Ethical and legal challenges of artificial intelligence-driven healthcare

Introduction

Economic forecasters have predicted **explosive growth in the AI health market** in the coming years.

This growth comes **many challenges**, and it is crucial that AI is implemented in the healthcare system **ethically and legally**.

Introduction

We will begin by briefly clarifying what AI is and giving an overview of the trends and strategies concerning ethics and law of AI in healthcare in the United States (US) and Europe.

This will be followed by an analysis of the ethical challenges of AI in healthcare. We will discuss 4 primary challenges:
(1) informed consent to use,
(2) safety and transparency,
(3) algorithmic fairness and biases, and
(4) data privacy

We then shift to 5 legal challenges in the US and Europe, namely,
(1) safety and effectiveness,
(2) liability,
(3) data protection and privacy,
(4) cybersecurity, and
(5) intellectual property law

Introduction

To realize the tremendous potential of AI to transform healthcare for the better,
stakeholders in the AI field, including:

- ***AI makers,***
- ***clinicians,***
- ***patients,***
- ***ethicists, and***
- ***legislators,***

must be engaged in ***the ethical and legal debate*** *on how AI is successfully implemented in practice.*

Machine learning (ML)*, is a subset of AI, and*
allows computational systems to learn from data and improve their performance.

Deep learning*, is a subset of ML, and*
employs artificial neural networks with multiple layers to identify patterns in very large datasets.

Most notably, as we will see below, ***there are additional ethical and legal challenges*** *in cases where ML algorithms are closer to* ***"black boxes"*** *(i.e., the results are very difficult for clinicians to interpret fully).*

Trends and strategies

We will discuss

- the **US and Europe's strategies** for AI,
- how they strive to compete against their biggest competitor **China**,
- **AI trends**, and
- some **examples of AI products that are already in clinical use** in the US and Europe

United States

During Barack Obama's presidency,
the US Government's reports on AI emphasized:

- *the applications of AI for the public good,*
- *the aspects of fairness, safety, and governance*
- *the need to improve fairness and transparency,*
- *building ethical AI*

Since Donald Trump's presidency,
the US AI strategy has shifted to a more free market-oriented approach.

In May 2018,
the White House,
for instance,
hosted the AI for American Industry Summit.

One of the key takeaways from the summit breakout discussions was that the Trump Administration aims to remove regulatory barriers to AI innovations.

United States

In July 2018,
the Executive Office of the President,
announced that ***American leadership in AI*** *is one of the* ***top*** *Administration R&D budget* ***priority*** *areas.*

In February 2019,
Trump signed the ***"Executive Order on Maintaining American Leadership in Artificial Intelligence"*** *to address the criticism that the US has taken a hands-off approach to AI in contrast to other countries such as China.*

With this executive order, Trump launched a coordinated Federal Government strategy, namely, ***the American AI Initiative****, guided by 5 key areas of emphasis:*
(1) investing in AI R&D,
(2) unleashing AI resources,
(3) setting AI governance standards,
(4) building the AI workforce, and
(5) international engagement and protecting the advantage of the US in AI

United States

In January 2020,
the White House
published ***draft guidance for the regulation of AI applications****.*

It contains 10 principles that agencies should consider when formulating approaches to AI applications:
(1) public trust in AI,
(2) public participation,
(3) scientific integrity and information quality,
(4) risk assessment and management,
(5) benefits and costs,
(6) flexibility,
(7) fairness and non-discrimination,
(8) disclosure and transparency,
(9) safety and security, and
(10) interagency coordination

United States

In February 2020,
the White House
also published
an annual report on the American AI Initiative,
summarizing the progress made since Trump signed the executive order.

This report,
for example,
highlights that
the US led historic efforts on the development of the Organization for Economic Co-operation and Development (OECD) Principles of AI
that were signed by over 40 countries in May 2019 to promote innovative and trustworthy AI and respect democratic values and human rights.

In June 2019,
the G-20
also released
AI Principles drawn from the OECD Principles of AI.

United States

The White House
has also launched a **website ("AI.gov")** that
focuses on
AI for the American people and
aims
to provide a platform for those who wish to learn more about AI and its opportunities.

There are also numerous **AI-related bills** that have been introduced in the US Congress since Trump's inauguration, such as

- the SELF DRIVE Act (H.R.3388),
- the FUTURE of Artificial Intelligence Act of 2017 (H.R.4625 and S.2217), and
- the AI JOBS Act of 2019 (H.R.827)

The SELF DRIVE Act,

- is the only bill that has passed **one chamber** (i.e., the US House of Representatives), and
- none of these bills are directly related to the ethical and legal aspects of AI in healthcare

United States

However,
the FUTURE of Artificial Intelligence Act of 2017,
for example,
stipulate the Secretary of Commerce to set up a ***Federal advisory committee*** *that shall provide advice to the Secretary.*

This committee shall also study and assess,

- *how to incorporate ethical standards in the development and implementation of AI, or*
- *how the development of AI can affect cost savings in healthcare*

There are also legal developments related to AI
at state and local levels*,*
for instance,
the State of California *adopted legislation in August 2018 (ACR-215) endorsing the 23 Asilomar AI principles.*

United States

AI is already in clinical use in the US. In particular, AI shows great promise in the areas of diagnostics and imaging.

The Food and Drug Administration (FDA) has already cleared or approved around 40 AI-based medical devices.

For example,
in January 2017,
Arterys received clearance from the US FDA for its medical imaging platform as the first ML application to be used in clinical practice.

