Ethical challenges of artificial intelligence-driven healthcare

Dr. Eva Van Steijvoort

KU Leuven

Ethical challenges

How to balance the risks and benefits of AI technology?

Improve efficiency of health care delivery and quality of patient care

vs.

Avoid/limit threats to privacy and confidentiality, informed consent, patient autonomy, patient safety

Key ethical principles for the use of AI for health

1. Protect autonomy
2. Promote human well-being, human safety & the public interest
3. Ensure transparency, explainability and intelligibility
4. Foster responsibility & accountability
5. Ensure inclusiveness and equity
6. Promote AI that is responsive & sustainable

Source: Ethics and governance of artificial intelligence for health: WHO guidance. Geneva 2021.

Key ethical principles for the use of AI for health

Grounded in four basis ethical principles

1. Non-maleficence (‘do no harm’)
2. Beneficence
3. Justice
4. Autonomy

Protect autonomy

Human autonomy should not be undermined

- Human should remain in full control of health-care systems and medical decisions
- AI systems should be designed to assist in making informed decisions
- Health care professionals should be able to override decisions made by AI systems

- Protect privacy and confidentiality & ensure informed, valid consent

Promote human well-being, human safety and the public interest

AI technologies should not harm people

- AI technologies should satisfy regulatory requirements for safety, accuracy & efficacy
- Specific measures should be in place to ensure quality control & quality improvement
  - *Are AI technologies working as designed?*
  - *Is there any detrimental effect on individual patients or groups?*

Ensure transparency and explainability

AI should be understandable to developers, users and regulators

Transparency

➔ sufficient information published or documented before the design and deployment

- *To facilitate meaning public consultation and debate on how is should be designed and how it should be used.*

Explainable

➔ information should be tailored, according to the capacity of those to whom the explanation is directed

➔ possible trade-off between full explainability of an AI algorithm (at the cost of accuracy) and improved accuracy (at the costs of explainability)

Foster responsibility and accountability

‘Human warranty’ – points of human supervision

- Responsibility of human stakeholders to ensure that AI technologies can perform specific tasks for which they are used under appropriate conditions
- Ensure that the algorithm remains on a machine-learning development path that is medically effective, can be interrogated and is ethically responsible
- Appropriate mechanism should be in place when something does go wrong (questioning and redress)
- Diffusion of responsibility: ‘everybody’s problem becomes nobody’s responsibility’ → collective responsibility

Ensure inclusiveness and equity

Widest possible appropriate, equitable used and access irrespective of age, gender, income, ability or other characteristics

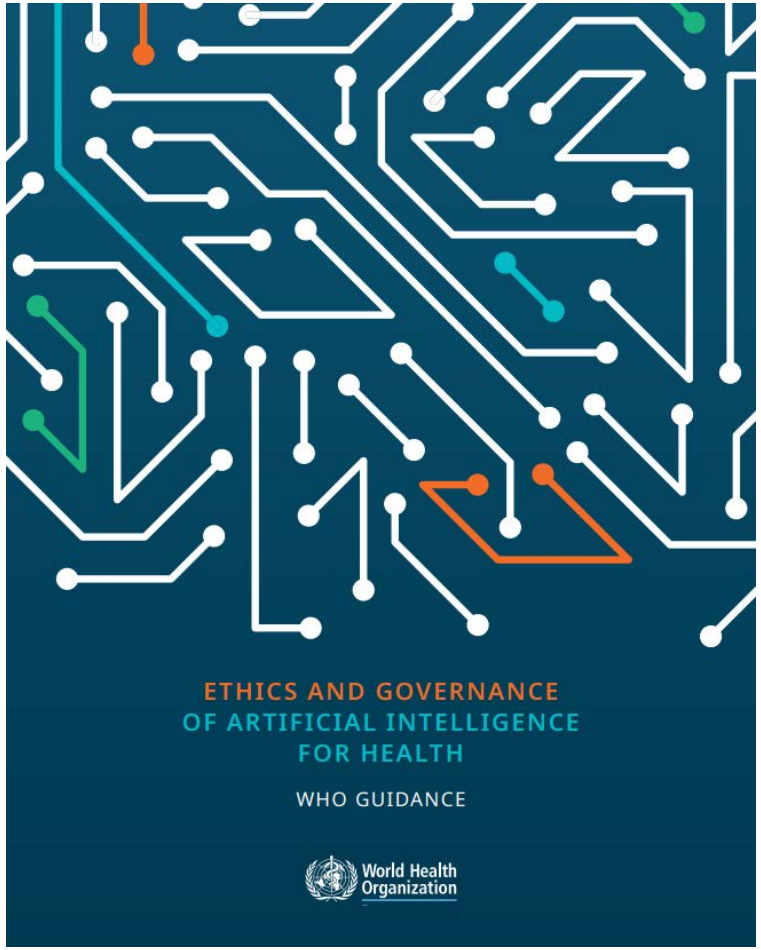
- Design and evaluation through stakeholders with diverse backgrounds
- AI technologies should be adaptable to context and needs of different settings (HIC vs LMIC)
- 'Digital divide' within and between countries should not be widened
- (Unintended) biases should be avoided or identified and mitigated
- Minimize inevitable power disparities (e.g. between providers & patients)
- Information technology literacy
- Open-source software – source codes publicly available

Promote artificial intelligence that is responsive and sustainable

Continuously, systematically and transparently monitoring

- Only introduce AI technologies that can be fully integrated and sustained in the health-care system
- Terminate use of technology if necessary
- Ecological footprint & energy efficiency
- Address anticipated disruptions to the workplace (e.g. training for health-care professionals, potential job losses, etc.)

WHO GUIDANCE



AI-POD

Trustworthy AI Tools for the Prediction of Obesity Related Vascular Diseases

Visit our website: www.ai-pod.eu

Follow us on twitter: @AIPOD_eu

Contact us:

Peter Gordebeke (EIBIR) – AI-POD Project Manager

pgordebeke@eibir.org

+43-1-533-4064-323

Funded by
the European Union

Thank you!

eva.vansteijvoort@kuleuven.be