

MEDICAL ETHICS LECTURE SERIES

Lecture 8

Special Ethical Cases, Including Artificial Intelligence

Department of Psychiatry • Jordan University Hospital
Psychiatry Residency Training Program

Lecture 8 of 8 • Duration: 40 minutes

Learning Objectives



Describe the WMA's ethical framework for AI in medical care



Explain why the physician remains accountable even when using AI tools



Identify the risks of bias, opacity, and data misuse in health care AI



Discuss the specific promise and risk of AI applications in psychiatry



Apply a general ethical framework to novel situations not covered by existing rules

Case Study



CASE VIGNETTE

A resident's patient with recurrent depression has been relying on a general-purpose AI chatbot for support between sessions. Separately, the hospital's newly deployed AI suicide-risk-prediction tool flags this same patient as “low risk.” The resident's own clinical assessment — weighing a recent bereavement — raises more concern than the algorithm does.

Discussion: Whose assessment should govern — the algorithm's or the resident's? And does the chatbot use need to be addressed at all? Hold this question; we will return to it.

Two AI tools, one patient, and a physician caught between them — exactly the terrain of this lecture.

Why “Special Cases” Still Need Ethics

Novel technology does not get a moral exemption. Every tool in this lecture is governed by the same four principles and declarations covered across this series — applied to unfamiliar terrain.



The principles still apply

Autonomy, beneficence, non-maleficence, and justice (Lecture 1) frame every question in this lecture, however novel the technology.



The declarations still apply

Confidentiality, consent, and non-discrimination obligations (Lectures 2–3) do not pause for new tools.



The physician's duty is unchanged

Whatever tool is in use, the physician's duties to the patient (Lecture 4) remain the anchor.

Artificial Intelligence in Medicine



Diagnostic support

Image analysis, pattern recognition, and decision-support tools that flag findings for clinician review.



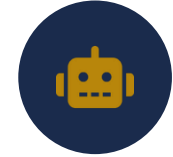
Risk prediction

Algorithms estimating risk of relapse, readmission, or self-harm from clinical and historical data.



Ambient documentation

AI tools that draft clinical notes from a recorded encounter, reducing documentation burden.



Conversational agents

Chatbots offering psychoeducation, symptom tracking, or supportive conversation to patients.

The WMA Framework

“AI is supporting rather than replacing human judgment, empathy, and accountability... the licensed physician retains ultimate authority over all AI-generated outputs influencing diagnosis, treatment, or care decisions.”¹

Why the WMA prefers “augmented” over “artificial” intelligence:

- The term signals a human-centred approach — the physician remains the final decision-maker
- AI tools are framed as extensions of clinical expertise, not replacements for it
- This applies even where a well-defined task is performed independently by AI — human accountability still governs

¹ WMA Statement on Artificial and Augmented Intelligence in Medical Care, 2025.

Duty 1: Physician Accountability



“Physician-in-the-loop” (PITL): a human physician remains accountable for decisions influenced by AI, even when the AI's role is substantial.¹

In practice, this means:

- An AI recommendation is an input to clinical judgement, not a substitute for it
- You cannot defer responsibility to “what the algorithm said” — the decision, and its consequences, remain yours
- Where multiple professionals are involved, the treating physician retains ultimate clinical responsibility
- Institutions must give physicians real, practical ability to question or override an AI recommendation

¹ WMA Statement on Artificial and Augmented Intelligence in Medical Care, 2025.

Duty 2: Transparency & Explainability



Patients are entitled to know when AI materially shapes their care, and physicians should be able to explain, in plain terms, why a tool reached its conclusion.

In practice, this means:

- Disclose to patients when a diagnostic, risk, or treatment recommendation involved an AI tool
- “Black box” tools that cannot be meaningfully explained deserve extra caution before relying on their output
- Ask vendors and institutions what validation and explainability standard a tool actually meets
- Documentation should reflect the physician's own reasoning, not simply restate the AI's output

Consistent with the informed-consent duty established in Lecture 1 and Lecture 4.

Duty 3: Bias, Equity & Validation



AI systems learn from historical data — and historical health data reflects historical inequities. An unvalidated tool can quietly encode and scale those inequities.

In practice, this means:

- Ask whether a tool was validated on a population resembling your own patients — performance does not automatically transfer
- Watch for systematically different outputs across demographic groups; report suspected bias through institutional channels
- Continuous-learning systems can drift over time — monitoring does not end at deployment
- Equitable access to beneficial AI tools is itself a justice concern, not only accuracy

See Lecture 1 (justice) and Lecture 5 (resource allocation) for the underlying principles.

