

TATA KELOLA ETIS KECERDASAN BUATAN DALAM PELAYANAN KESEHATAN: TRANSPARANSI ALGORITMA, AKUNTABILITAS KLINIS, DAN KEPERCAYAAN PASIEN

Ariful Hakki^{1*}

¹Administrasi Kesehatan, Institut Teknologi dan Kesehatan Aspirasi

*Email korespondensi: hakki44@gmail.com

Article Info	Abstract
Article History: Received: January 7, 2026 Accepted: February 12, 2026 Published: March 29, 2026	Artificial intelligence can support clinical diagnosis, operational efficiency, and personalized care, yet its use raises concerns about explainability, responsibility, privacy, fairness, and patient trust. This integrative literature review develops an ethical governance framework linking algorithmic transparency, clinical accountability, and patient trust. Recent peer-reviewed research and international guidance were examined through thematic synthesis. Five interconnected components emerged: clear intended use and limitations, meaningful explanations for clinicians and patients, explicit decision ownership, bias and privacy audits, and patient participation. Transparency alone is insufficient to generate justified trust; explanations must be understandable, performance must be validated locally, and effective correction pathways must exist. Clinical accountability requires meaningful human oversight, traceable documentation, incident reporting, and continuous post-deployment monitoring. The proposed framework conceptualizes trust as an assessable governance outcome rather than unconditional confidence in technology. Healthcare organizations should establish multidisciplinary oversight, equity and safety indicators, and transparent risk communication. Because this study is conceptual rather than a field trial, its proposed mechanisms require empirical validation across healthcare settings.
Keyword : Keyword 1 ; Ethical AI Governance Keyword 2 ; Algorithm Transparency Keyword 3 ; Patient Trust Kata Kunci: Kata Kunci 1; Tata Kelola AI Etis Kata Kunci 2; Transparansi Algoritma Kata Kunci 3; Kepercayaan Pasien	Abstrak Penerapan kecerdasan buatan dalam pelayanan kesehatan membuka peluang peningkatan ketepatan diagnosis, efisiensi proses klinis, dan personalisasi perawatan, tetapi juga memunculkan persoalan transparansi, tanggung jawab profesional, privasi, keadilan, dan kepercayaan pasien. Penelitian ini bertujuan merumuskan kerangka tata kelola etis kecerdasan buatan yang mengintegrasikan transparansi algoritma, akuntabilitas klinis, dan kepercayaan pasien dalam pengambilan keputusan kesehatan. Penelitian menggunakan kajian literatur integratif terhadap publikasi ilmiah mutakhir dan dokumen kebijakan yang relevan, dengan sintesis tematik atas aspek desain, penerapan, pengawasan, dan evaluasi sistem. Hasil sintesis mengidentifikasi lima komponen utama, yaitu keterjelasan tujuan dan batas penggunaan algoritma, penjelasan yang bermakna bagi klinisi dan pasien, penetapan tanggung jawab keputusan, audit bias serta keamanan data, dan mekanisme partisipasi pasien. Hubungan antarkomponen menunjukkan bahwa transparansi hanya meningkatkan kepercayaan ketika penjelasan dapat dipahami, kinerja sistem divalidasi secara kontekstual, serta tersedia jalur koreksi dan pertanggungjawaban. Akuntabilitas klinis menuntut supervisi manusia yang nyata, dokumentasi keputusan, pengelolaan insiden, dan pemantauan setelah implementasi. Kerangka yang diusulkan menempatkan kepercayaan sebagai hasil tata kelola yang dapat diuji, bukan sekadar persepsi positif terhadap teknologi. Implikasi praktisnya adalah perlunya komite multidisiplin, indikator keselamatan dan keadilan, serta komunikasi risiko yang terbuka. Kajian ini menawarkan dasar konseptual untuk merancang evaluasi empiris dan

pedoman organisasi kesehatan yang bertanggung jawab, berorientasi pada keselamatan, hak pasien, dan mutu pelayanan berkelanjutan. Temuan bersifat sintesis konseptual sehingga efektivitas kerangka perlu diuji melalui penelitian lapangan, audit berkala, dan pelibatan pemangku kepentingan pada berbagai tingkat fasilitas kesehatan di Indonesia secara sistematis. Pelaksanaan kebijakan membutuhkan evaluasi transparan dan

