Indonesian Health Justice Review
ISSN: XXXX-XXXXX (Online)
Vol. 3, No. 1, 2026, pp. 81-98
Published online since March 25, 2026

Informed Consent for AI-Assisted Diagnosis: Do Patients Have the Right to Refuse Artificial Intelligence in Clinical Decision-Making?

Satrio Agung Bagaskara*
Universitas Brawijaya, Malang, Indonesia
satriobagaskara@gmail.com

DOI: https://doi.org/10.65815/38p7x443

Submitted: 30-08-2025	Revised: 29-11-2025, 27-02-2026	Accepted: 10-03-2026

*Corresponding Author

Abstract

Kecerdasan buatan (AI) semakin banyak diintegrasikan ke dalam proses pengambilan keputusan diagnostik dan klinis, menciptakan tantangan baru bagi doktrin persetujuan berdasarkan informasi (informed consent) tradisional. Persetujuan berdasarkan informasi konvensional umumnya berfokus pada intervensi medis yang diusulkan, risiko, manfaat, dan alternatifnya. Namun, penggunaan AI memperkenalkan dimensi tambahan: apakah pasien harus diberitahu bahwa AI terlibat dalam diagnosis mereka dan apakah mereka berhak untuk menolak keterlibatan tersebut. Artikel ini mengkaji landasan hukum persetujuan berdasarkan informasi dalam perawatan kesehatan berbantuan AI di Indonesia dan mengevaluasi apakah hak pasien yang ada cukup memadai untuk menangani pengambilan keputusan algoritmik. Dengan menggunakan penelitian hukum normatif, studi ini menganalisis hukum kesehatan, etika medis, otonomi pasien, dan prinsip-prinsip tata kelola AI yang bertanggung jawab. Analisis menunjukkan bahwa persetujuan berdasarkan informasi yang bermakna membutuhkan lebih dari sekadar pengungkapan risiko medis; persetujuan tersebut juga harus mencakup informasi penting mengenai penggunaan dan keterbatasan AI di mana informasi tersebut dapat memengaruhi keputusan pasien untuk menjalani perawatan. Artikel ini mengusulkan standar pengungkapan kontekstual berdasarkan signifikansi AI dalam pengambilan keputusan klinis, potensi konsekuensi kesalahan algoritmik, dan ketersediaan alternatif yang wajar. Mengakui otonomi pasien dalam perawatan kesehatan yang dibantu AI akan memperkuat transparansi, akuntabilitas, dan kepercayaan sekaligus mempromosikan model tata kelola kesehatan digital yang lebih berpusat pada pasien.

[*Kecerdasan buatan (AI) semakin banyak diintegrasikan ke dalam proses pengambilan keputusan diagnostik dan klinis, menciptakan tantangan baru bagi doktrin persetujuan berdasarkan informasi (informed consent) tradisional. Persetujuan berdasarkan informasi konvensional umumnya*

berfokus pada intervensi medis yang diusulkan, risiko, manfaat, dan alternatifnya. Namun, penggunaan AI memperkenalkan dimensi tambahan: apakah pasien harus diberitahu bahwa AI terlibat dalam diagnosis mereka dan apakah mereka berhak untuk menolak keterlibatan tersebut. Artikel ini mengkaji landasan hukum persetujuan berdasarkan informasi dalam perawatan kesehatan berbantuan AI di Indonesia dan mengevaluasi apakah hak pasien yang ada cukup memadai untuk menangani pengambilan keputusan algoritmik. Dengan menggunakan penelitian hukum normatif, studi ini menganalisis hukum kesehatan, etika medis, otonomi pasien, dan prinsip-prinsip tata kelola AI yang bertanggung jawab. Analisis menunjukkan bahwa persetujuan berdasarkan informasi yang bermakna membutuhkan lebih dari sekadar pengungkapan risiko medis; persetujuan tersebut juga harus mencakup informasi penting mengenai penggunaan dan keterbatasan AI di mana informasi tersebut dapat memengaruhi keputusan pasien untuk menjalani perawatan. Artikel ini mengusulkan standar pengungkapan kontekstual berdasarkan signifikansi AI dalam pengambilan keputusan klinis, potensi konsekuensi kesalahan algoritmik, dan ketersediaan alternatif yang wajar. Mengakui otonomi pasien dalam perawatan kesehatan yang dibantu AI akan memperkuat transparansi, akuntabilitas, dan kepercayaan sekaligus mempromosikan model tata kelola kesehatan digital yang lebih berpusat pada pasien.]

Keywords: Informed consent, artificial intelligence, patient autonomy, clinical decision-making, digital health

Introduction

Artificial intelligence (AI) has moved from being primarily an experimental technology to becoming an increasingly relevant component of contemporary healthcare. Machine-learning systems can assist clinicians in interpreting medical images, identifying patterns in laboratory data, predicting clinical deterioration, supporting differential diagnosis, and recommending treatment options. The attraction of these technologies is understandable. AI can process large volumes of information, identify correlations that may be difficult for humans to detect, and potentially improve diagnostic consistency and efficiency (Topol 2019; Rajkomar, Dean, and Kohane 2019).

