmenkes RI No. 312, 2020).

Penggunaan SIMPUS bertujuan menjamin ketersediaan data dan informasi berkualitas, berkesinambungan, serta mudah diakses (Permenkes RI No. 19, 2024).

Hasil penelitian menunjukkan SOP pemberian kode tidak terlaksana di Puskesmas Umbulharjo I. Berdasarkan observasi, dokter menginputkan kode ke SIMPUS tanpa penelusuran ICD-10, sehingga menghasilkan kode manual yang cenderung berhenti di kelompok kategori. Menurut analisis peneliti, situasi ini melanggar pedoman ICD-10 Tahun 2010, karena pemberian kode diagnosis tidak sesuai SOP, padahal telah sesuai alurnya dengan di pedoman ICD-10 Tahun 2010. Oleh karena itu kode-kode yang dihasilkan tidak akurat dan tidak lengkap, sehingga melanggar Permenkes RI No. 24 Tahun 2022. Selain itu, pemberian kode diagnosis bukan tugas dokter, dan seharusnya merupakan kompetensi ke lima dari PMIK, yang menunjukkan pelanggaran dalam Kepmenkes RI No. HK.01.07/MENKES/312/2020.

Kondisi ini dipengaruhi oleh keterbatasan SIMPUS yang tidak memiliki fitur peringatan kode tidak spesifik maupun validasi untuk PMIK, serta gangguan *loading* yang sering terjadi. SIMPUS juga mengalami gangguan yang menyebabkan data-data sulit diakses dan menunjukkan pelanggaran Permenkes RI No. 31 Tahun 2019. Temuan serupa dilaporkan penelitian di salah satu puskesmas Indonesia, bahwa pemberian kode oleh tenaga selain PMIK dan keterbatasan Sistem Informasi Puskesmas akan

menghasilkan akurasi kode yang rendah (Mayang Sari et al., 2026).

Hasil penelitian menunjukkan akurasi kode manual hanya 23,37% sedangkan kode *AI* yang dihasilkan *Claude* mencapai 86,68%. Menurut analisis peneliti, *Claude* secara konsisten menambahkan kode *external cause* kendaraan bermotor (V23.49 dan V29.49) pedoman ICD-10 *Volume II* Tahun 2010, yang absen pada kode manual. *Claude* secara konsisten menerapkan kode kombinasi sesuai pedoman ICD-10 *Volume II* Tahun 2010 pada seluruh kasus. Misalnya, ketika kode I10 dan I50.0 yang ditulis terpisah pada kode manual, *Claude* justru mengkombinasi menjadi I11.0, dan kode N18.3 dengan I10 dikombinasi *Claude* menjadi I12.0.

Menurut analisis peneliti, pengabaian kode *external cause* berdampak langsung pada statistik morbiditas karena data penyebab luar cedera tidak tercatat, sehingga laporan epidemiologi kehilangan informasi yang seharusnya menjadi dasar kebijakan pencegahan kecelakaan. Di sisi lain, pengabaian kode kombinasi menghasilkan penghitungan ganda dalam laporan 10 besar penyakit yang mendistorsi statistik morbiditas. Kedua permasalahan ini merupakan dampak langsung dari tidak adanya verifikasi Perkam Medis dan Informasi Kesehatan (PMIK) sebagai penjaga kode akhir, serta diperparah SIMPUS yang tidak memiliki peringatan otomatis untuk mencegah kesalahan pemberian kode dan fitur validasi untuk PMIK.

Penelitian menemukan *AI* mampu meningkatkan akurasi pemberian kode secara

signifikan (Dai et al., 2024), dan dapat berperan sebagai alat bantu, namun tidak menggantikan peran PMIK (Mustafa et al., 2025). Penelitian lain juga menunjukkan bahwa kesalahan kode kombinasi sering ditemui (Susanah et al., 2025).

Perbaikan yang dibutuhkan bukan sekadar pembaruan SOP, melainkan perubahan alur kerja yang menempatkan PMIK sebagai penjaga gerbang kode akhir, didukung sistem informasi yang secara aktif mendorong pemberian kode yang akurat dengan penambahan peringatan otomatis saat kode berhenti dikategori, serta fitur validasi untuk PMIK.

2. Kelompok Kesalahan Kode Diagnosis pada Metode Manual dan *AI* Rekam Medis Rawat Jalan Puskesmas Umbulharjo I Tahun 2025

Pedoman ICD-10 *Volume II* Tahun 2010 menunjukkan bahwa kode-kode diagnosis harus diberikan lengkap. Oleh karena itu, dikelompokkan menjadi dua kelompok kesalahan kode, kesalahan kategori, yaitu kode tiga digit yang dipilih tidak sesuai dengan diagnosis yang ditegakkan, termasuk ketika kode kombinasi atau kode *external cause* diabaikan, dan kesalahan subkategori, yaitu kategori tiga digit sudah benar tetapi digit keempat atau kelima tidak digunakan atau keliru dipilih (WHO, 2010).

Hasil penelitian menunjukkan dari 305 kode tidak akurat pada manual, 291 rekam medis (95,41%) memiliki kelompok kesalahan subkategori dan 14 rekam medis (4,59%) memiliki kelompok kesalahan kategori.