Copyright © 2026, The Author(s)

How to cite: Hakki, A. (2026). TATA KELOLA ETIS KECERDASAN BUATAN DALAM PELAYANAN KESEHATAN: TRANSPARANSI ALGORITMA, AKUNTABILITAS KLINIS, DAN KEPERCAYAAN PASIEN. *Aspiration of Health Journal*, 4(1), 17–25. <https://doi.org/10.55681/aohj.v4i1.605>



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

PENDAHULUAN

Kecerdasan buatan semakin banyak digunakan untuk mendukung klasifikasi citra medis, penilaian risiko, pengelolaan alur kerja, dan komunikasi klinis. Namun, manfaat teknis tidak dengan sendirinya menjamin pelayanan yang aman atau adil. Boudierhem (2024) menekankan bahwa regulasi dan etika harus berkembang seiring penerapan teknologi. Dalam lingkungan klinis, keputusan algoritmik dapat memengaruhi keselamatan, otonomi, dan akses pasien terhadap layanan. Oleh karena itu, evaluasi kecerdasan buatan tidak cukup hanya berdasarkan akurasi prediksi, melainkan juga memerlukan kerangka pengawasan yang jelas.

Transparansi algoritma menjadi isu penting karena sistem pembelajaran mesin sering menghasilkan rekomendasi yang sulit ditafsirkan. Freyer et al. (2024) menguraikan alasan etis yang mendasari kebutuhan penjelasan dalam sistem pendukung keputusan klinis. Transparansi perlu dibedakan dari sekadar membuka kode sumber; penjelasan harus menjawab pertanyaan klinis tentang faktor relevan, keterbatasan prediksi, dan ketidakpastian. Tanpa penjelasan yang sesuai kebutuhan pengguna, informasi teknis justru dapat menambah beban kognitif tenaga kesehatan.

Kepercayaan pasien dan klinisi merupakan prasyarat penggunaan teknologi yang bertanggung jawab, tetapi kepercayaan berlebihan juga berisiko. Rosenbacke et al. (2024) menunjukkan bahwa penjelasan kecerdasan buatan dapat meningkatkan atau menurunkan kepercayaan klinisi bergantung pada mutu dan konsistensinya. Temuan tersebut mengindikasikan perlunya kepercayaan yang terkalibrasi, yakni kesediaan memanfaatkan sistem saat layak dan menolak rekomendasi ketika bukti klinis bertentangan. Kepercayaan pasien juga berkaitan dengan keterbukaan mengenai penggunaan data dan siapa yang bertanggung jawab atas keputusan.

Akuntabilitas klinis menuntut pembagian peran yang tegas antara pengembang, manajemen fasilitas kesehatan, dokter, perawat, dan penyedia teknologi. Solanki et al. (2023) menempatkan penerjemahan prinsip etika ke dalam proses pengembangan sebagai tugas operasional. Dalam praktik, perlu ditetapkan siapa yang memvalidasi model, menyetujui implementasi, memantau kejadian merugikan, dan menghentikan penggunaan bila risiko meningkat. Pengawasan manusia harus nyata dan dapat ditelusuri, bukan sekadar formalitas administratif.

Keadilan algoritmik menjadi dimensi tata kelola karena kinerja model dapat berbeda antarkelompok pasien. Chen et al. (2023) menjelaskan bahwa bias dapat berasal dari pengumpulan

data, pelabelan, dan variasi praktik klinis. Yang et al. (2024) menunjukkan tantangan generalisasi keadilan pada model pencitraan medis lintas lingkungan. Oleh sebab itu, audit yang hanya mengandalkan nilai akurasi agregat berpotensi menyembunyikan kerugian pada subpopulasi yang kurang terwakili.

Privasi dan keamanan informasi kesehatan merupakan syarat mendasar dalam pengembangan sistem yang dipercaya. Tayebi Arasteh et al. (2024) menelaah ketegangan antara perlindungan privasi, akurasi, dan keadilan dalam model pencitraan medis. Penggunaan data sensitif memerlukan pembatasan akses, pengelolaan persetujuan, pencatatan aktivitas, serta penilaian risiko pihak ketiga. Upaya perlindungan tidak boleh berhenti pada anonimisasi karena data dan keluaran model dapat tetap menimbulkan risiko pengungkapan.