Yet the incorporation of AI into clinical practice changes more than the technological infrastructure of healthcare. It potentially changes the legal and ethical relationship between patient and physician. Traditional medical decision-making is generally conceptualized as a relationship between a patient and a human healthcare professional. AI introduces another actor—or, more precisely, another decision-support mechanism—into that relationship. The resulting question is not simply whether AI is accurate. It is whether patients have a legally protected

interest in knowing who or what participates in decisions concerning their bodies, and whether that interest includes the ability to reject algorithmic participation.

This question is becoming increasingly important because AI can occupy very different positions in clinical decision-making. In one case, AI may perform a purely administrative function, such as scheduling or transcription. In another, it may provide an optional second reading of a radiological image. In a more consequential situation, an algorithm may materially influence a diagnosis, risk classification, treatment recommendation, or decision concerning whether further intervention is required. Treating all these applications identically would ignore the very different implications they have for patient autonomy.

The World Health Organization (WHO) has recognized this problem. Its 2021 guidance on AI in health identifies protection of human autonomy as a foundational principle and emphasizes that humans should remain in control of healthcare systems and medical decisions. It also links patient autonomy with valid informed consent and calls for transparency, accountability, and mechanisms through which individuals can seek redress when affected by algorithmic decisions (WHO 2021). The World Medical Association has similarly emphasized that patients should, where possible, be informed about the role AI plays in their care and that patient autonomy should be preserved through meaningful consent processes (WMA 2025).

The legal problem is particularly significant in Indonesia because the country's health-law framework predates the widespread deployment of generative and clinical AI. Indonesian health law recognizes patient autonomy and the right to receive health information, obtain adequate explanations concerning healthcare services, and accept or refuse medical treatment. Article 276 of Law No. 17 of 2023 concerning Health expressly recognizes the patient's rights to obtain information about their health, receive adequate explanations concerning healthcare services, and approve or reject medical actions, subject to statutory exceptions. The implementing Government Regulation No. 28 of 2024 further develops the legal architecture implementing Law No. 17 of 2023.

At the same time, Indonesian law protects the informational dimension of AI-assisted healthcare. Law No. 27 of 2022 concerning Personal Data Protection classifies health and medical information as specific personal data and establishes rights concerning information, processing, accountability, and protection of personal data (Republic of Indonesia 2022). Regulation of the Minister of Health No. 24 of 2022 concerning Medical Records further establishes rules for electronic medical records and confirms that the contents of medical records belong to the patient, while the physical or electronic medical-record document is held by the healthcare facility (Republic of Indonesia 2022b).

These rules, however, do not expressly answer the AI-specific question. The law clearly protects a patient's right to information and the right to accept or reject medical action, but it does not explicitly establish a general statutory right to refuse an AI system used in diagnosis. This creates an interpretive problem: can the

existing right to refuse medical action be extended to a right to refuse a technological component of clinical decision-making?

The distinction is important. A patient may be willing to undergo a biopsy but unwilling to have an AI system interpret the pathology. A patient may agree to treatment but object to a machine-learning risk score influencing the treatment recommendation. Alternatively, a patient may have no objection to AI being used as an internal tool provided that the final decision remains with a physician. These scenarios demonstrate that "AI consent" cannot simply be treated as another checkbox on a conventional consent form.

Existing scholarship has begun to address this issue. Cohen and Slottje (2024) examine the relationship between AI and informed consent and argue that existing consent doctrine must be reconsidered in light of shared decision-making and the uncertain role of AI in clinical judgment. Iserson (2024), focusing on emergency medicine, argues that patient autonomy may require physicians to disclose AI involvement and, where feasible, allow patients to express preferences regarding its use. By contrast, Sauerbrei et al. (2024) argue that AI-enhanced healthcare does not necessarily create an entirely new paradigm of informed consent and caution against assuming that every use of AI automatically creates an independent consent requirement.

This disagreement reveals an important research gap. The debate has largely developed in US, UK, and international bioethics contexts, while the specific implications for Indonesia's statutory patient-rights framework remain underdeveloped. More importantly, the existing debate frequently treats the question as binary: either patients should have a general right to refuse AI or AI should remain entirely within the professional discretion of clinicians. Such a binary framework is insufficient because AI can range from low-risk administrative assistance to systems that materially affect diagnosis and treatment. This article therefore asks: whether, and under what circumstances, Indonesian patients should have a legally protected right to refuse AI participation in clinical decision-making.

The article uses normative legal research. Primary legal materials include Law No. 17 of 2023 concerning Health, Government Regulation No. 28 of 2024, Law No. 27 of 2022 concerning Personal Data Protection, Regulation of the Minister of Health No. 24 of 2022 concerning Medical Records, and relevant international governance instruments. Secondary materials consist of academic literature on informed consent, patient autonomy, medical AI, algorithmic transparency, and digital-health governance. Comparative materials from WHO, WMA, the European Union, the United States, and other international frameworks are used not as binding law in Indonesia but as interpretive and policy references.