Duty 4: Data Privacy & Governance



“The rapidly developing use of big data has implications for confidentiality and privacy... it is essential to preserve high ethical standards.”²

In practice, this means:

- Patient data used to train or run an AI tool is still subject to the same confidentiality duties as any other record
- Understand, in general terms, where patient data goes when an AI tool is used — including any third-party vendor
- Consent for data use in AI systems should be as clear as consent for any other use of medical information
- Institutional data-governance policy, not individual improvisation, should set the rules here

² WMA Declaration of Taipei on Ethical Considerations regarding Health Databases and Biobanks.

AI in Psychiatry: Promise and Risk



Genuine promise

Risk-prediction tools can flag patients for closer follow-up; chatbots can extend psychoeducation and symptom tracking between visits; ambient scribes can free up time for the therapeutic relationship itself.



Specific risks to weigh

A chatbot cannot hold clinical responsibility or recognise a true emergency the way a clinician can. Risk-prediction scores can create false reassurance, or dilute a clinician's own judgement, particularly for low-frequency, high-stakes events like suicide.

When AI and Clinical Judgement Disagree

Per the WMA framework: AI augments judgement, and clinical judgement is entitled to prevail — that is what accountability means.



Other Special Ethical Cases



Genetics & genomic testing

Incidental findings, family implications of a single patient's results, and biobank data governance.³



Telemedicine

Verifying identity and consent remotely, and knowing the limits of what can be safely assessed by video.



Social media & online conduct

Maintaining patient confidentiality and professional boundaries in physicians' public digital presence.



Organ transplantation

Allocation justice, and separating the transplant team from the team determining death or withdrawal of care.

A General Approach to Any Novel Situation



Start with the four principles

Autonomy, beneficence, non-maleficence, justice — they apply before any specific rule exists (Lecture 1).



Check the relevant declarations

WMA and specialty declarations often already address the underlying value, even for a new technology (Lecture 3).



Return to your core duties

Duties to the patient, to society, and to colleagues (Lectures 4–6) don't disappear because the tool is new.



Use the conflict-resolution toolkit

Discussion, consultation, and escalation (Lecture 7) work as well for a novel dilemma as for a familiar one.

Back to the Case Study

The resident's clinical judgement is entitled to prevail over the “low risk” score — the algorithm is an input, not the decision-maker. The chatbot use also deserves a direct, non-judgemental conversation, not silent concern.

Applying today's framework:

- Document the specific clinical factors (recent bereavement) the algorithm may be underweighting
- Escalate the discrepancy to a supervisor rather than deferring silently to the “low risk” label
- Ask the patient directly, with curiosity rather than alarm, about how the chatbot is being used and what it offers her
- Treat the chatbot as informative context, not as a substitute for the therapeutic relationship or a crisis plan

KEY TAKEAWAYS

Special Ethical Cases, Including Artificial Intelligence

- New technology does not get a moral exemption — the same principles and declarations still govern
- “Augmented,” not “artificial,” intelligence: physicians remain accountable for AI-influenced decisions
- Transparency, bias awareness, and data governance are active duties, not passive assumptions
- In psychiatry specifically, AI can extend care but cannot hold clinical or therapeutic responsibility
- When facing any unfamiliar situation, the same four-step approach from this series still applies

COMPLETING THE SERIES

Eight Lectures in Medical Ethics

- 1. Definition & Fundamentals of Medical Ethics
- 2. Human Rights and the Law
- 3. Declarations: The Instruments of Medical Ethics
- 4. Physician's Duties to Patients
- 5. Physician's Duties to Society
- 6. Physician's Duties to Other Physicians
- 7. Conflicts, Reporting, Cooperation & Resolution
- 8. Special Ethical Cases, Including Artificial Intelligence

Every lecture used the same foundation: the four principles, the WMA declarations, and a structured way to reason through disagreement — tools built to outlast any single technology or case.

References

1. World Medical Association. WMA Statement on Artificial and Augmented Intelligence in Medical Care. Adopted 2025.
2. World Medical Association. WMA Declaration of Taipei on Ethical Considerations regarding Health Databases and Biobanks.
3. World Medical Association. WMA Declaration of Kigali on the Ethical Use of Medical Technology.
4. World Medical Association. WMA Declaration of Seoul on Professional Autonomy and Clinical Independence.
5. American Medical Association. Advancing Health Care AI through Ethics, Evidence and Equity.
6. World Medical Association. Medical Ethics Manual. 3rd ed. Ferney-Voltaire: WMA; 2015.
7. World Medical Association. WMA Declaration of Geneva. Adopted 1948; revised 2017.

This concludes the eight-lecture Medical Ethics series for the JUH Psychiatry Residency Program.