Perkembangan model generatif dan multimodal memperluas ruang lingkup pengawasan. World Health Organization (2024) menyoroti perlunya tata kelola model multimodal besar yang dapat menghasilkan berbagai bentuk keluaran kesehatan. Pfohl et al. (2024) menunjukkan pentingnya pendekatan penilaian yang beragam untuk mengungkap potensi kerugian terhadap kesetaraan. Rumah sakit dan layanan primer perlu memastikan bahwa keluaran otomatis tidak diperlakukan sebagai diagnosis final tanpa pemeriksaan klinis.

Literatur telah membahas transparansi, akuntabilitas, dan kepercayaan, tetapi hubungan operasional ketiganya sering diperlakukan terpisah. Kajian ini bertujuan menyusun kerangka konseptual tata kelola etis kecerdasan buatan dalam pelayanan kesehatan melalui sintesis publikasi tahun 2022–2025 dan pedoman relevan. Kontribusi utama kajian ialah pemetaan proses dari desain dan validasi sampai pemantauan pascapenerapan, dengan indikator transparansi, tanggung jawab klinis, dan kepercayaan pasien. Kerangka ini dimaksudkan sebagai dasar evaluasi organisasi, bukan sebagai bukti efektivitas empiris yang sudah teruji.

METODE

Penelitian menggunakan pendekatan kajian literatur integratif dan pengembangan kerangka konseptual. Sumber dipilih secara purposif dari publikasi ilmiah tahun 2022–2025 yang membahas etika kecerdasan buatan kesehatan, transparansi, keadilan, akuntabilitas, dan kepercayaan, dilengkapi pedoman World Health Organization. Penelusuran konseptual berfokus pada artikel yang menyediakan argumen teoritis, temuan empiris, atau rekomendasi tata kelola. Kajian ini bukan tinjauan sistematis dengan perhitungan meta-analisis dan tidak mengklaim penelusuran menyeluruh seluruh basis data.

Analisis dilakukan melalui pembacaan kritis, pengodean tema, dan perbandingan antarartikel. Prosedur sintesis mengikuti prinsip analisis data kualitatif yang dijelaskan Sugiyono (2022) dalam buku *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*, terutama pengorganisasian informasi, kategorisasi, dan penarikan kesimpulan. Tema awal meliputi penjelasan algoritmik, validasi, bias, perlindungan data, tanggung jawab profesional, dan keterlibatan pasien. Perbedaan konteks serta keterbatasan bukti dicatat agar tidak disamakan dengan hubungan kausal.

Tahap perumusan model menghubungkan tema ke dalam tiga konstruk utama: transparansi algoritma, akuntabilitas klinis, dan kepercayaan pasien. Setiap konstruk dijabarkan menjadi tindakan organisasi dan indikator yang dapat dievaluasi. Konsistensi kerangka diperiksa dengan membandingkan temuan penelitian serta pedoman etika internasional. Tidak dilakukan pengumpulan data pasien, survei tenaga kesehatan, uji hipotesis, maupun penghitungan signifikansi statistik. Karena itu, hasil disajikan sebagai sintesis literatur dan usulan mekanisme tata kelola yang masih memerlukan validasi lapangan.

HASIL DAN PEMBAHASAN

1. Transparansi sebagai kewajiban klinis

Transparansi bukan sekadar membuka cara kerja perangkat lunak, tetapi memastikan bahwa alasan rekomendasi dapat dikomunikasikan dalam bahasa klinis yang berguna. Freyer et al. (2024) menjelaskan bahwa kebutuhan penjelasan berkaitan dengan pembenaran keputusan, penghormatan otonomi, dan pengawasan. Pada pelayanan kesehatan, penjelasan perlu memuat variabel penting, tingkat ketidakpastian, kondisi ketika model tidak berlaku, dan alternatif tindakan. Informasi tersebut harus tersedia sebelum rekomendasi dipakai dalam keputusan yang berisiko tinggi. Pengembang sebaiknya menyediakan dokumentasi model, sedangkan fasilitas kesehatan menerjemahkannya menjadi prosedur kerja. Pasien tidak harus memahami seluruh matematika algoritma, tetapi berhak mengetahui bahwa teknologi digunakan dan bagaimana hasilnya memengaruhi perawatan. Dengan demikian, transparansi dinilai dari pemahaman pengguna serta kemampuan mempertanyakan keluaran, bukan hanya jumlah dokumen teknis yang diterbitkan. Pemantauan rutin tetap diperlukan untuk menjamin