The article argues that Indonesian law can support a contextual right to refuse AI involvement, even though it does not currently establish an explicit absolute right to reject every form of AI. The stronger interpretation is that patients should have a right to refuse AI where its use is material to the clinical decision, creates meaningful additional risks, involves significant processing of sensitive health data,

or where a reasonable human-only alternative exists. Conversely, requiring individualized consent for every low-risk computational tool would impose an excessive burden and could undermine beneficial healthcare innovation. The appropriate legal model is therefore not "AI always requires separate consent", but "material AI involvement triggers enhanced disclosure and, in appropriate circumstances, a right to object."

Reconstructing Informed Consent in the Age of Clinical AI

The emergence of artificial intelligence in clinical decision-making challenges an assumption that has long been embedded in the doctrine of informed consent: that the principal decision-making relationship exists between the patient and a human healthcare professional. Conventional informed consent doctrine is structured around the disclosure of the proposed medical intervention, its foreseeable risks and benefits, available alternatives, and the consequences of accepting or refusing treatment. The increasing use of AI complicates this structure because the clinical decision may now be produced through an interaction between professional judgment and an algorithmic system. The legal question is consequently no longer limited to whether a patient has consented to a particular medical intervention. It increasingly concerns whether the patient is entitled to know the technological conditions under which that intervention has been recommended and whether those conditions are sufficiently material to affect the patient's autonomous decision (Beauchamp and Childress 2019; Faden and Beauchamp 1986).

The significance of this problem depends on the role performed by AI. It would be analytically mistaken to treat every software application used within a healthcare institution as part of clinical decision-making. An algorithm that schedules appointments, converts speech into clinical notes, or assists with administrative management does not ordinarily alter the substantive medical judgment exercised by the physician. Requiring a separate informed-consent procedure for every such technological application would transform informed consent into an exhaustive inventory of hospital technology rather than a mechanism for protecting patient autonomy. The legal relevance of AI therefore depends not simply on its presence but on its material contribution to the clinical decision.

This distinction becomes considerably more important where an AI system interprets diagnostic information, generates a risk prediction, prioritizes patients, recommends a treatment, or influences the decision whether additional testing is necessary. In such circumstances, AI becomes part of the epistemic process through which the clinician understands the patient's condition. A patient who is told that a diagnosis has been reached through conventional clinical assessment is receiving materially different information from a patient whose diagnosis has been significantly influenced by an algorithm trained on large datasets. The difference is not necessarily that one method is better than the other. Rather, the difference lies

in the source, limitations, and structure of the information on which the medical decision is based.

This is particularly relevant because the risks associated with clinical AI are not identical to conventional clinical risks. Machine-learning systems can perform differently across patient populations, encounter difficulties when deployed outside their original validation environment, and reproduce biases embedded within training data (Gianfrancesco et al. 2018; Kelly et al. 2019). The problem of opacity further complicates clinical judgment. A physician may know that an algorithm has classified a patient as high risk without being able to provide an intuitive account of the precise factors that generated the classification (Burrell 2016; Rudin 2019). Consequently, the legal relevance of AI cannot be determined solely by asking whether the technology has been approved or whether its statistical performance is superior to that of human clinicians. The more fundamental question is whether AI changes the informational circumstances in which the patient is being asked to authorize medical care.

The concept of materiality provides a more appropriate doctrinal basis. Information should be regarded as material when a reasonable patient could consider it relevant to deciding whether to undergo the proposed healthcare intervention. Under this approach, disclosure of AI involvement is not automatically required merely because an algorithm has been used somewhere in the healthcare process. Disclosure becomes more important as AI moves closer to the core of diagnosis or treatment and as the consequences of algorithmic error become more serious. The patient need not understand the technical architecture of a neural network or the statistical parameters of a machine-learning model. Meaningful autonomy requires something different: the patient should understand whether AI was used, what role it performed, whether its output materially influenced the clinician, whether the physician independently evaluated that output, and whether the system presents limitations that could reasonably affect the decision.

This approach is consistent with emerging international principles of responsible medical AI. The World Health Organization places protection of human autonomy at the center of AI governance in healthcare and emphasizes that individuals should retain control over healthcare decisions involving AI (World Health Organization 2021). The United States Food and Drug Administration, together with Health Canada and the United Kingdom's Medicines and Healthcare products Regulatory Agency, similarly emphasizes transparency concerning the intended use, performance, limitations, and human-AI interaction of machine-learning-enabled medical devices (FDA, Health Canada, and MHRA 2024). These principles suggest that transparency should be calibrated to clinical significance rather than imposed indiscriminately.

The resulting conception of informed consent is therefore broader than conventional risk disclosure but narrower than technological disclosure in its most expansive form. AI-sensitive informed consent should not require physicians to disclose every software component used in a hospital. It should require disclosure

when the AI system becomes a material part of the clinical reasoning that supports a decision affecting the patient's body, health, or future treatment. This distinction allows informed consent to retain its central function as an instrument of autonomy while avoiding the impracticality of converting every technological interaction into a separate consent event.