Sedangkan kode *AI* yang dihasilkan *Claude* hanya mengalami kelompok kesalahan subkategori, yaitu pada 53 rekam medis (100%). Hasil penelitian menunjukkan lima besar kesalahan kode manual yang didominasi oleh penetapan kode E11. Menurut analisis peneliti, petugas puskesmas yang menghasilkan kode manual telah memahami kondisi klinis pasien, namun pemilihan kode tidak spesifik, serta tidak menggunakan aturan penambahan kode *external cause* dan juga kombinasi kode. Temuan ini sejalan dengan penelitian bahwa E11 merupakan kasus dengan kesalahan kode terbanyak di salah satu puskesmas Indonesia (Ornela et al., 2025).

Lima besar kesalahan kode *AI* yang diberikan oleh *Claude* dan didominasi oleh E78.9 yang ditemukan pada 31 rekam medis. Menurut analisis peneliti, *AI* sudah mampu memberikan kode spesifik, kode *external cause*, ataupun melakukan kombinasi kode. Tetapi, *AI* masih rentan pada level subkategori, terutama Ketika kondisi ambigu, saat anamnesis terlalu singkat untuk membedakan antara *other specified* (.8) dan *unspecified* (.9). Penelitian internasional mendukung temuan ini, bahwa *AI* memerlukan anamnesis yang kaya untuk menghasilkan kode akurat pada level subkategori (Karila Cohen et al., 2025).

Secara keseluruhan, *Claude* mampu menerapkan aturan kode kombinasi dan melengkapi kode *external cause*, namun memiliki keterbatasan pada subkategori ketika anamnesis singkat. Hal ini menegaskan bahwa *AI* berperan sebagai *filter* pertama yang meningkatkan akurasi

pemberian kode, sementara PMIK tetap menjadi validator akhir pada kondisi ambigu.

PENUTUP

Metode pemberian kode berpengaruh terhadap akurasi dan kelompok kesalahan kode ICD-10. Metode *AI* menggunakan *Claude* menghasilkan akurasi 345 rekam medis (86,68%) dan hanya mengalami kesalahan subkategori, sedangkan metode manual hanya mencapai 93 rekam medis (23,37%) dengan kesalahan kategori maupun subkategori akibat ketidakpatuhan terhadap SOP ICD-10. Tingginya akurasi *Claude* menjadikannya potensial sebagai alat bantu pemberian kode, namun verifikasi akhir oleh PMIK tetap diperlukan. Oleh karena itu, disarankan agar Puskesmas Umbulharjo I merevisi SOP dengan menetapkan PMIK sebagai verifikator kode akhir, pengembang SIMPUS menambahkan fitur peringatan kode tidak spesifik dan validasi untuk PMIK, serta peneliti selanjutnya membandingkan beberapa *AI* atau variasi *prompt* di fasilitas kesehatan yang berbeda.

DAFTAR PUSTAKA

- Alvionita, C. V., & Renata, E. C. (2025). Analisis Kesesuaian SOP Pengodean dan Tindakan Berdasarkan ICD di RS Universitas Brawijaya. *Jurnal Informasi Kesehatan Indonesia*, 11(2), 132–139. <https://ojs.poltekkes-malang.ac.id/JIKI/article/download/5760/1121>
- BPS Yogyakarta. (2025). *BPS Yogyakarta Tahun 2025* (D. R. Wati & F. T. Hastuti, Eds.; Vol. 46). <https://jogjakota.bps.go.id/id/publication/2026/02/27/59d960318da3438db4a83474/kota-yogyakarta-dalam-angka-2026.html>
- Budiyanti, N. (2025). Perancangan website INCLADIS sebagai alat bantu dalam