2. Mutu penjelasan dan kalibrasi kepercayaan

Kepercayaan yang tepat muncul ketika pengguna memahami manfaat dan batas kecerdasan buatan secara proporsional. Rosenbacke et al. (2024) menemukan bahwa penjelasan dapat meningkatkan, tidak mengubah, bahkan mengurangi kepercayaan klinisi sesuai mutu penyajiannya. Oleh karena itu, desain antarmuka tidak boleh memaksakan keyakinan melalui visualisasi yang tampak meyakinkan tetapi menutupi ketidakpastian. Rumah sakit dapat mengevaluasi apakah tenaga kesehatan mampu mengenali rekomendasi yang keliru melalui skenario simulasi. Ukuran keberhasilan bukan tingkat penerimaan algoritma tertinggi, melainkan ketepatan menerima atau menolak saran. Informasi tentang keterbatasan data pelatihan dan perubahan kondisi pasien harus tersedia. Pengawasan semacam ini menjaga agar kepercayaan tidak berubah menjadi ketergantungan otomatis yang mengurangi penalaran profesional dan keselamatan pasien dalam praktik. Keterbatasan bukti harus disampaikan terbuka agar rekomendasi tidak dipahami sebagai kepastian empiris.

3. Pembagian tanggung jawab klinis

Akuntabilitas klinis memerlukan rantai tanggung jawab yang dapat diperiksa ketika sistem menghasilkan rekomendasi yang keliru. Solanki et al. (2023) mengusulkan penerjemahan prinsip etika menjadi tindakan konkret dalam siklus pengembangan. Pada tingkat fasilitas, manajemen menetapkan kebijakan persetujuan teknologi, tim klinis menentukan kesesuaian penggunaan, dan unit mutu mengaudit dampaknya. Penyedia sistem wajib menyediakan informasi validasi serta pembaruan yang memengaruhi kinerja. Dokter atau tenaga profesional tetap menjalankan pertimbangan klinis sesuai kewenangannya, tanpa menganggap perangkat sebagai pengganti tanggung jawab. Dokumentasi perlu mencatat kapan keluaran digunakan, diabaikan, atau dikoreksi beserta alasannya. Pembagian peran ini penting agar investigasi insiden berfokus pada pembelajaran sistem, bukan sekadar menyalahkan individu yang bekerja dalam proses kompleks. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan

4. Audit bias pada kelompok pasien

Kinerja kecerdasan buatan yang baik secara keseluruhan belum tentu adil untuk setiap kelompok pasien. Chen et al. (2023) menjelaskan sumber bias dalam data dan praktik kesehatan, sedangkan Hu et al. (2024) menawarkan pendekatan teknis untuk mengurangi pengaruh atribut sensitif. Tata kelola etis mensyaratkan evaluasi terstratifikasi berdasarkan kelompok yang relevan

secara klinis dan sosial. Ukuran sensitivitas, spesifisitas, kalibrasi, serta dampak keputusan perlu dibandingkan tanpa membuka identitas pribadi. Jika ditemukan kesenjangan, organisasi harus menilai akar masalah pada representasi data, pelabelan, dan prosedur layanan. Perbaikan model tidak boleh hanya mengejar keseimbangan statistik tetapi mengabaikan keselamatan. Hasil audit sebaiknya dibahas bersama klinisi, ahli data, dan perwakilan pasien agar keputusan koreksi mempertimbangkan konsekuensi nyata. Pemantauan rutin tetap diperlukan untuk menjamin mutu layanan kesehatan secara berkelanjutan

5. Validasi eksternal dan perubahan konteks

Model yang akurat di lokasi pengembangan dapat menunjukkan performa berbeda ketika dipakai pada rumah sakit lain. Yang et al. (2024) menunjukkan bahwa perbaikan keadilan pada satu distribusi data tidak selalu bertahan pada pengujian eksternal. Karena itu, keputusan implementasi membutuhkan validasi populasi lokal, penilaian perubahan distribusi data, dan pemantauan berkala. Indikator operasional perlu mencakup kesalahan menurut subkelompok serta perubahan pola penggunaan klinis. Pengelola harus menetapkan ambang yang memicu pemeriksaan ulang atau penghentian sementara. Dokumentasi versi model menjadi penting agar perubahan perangkat lunak tidak menghilangkan jejak tanggung jawab. Proses validasi berkelanjutan membantu mencegah rasa aman palsu akibat hasil pengujian awal yang baik tetapi tidak lagi mencerminkan praktik pelayanan terkini. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan.