Patient Autonomy and the Interpretive Capacity of Indonesian Health Law

Indonesian health law does not expressly establish a statutory "right to refuse artificial intelligence." This absence, however, should not be interpreted to mean that patients have no legal interest in the use of AI in their healthcare. The stronger interpretation is that existing patient-rights provisions provide a normative foundation from which AI-related autonomy can be developed through systematic interpretation. Law No. 17 of 2023 concerning Health recognizes the patient's right to receive information concerning health, to receive adequate explanations concerning healthcare services, and to accept or refuse medical action subject to statutory limitations. The statute therefore establishes the essential elements from which an AI-sensitive doctrine of informed consent can be constructed, even though it does not identify AI as a separate object of consent.

The most important element is the patient's right to adequate information and explanation. These provisions should not be understood as limited to information about the biological condition of the patient or the physical procedure proposed by a physician. Healthcare is a process, and the circumstances through which a medical judgment is reached can become legally relevant when they materially affect the patient's decision. If an AI system substantially contributes to the diagnosis of a potentially serious disease, information about that contribution may reasonably fall within the scope of information necessary for the patient to understand the basis of the proposed intervention.

The argument is stronger when the AI system does more than merely provide supplementary information. Suppose, for example, that an algorithm identifies a patient as having a high probability of malignancy and the physician recommends an invasive procedure principally because of that classification. The patient's decision to undergo the procedure is then connected to the algorithmic assessment. Treating AI involvement as legally irrelevant merely because the physician formally signs the final recommendation would reduce the right to explanation to a procedural fiction. The patient may have consented to the intervention, but the quality of that consent depends upon whether the patient was provided with information sufficiently relevant to understand the basis of the recommendation.

At the same time, extending the statutory right to refuse medical action into an unrestricted right to refuse any technological component of healthcare would be difficult to justify. AI is not necessarily a medical action in itself. It may function as a diagnostic instrument in much the same way that physicians rely upon laboratory

equipment, statistical models, imaging technologies, or clinical guidelines. If every patient could veto the use of any technological tool used in healthcare, the autonomy principle could become incompatible with the institutional realities of modern medicine.

The better approach is therefore to distinguish between refusing treatment because AI is involved and objecting to material AI participation in the clinical decision-making process. The former would imply a general technological veto that Indonesian law does not clearly recognize. The latter can be derived more plausibly from the interaction between the patient's rights to information, adequate explanation, and informed acceptance or refusal of medical treatment.

This distinction also responds to an important counterargument. One might contend that because the physician remains legally responsible for the final clinical decision, disclosure of AI involvement is unnecessary. That argument assumes, however, that the physician's involvement automatically neutralizes the significance of the algorithm. It does not. Human oversight is legally meaningful only when the physician exercises genuine independent judgment. If the physician simply accepts an algorithmic recommendation without meaningful scrutiny, describing the process as a human decision may obscure rather than resolve the autonomy problem.

Accordingly, Indonesian law should distinguish between formal human oversight and meaningful human oversight. A physician should be able to evaluate whether the AI output is clinically appropriate, recognize relevant limitations, depart from the algorithm when professional judgment requires it, and explain the material basis of the final recommendation. The physician's legal responsibility should remain intact regardless of whether the patient has been informed of AI involvement. Consent cannot convert an unsafe or negligent deployment of AI into lawful medical practice.

The patient's autonomy therefore operates at two interconnected levels. The first concerns the right to decide whether to undergo medical treatment. The second concerns the informational circumstances in which that decision is made. AI becomes legally relevant when it materially changes those circumstances. The absence of explicit AI terminology in Law No. 17 of 2023 should not prevent the law from responding to technological change because rights expressed in general language must remain capable of application to new forms of healthcare delivery. Otherwise, technological innovation would paradoxically become a mechanism for reducing the practical scope of rights that were intended to protect patients regardless of the technology used.

Informational Autonomy, Health Data, and Artificial Intelligence

The autonomy problem becomes more complex when AI-assisted diagnosis is examined through the lens of personal data protection. Clinical AI systems generally depend upon extensive health information, including medical histories, diagnostic

images, laboratory results, medication information, demographic characteristics, and other forms of sensitive information. Indonesian Law No. 27 of 2022 concerning Personal Data Protection expressly classifies health data as specific personal data and establishes enhanced legal protection for such information. This framework is particularly important because AI may process health information not only for the immediate treatment of the patient but also for model development, validation, monitoring, and improvement.

The legal distinction between these activities should not be overlooked. A patient who authorizes the use of medical information for diagnosis does not necessarily understand that the same information may subsequently be used for a secondary technological purpose. Conversely, data protection law does not necessarily require a separate consent for every individual computational operation undertaken in the course of legitimate healthcare. The existence of AI therefore does not automatically resolve the question of lawful processing. Rather, it requires healthcare institutions to identify the specific purpose, legal basis, scope, and safeguards associated with the processing.

This distinction reveals why informed consent and data protection should be treated as complementary rather than interchangeable. Informed consent protects clinical autonomy by enabling the patient to decide whether to undergo healthcare. Data protection protects informational autonomy by regulating the collection, use, disclosure, retention, and other processing of personal information. AI-assisted healthcare can implicate both forms of autonomy simultaneously.