- menunjang akurasi kodifikasi penyakit sesuai ICD 10 dan ICD 11. *Health Education Laboratory*, 1(2), 41–53. <https://e-journal.poltekkesjogja.ac.id/index.php/HEL-AB/article/download/2785/1266>
- Dai, H., Wang, C., Chen, C., Liou, C., Lu, A., Lai, C., Shain, B., Ke, C., Wang, W., Mir, T., Simanjuntak, M., Kao, H., Tsai, M., & Tseng, V. (2024). Evaluating a Natural Language Processing–Driven, AI-Assisted International Classification of Diseases, 10th Revision, Clinical Modification, Coding System for Diagnosis Related Groups in a Real Hospital Environment: Algorithm Development and Validation Study. *Journal of Medical Internet Research*, 26, e58278. <https://doi.org/10.2196/58278>
- D’Urso, F., Paladini, F., Pollini, M., & Broccolo, F. (2025). Use of Artificial Intelligence Chatbots in Interpretation of Clinical Chemistry and Laboratory Medicine Reports: A Standardized Approach. *Applied Sciences (Switzerland)*, 15(8). <https://www.mdpi.com/2076-3417/15/8/4232/pdf?version=1744621229>
- Jin, H., Guo, J., Lin, Q., Wu, S., Hu, W., & Li, X. (2024). Comparative study of Claude 3.5-Sonnet and human physicians in generating discharge summaries for patients with renal insufficiency: assessment of efficiency, accuracy, and quality. *Frontiers in Digital Health*, 6. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11655460/pdf/fdgth-06-1456911.pdf>
- Kaczmarczyk, R., & al., et. (2024). Evaluating multimodal AI in medical diagnostics. *Npj Digital Medicine*, 7, 205. <https://doi.org/10.1038/s41746-024-01208-3>
- Karila Cohen, J., Garcelon, N., & Quennelle, S. (2025). Assigning Clinical Codes with Semi-Automated Medical Language Processing. *Studies in Health Technology and Informatics*, 328, 434–438. <https://doi.org/10.3233/SHTI250754>
- Kepmenkes RI No. 312. (2020). *Keputusan Menteri Kesehatan Republik Indonesia Nomor HK.01.07/MENKES/312/2020 tentang Standar Profesi Perkam Medis dan Informasi Kesehatan*. <https://peraturan.bpk.go.id/Download/174833/KMK%20Nomor%20312%20Tahun%202020.pdf>
- Lu, A. T., & al., et. (2025). Application of Clinical Department–Specific AI-Assisted Coding Using Taiwan Diagnosis-Related Groups: Retrospective Validation Study. *JMIR Human Factors*, 12, e59961. <https://doi.org/10.2196/59961>
- Mayang Sari, S., Ermasari Muzuh, M., Amiruddin, A., Safitria Rahmadi, D., & Dzulkarnain, K. (2026). Analysis of the Accuracy of ICD-10 Diagnosis Coding and Its Contributing Factors in Outpatient Medical Records at a Primary Health Center in Baubau City. *Jurnal Promotif Preventif*, 9(1), 60–69. <https://journal.unpacti.ac.id/index.php/JPP/article/view/2555/1313>
- Mousavi Baigi, S. F., & al., et. (2025). Artificial Intelligence-based Automated International Classification of Diseases Coding: A Systematic Review. *Journal of Medical Signals & Sensors*, 15, 22. https://doi.org/10.4103/jmss.jmss_76_24
- Munandziroh, Asmorowati, A., Rahadiyanto, C., & Prasetyowati, A. (2024). Analysis of the Accuracy of Diagnosis Codes Based on ICD-10 with the Application of the 4th Character at the Puskesmas of Semarang City. *International Scientific Meeting on Health Information Management*, 6. <https://pels.umsida.ac.id/index.php/PELS/article/view/1931/1255>
- Mustafa, A., Naseem, U., & Rahimi Azghadi, M. (2025). Large language models vs human for classifying clinical documents. *International Journal of Medical Informatics*, 195. <https://doi.org/10.1016/j.ijmedinf.2025.105800>
- Nahor, A. B., Putri, R., & Simanjuntak, S. (2024). Analisis Ketepatan Kode Diagnosa Penyakit

- Terkait di Puskesmas Kecamatan Makasar, Jakarta Timur. *Jurnal Informatika Rekam Medis Dan Kesehatan Indonesia*, 10(1). <https://journal.khj.ac.id/index.php/jirmiki/article/view/87>
- Ornela, O., Fadhila, W., Leonar, D., & Yasli, Z. (2025). Tinjauan Ketepatan Kode Diagnosis ICD-10 Diabetes Melitus Pada Rekam Medis Rawat Jalan Di Puskesmas Taram Kabupaten Lima Puluh Kota Tahun 2023. *Jurnal Rekam Medis Dan Informasi Kesehatan*, 4(1), 29–34. <https://ejurnal2.poltekkestasikmalaya.ac.id/index.php/jremikes/article/view/766/467>
- Permenkes RI No. 19. (2024). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 19 Tahun 2024 tentang Puskesmas*. <https://peraturan.bpk.go.id/Download/373654/permenkes-no-19-tahun-2024.pdf>
- Permenkes RI No. 24. (2022). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 24 Tahun 2022 tentang Rekam Medis*. <https://peraturan.bpk.go.id/Download/235921/Permenkes%20Nomor%2024%20Tahun%202022.pdf>
- Permenkes RI No. 31. (2019). *Peraturan Menteri Kesehatan Republik Indonesia tentang SIMPUS Nomor 31 Tahun 2019*. <https://peraturan.bpk.go.id/Download/129887/Permenkes%20Nomor%2031%20Tahun%202019.pdf>
- Susanah, S., Aryani, B., Haryanto, Y., & Dewi Rahmawati, F. (2025). Review of the Accuracy of Codification of Chronic Kidney Disease Cases Based on ICD-10 at Hospital X in 2023. *JURNAL KESEHATAN, SAINS, DAN TEKNOLOGI (JAKASAKTI)*, 4(2), 685–693. <https://doi.org/10.36002/js.v4i2.3799>
- Uppalapati, V. K., & Nag, D. S. (2024). A Comparative Analysis of AI Models in Complex Medical Decision-Making Scenarios: Evaluating ChatGPT, Claude AI, Bard, and Perplexity. *Cureus*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10874112/pdf/cureus-0016-00000052485.pdf>
- WHO. (2010). *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)*. World Health Organization.