6. Perlindungan privasi sebagai dasar kepercayaan

Keamanan data kesehatan tidak dapat dipisahkan dari legitimasi penggunaan kecerdasan buatan. Tayebi Arasteh et al. (2024) menunjukkan bahwa teknik perlindungan privasi perlu dinilai bersama dampaknya terhadap akurasi dan keadilan model. Organisasi sebaiknya menerapkan minimisasi data, pembatasan hak akses, enkripsi, audit log, serta evaluasi risiko vendor. Penggunaan data sekunder untuk pelatihan memerlukan dasar hukum dan komunikasi yang dapat dipahami pasien. Keamanan juga mencakup kemungkinan kebocoran melalui keluaran sistem atau integrasi antaraplikasi. Komite tata kelola harus meninjau konsekuensi setiap perubahan arsitektur penyimpanan. Jika insiden terjadi, mekanisme pemberitahuan dan pemulihan perlu jelas. Kepercayaan pasien lebih layak dibangun melalui perlindungan yang dapat dibuktikan daripada pernyataan umum bahwa sistem telah aman. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan

7. Keadilan dalam model bahasa kesehatan

Model bahasa besar dapat memperluas akses informasi, tetapi juga berpotensi memperkuat stereotip dan kesenjangan pelayanan. Pfohl et al. (2024) mengembangkan pendekatan evaluasi untuk mengungkap potensi kerugian kesetaraan dalam jawaban medis. Temuan tersebut menunjukkan perlunya pengujian dengan skenario yang mencerminkan variasi latar pasien, kemampuan literasi, dan kebutuhan komunikasi. Penilaian tidak cukup menggunakan satu metrik otomatis karena dampak sosial sering muncul dalam cara informasi disajikan. Tim penilai sebaiknya mencakup klinisi, pakar kesetaraan, dan masyarakat pengguna. Keluaran model harus dapat diperiksa sumber serta ketidakpastiannya. Kebijakan organisasi perlu melarang penggunaan jawaban generatif sebagai keputusan final tanpa validasi profesional, terutama untuk keadaan gawat darurat atau terapi berisiko tinggi. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin tetap diperlukan

8. Pengawasan manusia yang bermakna

Pengawasan manusia sering disebut sebagai jaminan keselamatan, tetapi keberadaannya harus dibuktikan dalam alur kerja. Nagendran et al. (2024) meneliti perilaku dokter ketika menerima rekomendasi kecerdasan buatan yang aman dan tidak aman, termasuk perhatian terhadap penjelasan. Temuan semacam itu mengingatkan bahwa menyediakan fitur penjelasan tidak otomatis membuat pengguna memanfaatkannya secara kritis. Organisasi perlu merancang titik pemeriksaan sebelum tindakan penting, pelatihan mengenali kegagalan, dan prosedur eskalasi. Beban kerja serta tekanan waktu juga harus dipertimbangkan karena dapat mengurangi kualitas evaluasi. Indikator pengawasan dapat mencakup dokumentasi penolakan rekomendasi dan tindak lanjut insiden. Dengan demikian, supervisi menjadi kegiatan profesional yang terukur, bukan tanda persetujuan administratif pada layar komputer. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin tetap diperlukan

9. Penilaian etika sebelum implementasi

Pemeriksaan etika idealnya dimulai sebelum pembelian atau penerapan teknologi. Keane et al. (2024) membahas alat penilaian CARE-AI untuk membantu organisasi mengidentifikasi persoalan etis dalam implementasi model prediksi. Komite multidisiplin dapat menilai tujuan klinis, manfaat yang diharapkan, risiko kesalahan, ketimpangan akses, dan alternatif nonalgoritmik. Penilaian juga perlu mempertimbangkan siapa yang menanggung biaya kesalahan serta apakah pasien memiliki kesempatan menyampaikan keberatan. Setiap risiko harus memiliki pemilik, tindakan mitigasi, dan tenggat evaluasi. Dokumen persetujuan implementasi sebaiknya dapat ditinjau kembali ketika model diperbarui. Pendekatan ini mencegah keputusan teknologi didominasi klaim pemasaran dan membantu manajemen menempatkan keselamatan serta hak pasien di atas dorongan otomatisasi yang tidak teruji. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin tetap diperlukan untuk