The distinction has practical significance. A patient might be entirely comfortable with an AI system analyzing an MRI to assist the physician's diagnosis but object to the MRI subsequently being incorporated into a dataset used to develop a commercial diagnostic model. Another patient might accept secondary data processing but object to an algorithm being permitted to determine the clinical recommendation. These are not contradictory preferences. They represent different dimensions of autonomy that should not be collapsed into a single generalized consent.

Law No. 27 of 2022 provides an important normative foundation for this distinction through its principles of protection, legal certainty, benefit, public interest, prudence, balance, accountability, and confidentiality. These principles are particularly relevant to clinical AI because the risks of data processing can extend beyond the immediate clinical encounter. Once health information enters an AI system, questions arise concerning access, security, retention, reuse, model development, and potential inference of additional information from the original dataset.

The importance of these concerns is reinforced by the broader international approach to trustworthy AI. The WHO emphasizes privacy, confidentiality, transparency, and accountability as central principles for AI in healthcare, while the National Institute of Standards and Technology identifies privacy, fairness, transparency, explainability, safety, and accountability as interconnected

dimensions of trustworthy AI governance (World Health Organization 2021; Tabassi 2023). These principles support an interpretation of Indonesian data-protection law in which AI systems handling health information should be subject to governance proportionate to the sensitivity and consequences of their processing.

The regulatory challenge is therefore not simply whether a patient "consented to AI." The more precise legal questions concern what the AI was permitted to do, what data it processed, for what purpose, with what safeguards, and with what consequences for the patient. A consent form that merely states that "artificial intelligence may be used in healthcare" would provide little meaningful protection because it does not communicate the particular implications of the technology.

An effective legal framework should consequently require purpose-specific transparency. Where AI is used directly to support diagnosis, the patient should be able to understand its clinical role. Where health information is processed for secondary AI development, the relevant data-protection requirements should apply independently. Where the AI system produces an output that materially affects treatment, the patient should have access to meaningful human review. These requirements do not create a technologically hostile healthcare environment. They instead recognize that the more powerful the technology becomes, the more important it is to identify the precise legal relationship between data processing, clinical judgment, and individual autonomy.

The Materiality Threshold and the Emerging Right to Object

The central legal question is not whether patients should possess an unrestricted veto over AI, but when AI involvement becomes sufficiently material to trigger a legally protected opportunity to object. A materiality-based standard provides a more defensible solution because it recognizes both the value of autonomy and the operational realities of healthcare.

The first consideration should be the extent to which AI influences the clinical outcome. An algorithm that merely organizes information is fundamentally different from an algorithm that determines a patient's diagnostic category or substantially influences the selection of treatment. As the algorithmic output moves closer to the final clinical decision, the justification for enhanced disclosure becomes stronger because the technology becomes more consequential to the patient's medical interests.

The second consideration is the severity of potential harm. The legal significance of AI should increase where an error could result in serious or irreversible consequences. A mistaken administrative classification is unlikely to raise the same autonomy concerns as an erroneous diagnosis of cancer, an incorrect assessment of stroke risk, or a treatment recommendation that could expose a patient to significant harm. A risk-sensitive approach is therefore preferable to a technology-sensitive approach. What matters is not whether AI was used but what could happen if the AI-assisted decision is wrong.

The third consideration is whether the AI system has limitations that are particularly relevant to the patient. Machine-learning systems can exhibit different levels of performance across populations, clinical settings, and data environments. Where such limitations are material to the patient's circumstances, disclosure becomes especially important. This does not require the physician to provide an exhaustive technical explanation. It requires the physician to communicate clinically relevant limitations in language the patient can reasonably understand.

The fourth consideration is whether a reasonable human-led alternative exists. This factor is critical because the practical value of a refusal right depends upon whether refusal produces an actual alternative or merely forces the patient to abandon healthcare. If a patient can receive substantially equivalent assessment from a physician without AI, refusing AI represents a meaningful exercise of autonomy. If no reasonable alternative exists, the legal analysis must balance autonomy against the patient's interest in receiving timely and effective care.

The fifth consideration is the degree of human oversight. AI should not be treated as an autonomous legal subject. Responsibility for the clinical decision must remain with identifiable healthcare professionals and institutions. Nevertheless, the existence of a physician somewhere in the decision-making chain should not automatically defeat a patient's objection. What matters is whether the physician is genuinely capable of evaluating and overriding the AI output.

These considerations support recognition of a contextual right to object, rather than an absolute right to refuse. Such a right would be strongest where AI materially influences diagnosis or treatment, the potential consequences of algorithmic error are significant, and a reasonable human-led alternative is available. It would be weaker where AI performs a low-risk supporting function or where refusal would create disproportionate clinical consequences.

This approach is consistent with the emerging scholarly debate. Iserson (2024) emphasizes the relevance of patient autonomy where AI is integrated into emergency clinical decision-making, while Sauerbrei et al. (2024) caution against treating AI as automatically requiring a new form of consent in every clinical interaction. These positions need not be viewed as mutually exclusive. The disagreement largely disappears when the legal analysis focuses on materiality. AI does not necessarily create a new consent category; rather, certain forms of AI involvement make existing principles of informed consent more demanding.

This is also why the phrase “right to refuse AI” can be misleading. It suggests that AI itself is the object of a general patient veto. A more accurate formulation is a right to object to material AI-mediated clinical decision-making. The distinction is doctrinally significant. It recognizes the patient's autonomy without assuming that patients possess legal control over every instrument used by healthcare professionals.

Such a right should also include the possibility of human reassessment. If an AI system materially contributes to a consequential decision and the patient objects, the healthcare institution should, where reasonably practicable, provide a human-

led review. This does not mean that the physician must reach the conclusion preferred by the patient. It means that the patient should not be trapped within an algorithmic decision pathway without access to accountable professional judgment. The principle becomes especially important as healthcare institutions increasingly adopt AI systems for triage, diagnosis, and treatment recommendations. Without a contestability mechanism, informed consent risks becoming purely informational. A patient may be told that AI was used but have no meaningful opportunity to challenge the consequences. Autonomy requires more than disclosure; it requires a reasonable capacity to exercise choice.

Emergency Care, Clinical Necessity, and the Limits of Patient Refusal

A contextual right to object must nevertheless recognize circumstances in which immediate healthcare takes priority over contemporaneous technological choice. Emergency medicine provides the clearest example. A patient may arrive unconscious, unable to communicate preferences, while an AI system is capable of rapidly identifying a potentially fatal condition. Requiring prior authorization before the system can be used could create a conflict between autonomy and the preservation of life.

The existence of an emergency exception does not mean that patient autonomy disappears. Rather, the legal justification for immediate AI use must arise from necessity and proportionality. The healthcare provider should be able to demonstrate that the situation was genuinely urgent, that AI use was reasonably connected to the clinical objective, and that the system remained subject to professional oversight. The exception should not become a general mechanism through which healthcare institutions avoid ordinary transparency obligations.

Post-treatment transparency is therefore important. When a patient subsequently becomes capable of participating in decision-making, the healthcare provider should disclose the material use of AI and explain its clinical role where reasonably practicable. Such deferred disclosure preserves the patient's informational autonomy even where contemporaneous consent was impossible.

The emergency context also demonstrates why the right to refuse AI cannot be understood as an absolute property right. Healthcare law balances individual autonomy with other legally protected interests, including preservation of life, prevention of serious harm, and protection of public health. The appropriate objective is not to maximize refusal rights regardless of circumstances but to ensure that technological necessity does not become an automatic justification for disregarding patient autonomy.

The same principle should apply to situations in which a healthcare provider argues that AI is necessary because of resource constraints. Technological efficiency may constitute a legitimate consideration, but institutional convenience should not automatically override individual autonomy. If a hospital chooses to make AI a routine component of its diagnostic pathway, it should also consider whether

patients who reasonably object can access an alternative form of clinical review. Otherwise, the right to object would exist formally but disappear practically.

This concern is particularly important in developing digital-health systems because unequal access to human-led alternatives may create a new form of technological dependency. Wealthier patients may be able to seek independent human assessment elsewhere, while economically vulnerable patients may be required to accept AI-mediated healthcare as the only available option. A patient-centered regulatory model must therefore consider not only whether refusal is legally permitted but also whether the healthcare system makes meaningful choice practically possible.

Accountability, Liability, and the Prohibition of Responsibility Transfer

The recognition of AI-related patient autonomy must not inadvertently create a mechanism for shifting responsibility from healthcare providers to patients. There is a significant legal difference between informing a patient about AI and asking the patient to assume the risks of AI. The first is an essential component of transparency. The second could undermine professional accountability.

A consent form that states that the patient "accepts all risks associated with artificial intelligence" should therefore be treated with considerable skepticism. The patient is not normally in a position to assess whether the AI system was appropriately validated, whether its training data were representative, whether it was deployed within its intended context, or whether the physician adequately interpreted its output. Those responsibilities belong primarily to the actors who design, procure, deploy, supervise, and use the system.

The allocation of responsibility should consequently remain functional. Developers should be responsible for the design and documented limitations of their systems. Healthcare institutions should be responsible for procurement, validation, deployment, monitoring, staff training, and governance. Physicians should remain responsible for exercising professional judgment and appropriately interpreting AI outputs. Regulators should establish standards and enforcement mechanisms. Patients should exercise autonomous choices concerning their healthcare without being transformed into guarantors of technological safety.

This allocation is especially important for medical liability. Consider a case in which an AI system fails to identify a serious condition, the physician accepts the output without adequate independent evaluation, and the patient subsequently suffers harm. The fact that the patient consented to AI use cannot by itself eliminate the physician's potential responsibility. The relevant legal questions remain whether the AI system was appropriate for the clinical context, whether the physician exercised reasonable professional judgment, whether the healthcare institution implemented adequate safeguards, and whether material information was disclosed.

This distinction also prevents the emergence of what may be described as algorithmic responsibility displacement. AI creates a risk that each actor will characterize the system as belonging to someone else: the physician may blame the software, the hospital may blame the physician, and the developer may argue that the system was merely an assistive tool. Effective legal governance must reject this circular allocation of responsibility.