10. Tata kelola model multimodal

Penggunaan sistem multimodal menghadirkan risiko tambahan karena satu model dapat memproses teks, gambar, dan informasi kesehatan lainnya. World Health Organization (2024) menguraikan pertimbangan tata kelola bagi model multimodal besar dalam bidang kesehatan. Rumah sakit perlu membatasi penggunaan berdasarkan tujuan yang telah divalidasi, termasuk siapa yang boleh memasukkan data dan jenis keluaran yang dapat digunakan. Kebijakan harus melarang memasukkan informasi identitas ke layanan yang tidak memenuhi standar perlindungan. Selain itu, keluaran sintetis dan ringkasan otomatis perlu diverifikasi terhadap rekam medis sumber. Pelatihan pengguna harus mencakup halusinasi, kesalahan konteks, dan risiko penggunaan ulang informasi. Tata kelola yang adaptif memungkinkan pembaruan aturan ketika kemampuan sistem berubah, tanpa mengorbankan prinsip kehati-hatian dan akuntabilitas klinis. Pemantauan rutin tetap diperlukan untuk menjamin mutu layanan kesehatan secara

11. Dokumentasi dan pelaporan model

Pelaporan yang jelas memudahkan evaluasi independen terhadap sistem prediksi klinis. Cohen dan Bossuyt (2024) menyoroti pentingnya pedoman TRIPOD+AI bagi keterbukaan laporan model. Informasi mengenai sumber data, definisi hasil, prosedur validasi, dan keterbatasan seharusnya tersedia bagi komite klinis sebelum penggunaan. Dokumentasi juga perlu menjelaskan perubahan model setelah implementasi dan alasan pembaruannya. Transparansi pelaporan berbeda dari penjelasan untuk pasien, tetapi keduanya saling melengkapi dalam rantai akuntabilitas. Penilai independen membutuhkan bukti teknis, sedangkan pasien memerlukan informasi yang

relevan bagi keputusan perawatan. Penyelarasan kedua lapisan komunikasi membantu organisasi menghindari klaim kinerja berlebihan. Dengan catatan yang terstruktur, kesalahan dapat ditelusuri dan rekomendasi perbaikan dapat diuji secara lebih objektif. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin tetap diperlukan

12. Etika penelitian dan keterlibatan komunitas

Tata kelola kecerdasan buatan memerlukan partisipasi pihak yang akan terdampak oleh keputusan algoritmik. Hunt et al. (2024) membahas tantangan etika penelitian kecerdasan buatan dalam kesehatan global, termasuk perhatian terhadap konteks dan kepentingan masyarakat. Partisipasi pasien dapat diwujudkan melalui konsultasi rancangan informasi, evaluasi keluhan, dan penilaian manfaat yang dirasakan. Keterlibatan tersebut sebaiknya mencakup kelompok yang sering kurang terwakili dalam data. Komunikasi risiko perlu menjelaskan bahwa pasien dapat mengajukan pertanyaan atau meminta peninjauan klinis. Dalam penelitian pengembangan, persetujuan partisipasi tidak boleh diperlakukan sebagai formalitas. Keterbukaan proses memperkuat legitimasi penggunaan teknologi dan menyediakan informasi praktis yang tidak selalu terdeteksi oleh metrik performa algoritma. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin tetap diperlukan untuk menjamin mutu

13. Mekanisme keamanan dan audit insiden

Sistem kecerdasan buatan membutuhkan pengawasan berkelanjutan setelah diluncurkan. Kajian dalam International Journal of Medical Informatics (2025) mengenai AI yang aman dan dipercaya menyoroti keamanan, bias, serta transparansi sebagai persoalan yang saling terkait. Organisasi dapat membentuk daftar risiko, kanal pelaporan kejadian, penilaian tingkat keparahan, dan prosedur penarikan sistem. Audit sebaiknya menilai dampak pada keputusan klinis, bukan hanya ketersediaan aplikasi. Hasil investigasi perlu diterjemahkan menjadi pembaruan prosedur, pelatihan, atau perubahan kontrak dengan vendor. Pelaporan yang tidak menghukum secara otomatis dapat membantu tenaga kesehatan mengungkapkan kegagalan tersembunyi. Mekanisme umpan balik juga perlu menjangkau pasien yang mengalami konsekuensi penggunaan sistem. Siklus ini menjadikan akuntabilitas sebagai praktik organisasi yang berulang. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan. Pemantauan rutin