Human oversight therefore has to be substantive rather than symbolic. A physician cannot satisfy the requirement of human control simply by clicking "approve" after an algorithm generates a recommendation. Meaningful oversight requires the ability and willingness to critically evaluate the output. If a healthcare institution deploys a system that physicians cannot adequately understand, monitor, or override, the institution should bear heightened responsibility for that governance failure.

The same principle strengthens rather than weakens the case for AI-related disclosure. When patients are told that AI materially influenced their care, they can better understand the structure of the decision and exercise their rights to ask questions, request human review, or seek another medical opinion. Transparency therefore serves not merely as an ethical courtesy but as an accountability mechanism.

Explainability, Contestability, and Procedural Autonomy

A final dimension of AI-assisted informed consent concerns explainability. It would be unrealistic to require every healthcare provider to provide patients with a complete technical explanation of an AI model. Many contemporary systems are highly complex, and even developers may not be able to provide a simple causal explanation for every individual output. The law should therefore distinguish between technical explainability and clinically meaningful transparency.

Technical explainability concerns the internal operation of the algorithm. Clinically meaningful transparency concerns whether the patient can understand how the system participated in the decision and what significance should be attributed to its output. For purposes of informed consent, the latter is generally more important. A patient deciding whether to undergo a biopsy does not need to understand the mathematical parameters of the model. The patient does need to know that AI identified a pattern associated with elevated risk, that the physician reviewed the result, and that further testing is being recommended because of the combined clinical assessment.

This approach also aligns with the distinction between explainability and interpretability in responsible AI governance. NIST's AI Risk Management Framework recognizes explainability and interpretability as important elements of trustworthy AI while acknowledging that these values operate alongside fairness, privacy, safety, and accountability (Tabassi 2023). The same integrated approach should inform Indonesian healthcare regulation.

Procedural autonomy consequently becomes as important as informational disclosure. Patients should have a meaningful opportunity to ask whether AI was used, understand its clinical role, request clarification concerning material limitations, and seek human reassessment where the AI output has significant consequences. This creates a procedural safeguard against the possibility that AI systems become effectively unchallengeable within healthcare institutions.

The right to contest should not, however, be confused with a right to demand that clinicians disregard scientifically supported AI outputs. The purpose of contestability is to preserve human deliberation, not to guarantee a particular medical conclusion. A patient who objects to AI use should receive an explanation of whether an alternative assessment is clinically reasonable. The physician may ultimately reach the same diagnosis, but the decision should emerge from accountable professional judgment rather than unreviewable algorithmic authority.

For Indonesia, this suggests that future regulation should move beyond a simple disclosure requirement. Healthcare institutions should establish procedures through which patients can identify material AI involvement and request human review when appropriate. Such procedures would operationalize the patient's existing rights under health law while integrating them with the accountability and transparency principles of data protection and responsible AI governance.

The resulting legal architecture would be substantially more sophisticated than a generic "AI consent" form. It would recognize that the central legal issue is not whether technology is present but whether technology has altered the conditions under which an individual makes a consequential healthcare decision. That is the point at which AI becomes not merely a technological instrument but a matter of legal autonomy.

Conclusion

Artificial intelligence does not justify replacing the conventional doctrine of informed consent with an entirely separate regime of "AI consent." Nor does the absence of an explicit reference to artificial intelligence in Indonesia's health legislation mean that patients have no legally relevant interest in how AI participates in their care. Indonesian law already establishes a substantial normative foundation through the patient's rights to health information, adequate explanation, and acceptance or refusal of medical action under Law No. 17 of 2023, complemented by the heightened protection of health data under Law No. 27 of 2022 and the governance of medical records under Permenkes No. 24 of 2022. These provisions should be interpreted dynamically in light of the changing conditions of healthcare delivery. Where AI materially influences diagnosis or treatment, information concerning its use, clinical role, significant limitations, and human oversight can become material to autonomous decision-making. The appropriate legal response is therefore not to recognize an unlimited technological veto but to establish a contextual right to object to material AI-mediated clinical decision-making.

Such a framework would permit Indonesia to reconcile technological innovation with the foundational principle that medical care remains oriented toward human autonomy. The strength of a patient's objection should depend upon the degree of AI influence, the severity of potential harm, the sensitivity of the data involved, the availability of a reasonable human-led alternative, and the quality of meaningful human oversight. Emergency circumstances may justify temporary departure from contemporaneous consent where necessary to protect life or prevent serious harm, but such exceptions should remain proportionate and followed by transparency where practicable. Most importantly, patient consent must never operate as a transfer of responsibility for algorithmic error from healthcare professionals or institutions to the patient. The future of AI governance in Indonesian healthcare should therefore be built around materiality, transparency, contestability, and human accountability. Under this model, patients do not possess an absolute right to prohibit technology from entering the clinical environment; they possess the more defensible and practically meaningful right to know when AI materially participates in decisions concerning them, to understand the significance of that participation, and, where a reasonable alternative exists, to refuse or seek human reassessment of an AI-mediated clinical decision.

References

Amann, Julia, Alessandro Blasimme, Effy Vayena, and others. 2020. "Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective." *BMC Medical Informatics and Decision Making* 20: 310. https://doi.org/10.1186/s12911-020-01332-6.