14. Konteks kebijakan dan kesiapan organisasi

Pengaturan kecerdasan buatan harus disesuaikan dengan kapasitas fasilitas kesehatan dan lingkungan kebijakan. Boudierhem (2024) menunjukkan adanya tantangan regulasi, perlindungan data, dan kerja sama lintas institusi. Fasilitas dengan infrastruktur terbatas sebaiknya memulai dari penggunaan berisiko lebih rendah dan memastikan mekanisme rujukan manusia tetap tersedia. Kesiapan organisasi dapat dinilai melalui kualitas data, literasi digital tenaga kesehatan, tata kelola vendor, serta kemampuan audit. Kebijakan tertulis tidak akan efektif tanpa alokasi sumber daya dan kepemimpinan yang konsisten. Komite pengarah perlu meninjau manfaat, biaya, dan dampak distribusional secara berkala. Pendekatan bertahap memungkinkan pembelajaran sebelum sistem diterapkan secara luas. Dalam konteks Indonesia, harmonisasi prosedur dengan regulasi kesehatan dan perlindungan data menjadi kebutuhan utama. Indikator keberhasilan harus dikaitkan dengan keselamatan pasien dan keputusan klinis yang dapat dipertanggungjawabkan.

15. Kerangka integratif dan agenda evaluasi

Sintesis kajian ini mengusulkan hubungan bertahap antara transparansi, akuntabilitas, dan kepercayaan pasien. Eke dan Shuib (2025) menekankan peran penjelasan serta transparansi dalam pembentukan kepercayaan terhadap sistem kesehatan berbasis AI. Transparansi menyediakan dasar pemahaman, akuntabilitas menjamin adanya pihak yang dapat dimintai pertanggungjawaban, sedangkan pengalaman pasien menjadi ukuran legitimasi penerapan. Kerangka dapat diuji melalui survei pemahaman pasien, audit keputusan klinis, analisis kesenjangan kinerja, dan penilaian insiden. Penelitian berikutnya perlu menguji apakah perbaikan penjelasan benar-benar meningkatkan keputusan yang aman, bukan sekadar sikap positif. Perbandingan lintas rumah sakit dan layanan primer juga diperlukan untuk mengetahui pengaruh kapasitas organisasi. Usulan ini merupakan model konseptual, sehingga belum dapat dianggap sebagai bukti hubungan sebab akibat. Keterbatasan bukti harus disampaikan terbuka agar rekomendasi tidak dipahami sebagai kepastian empiris.

KESIMPULAN

Kajian ini menyimpulkan bahwa tata kelola etis kecerdasan buatan dalam pelayanan kesehatan memerlukan keterpaduan transparansi algoritma, akuntabilitas klinis, dan kepercayaan pasien. Transparansi yang bermakna mencakup penjelasan yang dapat dipahami, batas penggunaan, serta informasi tentang ketidakpastian dan risiko. Ketiganya perlu dinilai sebagai komponen yang saling memengaruhi, bukan sekadar kewajiban administratif yang berdiri sendiri.

Akuntabilitas klinis menuntut pembagian tanggung jawab, pengawasan manusia yang efektif, validasi lokal, audit bias, perlindungan data, dan mekanisme pelaporan insiden. Kepercayaan pasien seharusnya tumbuh dari bukti keamanan, keadilan, dan keterbukaan, bukan dari klaim teknologi yang tidak diuji. Kerangka konseptual ini dapat menjadi acuan awal bagi rumah sakit dan fasilitas kesehatan dalam menyusun kebijakan penerapan kecerdasan buatan.

Penelitian lanjutan perlu memvalidasi kerangka melalui wawancara pasien dan tenaga kesehatan, evaluasi alur keputusan klinis, serta audit performa lintas kelompok dan fasilitas. Karena artikel ini merupakan kajian literatur integratif, tidak ada klaim hubungan kausal atau peningkatan hasil klinis yang telah dibuktikan secara langsung. Pengembangan indikator operasional dan pengujian prospektif menjadi prioritas agar tata kelola etis dapat dievaluasi secara objektif.