Beauchamp, Tom L., and James F. Childress. 2019. *Principles of Biomedical Ethics*. 8th ed. New York: Oxford University Press.

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3 (1): 1–12. https://doi.org/10.1177/2053951715622512.

Char, Danton S., Nigam H. Shah, and David Magnus. 2018. "Implementing Machine Learning in Health Care—Addressing Ethical Challenges." *New England Journal of Medicine* 378 (11): 981–983. https://doi.org/10.1056/NEJMp1714229.

Cohen, I. Glenn, and Andrew Slottje. 2024. "Artificial Intelligence and the Law of Informed Consent." In *Research Handbook on Health, AI and the Law*, edited by I. Glenn Cohen and others. Cheltenham: Edward Elgar Publishing.

European Union. 2024. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).* Official Journal of the European Union.

Faden, Ruth R., and Tom L. Beauchamp. 1986. *A History and Theory of Informed Consent*. New York: Oxford University Press.

Food and Drug Administration, Health Canada, and Medicines and Healthcare products Regulatory Agency. 2024. *Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles*. Silver Spring, MD: U.S. Food and Drug Administration.

Gianfrancesco, Milena A., Suzanne Tamang, Jino Yazdany, and Gabriela Schmajuk. 2018. "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data." *JAMA Internal Medicine* 178 (11): 1544–1547. https://doi.org/10.1001/jamainternmed.2018.3763.

Iserson, Kenneth V. 2024. "Informed Consent for Artificial Intelligence in Emergency Medicine: A Practical Guide." *American Journal of Emergency Medicine* 76: 225–230. https://doi.org/10.1016/j.ajem.2023.11.022.

Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1: 389–399. https://doi.org/10.1038/s42256-019-0088-2.

Kelly, Christopher J., Alan Karthikesalingam, Mustafa Suleyman, Greg Corrado, and Dominic King. 2019. "Key Challenges for Delivering Clinical Impact with Artificial Intelligence." *BMC Medicine* 17: 195. https://doi.org/10.1186/s12916-019-1426-2.

Kiseleva, Anastasia, and Paul De Hert. 2020. "If You Don't Know What You're Doing, Don't Do It! Artificial Intelligence and the Right to Explanation." *International Data Privacy Law* 10 (2): 128–141. https://doi.org/10.1093/idpl/ipaa007.

Morley, Jessica, Luciano Floridi, Libby Kinsey, and Anat Elhalal. 2020. "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices." *Science and Engineering Ethics* 26: 2141–2168. https://doi.org/10.1007/s11948-019-00165-5.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366 (6464): 447–453. https://doi.org/10.1126/science.aax2342.

Rajkomar, Alvin, Jeffrey Dean, and Isaac Kohane. 2019. "Machine Learning in Medicine." *New England Journal of Medicine* 380: 1347–1358. https://doi.org/10.1056/NEJMra1814259.

Republic of Indonesia. 2022a. *Law Number 27 of 2022 concerning Personal Data Protection*. Jakarta: Government of the Republic of Indonesia.

Republic of Indonesia. 2022b. *Regulation of the Minister of Health Number 24 of 2022 concerning Medical Records*. Jakarta: Ministry of Health of the Republic of Indonesia.

Republic of Indonesia. 2023. *Law Number 17 of 2023 concerning Health*. Jakarta: Government of the Republic of Indonesia.

Republic of Indonesia. 2024. *Government Regulation Number 28 of 2024 concerning the Implementing Regulation of Law Number 17 of 2023 concerning Health*. Jakarta: Government of the Republic of Indonesia.

Rudin, Cynthia. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1: 206–215. https://doi.org/10.1038/s42256-019-0048-x.

Sauerbrei, A., and others. 2024. "AI-Enhanced Healthcare: Not a New Paradigm for Informed Consent." *Journal of Bioethical Inquiry* 21: 475–489.

Sendak, Mark P., Michael C. D'Arcy, Michael J. Kashyap, David Gao, Suresh K. C. K. K. ... 2020. "A Path for Translation of Machine Learning Products into Healthcare Delivery." *EMJ Innovations* 4 (1): 43–49.

Tabassi, Elham. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1.

Topol, Eric. 2019. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books.

Vayena, Effy, Alessandro Blasimme, and I. Glenn Cohen. 2018. "Machine Learning in Medicine: Addressing Ethical Challenges." *PLoS Medicine* 15 (11): e1002689. https://doi.org/10.1371/journal.pmed.1002689.

World Health Organization. 2021. *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*. Geneva: World Health Organization.

World Medical Association. 2025. *WMA Statement on Artificial and Augmented Intelligence in Medical Care*. Ferney-Voltaire: World Medical Association.

***

Acknowledgment
None

Funding Information
None

Conflicting Interest Statement
The authors state that there is no conflict of interest in the publication of this article.

Publishing Ethical and Originality Statement
All authors declared that this work is original and has never been published in any form and in any media, nor is it under consideration for publication in any journal, and all sources cited in this work refer to the basic standards of scientific citation.

Generative AI Statement
N/A