DAFTAR PUSTAKA

- Bouderhem, R. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*, 11, 416. <https://doi.org/10.1057/s41599-024-02894-w>
- Chen, R. J., Wang, J. J., Williamson, D. F. K., et al. (2023). Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nature Biomedical Engineering*, 7, 719–742. <https://doi.org/10.1038/s41551-023-01056-8>
- Cohen, J. F., & Bossuyt, P. M. M. (2024). TRIPOD+AI: An updated reporting guideline for clinical prediction models. *BMJ*, 385, q824. <https://doi.org/10.1136/bmj.q824>
- Eke, C. I., & Shuib, L. (2025). The role of explainability and transparency in fostering trust in AI healthcare systems: A systematic literature review, open issues and potential solutions. *Neural Computing and Applications*, 37, 1999–2034. <https://doi.org/10.1007/s00521-024-10868-x>
- Freyer, N., Groß, D., & Lipprandt, M. (2024). The ethical requirement of explainability for AI-DSS in healthcare: A systematic review of reasons. *BMC Medical Ethics*, 25, 104. <https://doi.org/10.1186/s12910-024-01103-2>

- Hu, L., Li, D., Liu, H., et al. (2024). Enhancing fairness in AI-enabled medical systems with the attribute neutral framework. *Nature Communications*, 15, 8767. <https://doi.org/10.1038/s41467-024-52930-1>
- Keane, P. A., et al. (2024). An ethics assessment tool for artificial intelligence implementation in healthcare: CARE-AI. *Nature Medicine*, 30, 3038–3039. <https://doi.org/10.1038/s41591-024-03310-1>
- Nagendran, M., Festor, P., Komorowski, M., Gordon, A. C., & Faisal, A. A. (2024). Eye tracking insights into physician behaviour with safe and unsafe explainable AI recommendations. *npj Digital Medicine*, 7, 202. <https://doi.org/10.1038/s41746-024-01200-x>
- Pfohl, S. R., Cole-Lewis, H., Sayres, R., et al. (2024). A toolbox for surfacing health equity harms and biases in large language models. *Nature Medicine*, 30, 3590–3600. <https://doi.org/10.1038/s41591-024-03258-2>
- Rosenbacke, R., Melhus, Å., McKee, M., & Stuckler, D. (2024). How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: Systematic review. *JMIR AI*, 3, e53207. <https://doi.org/10.2196/53207>
- Solanki, P., Grundy, J., & Hussain, W. (2023). Operationalising ethics in artificial intelligence for healthcare: A framework for AI developers. *AI and Ethics*, 3, 223–240. <https://doi.org/10.1007/s43681-022-00195-z>
- Tan, W., Wei, Q., Xing, Z., et al. (2024). Fairer AI in ophthalmology via implicit fairness learning for mitigating sexism and ageism. *Nature Communications*, 15, 4750. <https://doi.org/10.1038/s41467-024-48972-0>
- Tayebi Arasteh, S., Ziller, A., Kuhl, C., et al. (2024). Preserving fairness and diagnostic accuracy in private large-scale AI models for medical imaging. *Communications Medicine*, 4, 46. <https://doi.org/10.1038/s43856-024-00462-6>
- Yang, Y., Zhang, H., Gichoya, J. W., Katabi, D., & Ghassemi, M. (2024). The limits of fair medical imaging AI in real-world generalization. *Nature Medicine*, 30, 2838–2848. <https://doi.org/10.1038/s41591-024-03113-4>
- Zhang, H., Gerych, W., & Ghassemi, M. (2024). A data-centric perspective to fair machine learning for healthcare. *Nature Reviews Methods Primers*, 4, 86. <https://doi.org/10.1038/s43586-024-00371-x>
- Pathni, R. K. (2024). Beyond algorithms: Ethical implications of AI in healthcare. *Medical Journal Armed Forces India*. <https://doi.org/10.1016/j.mjafi.2024.10.014>
- World Health Organization. (2024). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. World Health Organization. <https://www.who.int/publications/b/70584>
- World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. <https://www.who.int/publications/i/item/9789240029200>
- Sugiyono. (2022). Metode penelitian kuantitatif, kualitatif, dan R&D. Alfabeta.