It was initially cleared for **cardiac magnetic resonance image analysis**, but Arterys has meanwhile also received clearance from the FDA for other equivalent devices.

IDx-DR is the first FDA-authorized AI diagnostic system that provides an autonomous screening decision **without the need for a human** being to interpret the image or results additionally.

Estados Unidos

In April 2018,
the FDA permitted marketing of this AI-based device to detect ***more than a mild level of the eye condition diabetic retinopathy*** *in adult patients (ages 22 and older) diagnosed with diabetes.*

The physician uploads the images of the patient's retinas to a cloud server, and the IDx-DR software then provides the physician with the recommendation either to rescreen in 12 months or to refer the patient to an eye specialist ***when more than mild diabetic retinopathy is detected.***

In May 2018,
the FDA also granted marketing authorization for ***Imagen's software OsteoDetect*** *for helping clinicians in detecting a common type of wrist fracture, called* ***distal radius fracture****, in adult patients.*

OsteoDetect *uses ML techniques to analyze two-dimensional X-ray images to identify and highlight this type of fracture.*

Europe

The European Commission adopted its AI strategy for Europe in April 2018.

In this Communication, the Commission launched an European initiative on AI that ***aims*** *to ensure* ***an appropriate ethical and legal framework****, for example, by creating an European AI Alliance and developing AI ethics guidelines.*

The Commission also stresses in this Communication that the entire European Union (EU) should strive to increase the (public and private) investment in AI.

The European Commission's High-Level Expert Group on AI (AI HLEG)*,*
which
was appointed by the European Commission in June 2018, and is also the steering group for the European AI Alliance,
published the Ethics Guidelines in April 2019.

Europe

The Guidelines
promote the slogan "Trustworthy AI" and
contain 7 key requirements that AI systems need to fulfill in order to be trustworthy:
(1) human agency and oversight,
(2) technical robustness and safety,
(3) privacy and data governance,
(4) transparency,
(5) diversity, nondiscrimination and fairness,
(6) environmental and societal well-being, and
(7) accountability

The AI HLEG,
also published

- a document on the definition of AI, and
- another document that provides "Policy and Investment Recommendations for Trustworthy AI"

The European Commission encourages all EU Member States to develop a national AI strategy, and several states have already released one such as the United Kingdom (UK) and Germany.

Europe

In December 2018,
the European Commission,
also agreed upon
a coordinated plan on AI
with
EU Member States, Norway, and Switzerland
to promote the development and use of AI in Europe.

The ***overall goal*** *of working together is*
to ensure that Europe becomes the world-leading region for the development and application of ***"cutting-edge, ethical and secure AI"****.*

In February 2020,
the European Commission,
released

- *a* ***White Paper on AI*** *that contains a European approach to excellence and trust,*
- *a* ***Communication*** *on a European strategy for data, and*
- *a* ***Report*** *on the liability implications and safety of AI, the Internet of Things (IoT), and robotics*

Europe

The Commission's White Paper,
emphasizes that
*"Europe can combine **its technological and industrial strengths** with **a high-quality digital infrastructure and a regulatory framework** based on its fundamental values **to become a global leader** in innovation in the data economy and its applications"*

There are already AI health applications in Europe, and more are in the pipeline.

For example,
***Ada** is an AI health app that assesses an individual's symptoms and gives guidance (e.g., suggest to the user a visit to a doctor or to seek emergency care).*

***Ada** has been CE-marked **(class I)** in Europe—a basic requirement to putting a medical device on the market within Europe—and **complies with the EU General Data Protection Regulation 2016/679 (GDPR)**.*

Europe

In August 2018,
researchers at ***DeepMind and Moorfields Eye Hospital****,*
in London,
published in ***Nature Medicine*** *the study results of an AI system that*
can read ***eye scans*** *and make referral recommendation,*
comprising more than 50 common diagnoses.

The system was trained on 14884 scans and showed a success rate of 94%.

DeepMind's health team has meanwhile transitioned to Google Health.

Europe

Other examples:

- ***Ultromics***

The team at the University of Oxford,
is dedicated to reducing misdiagnosis and enabling earlier prevention of cardiovascular disease.

***Ultromics's EchoGo Pro**,*
*is an outcome-based AI system with CE marking in Europe that **predicts coronary artery disease at an early stage**.*

- ***Corti***

*Is a software developed by a **Danish company**,*
that leverages ML to help emergency dispatchers make decisions.

***Corti** can detect **outof-hospital cardiac arrests** (i.e., those that occur in the public or home) during emergency calls,*
faster and more accurately than humans,
***by listening in to calls and analyzing symptoms, the tone of voice, breathing patterns, and other metadata in real time**.*

Ethical challenges

1. Informed consent to use

How will the use of AI to assist with the care of patients interface with the principles of **informed consent**?

This is a pressing question that has not received enough attention in the ethical debate.

There is a need to examine…
Under what circumstances (if at all), the principles of informed consent should be deployed in the clinical AI space?

To what extent do clinicians have a responsibility to educate the patient around the complexities of AI, including the form(s) of ML used by the system, the kind of data inputs, and the possibility of biases or other shortcomings in the data that is being used?

Under what circumstances must a clinician notify the patient that AI is being used at all?

These questions are especially challenging to answer in cases where the AI operates using **"black-box" algorithms**, which may result from noninterpretable machine-learning techniques that are very difficult for clinicians to understand fully.

Ethical challenges

For instance,
***Corti's algorithms are "black box"**,*
*because even Corti's inventor **does not know** how the software reaches its decisions to alert emergency dispatchers that someone has a cardiac arrest.*

This lack of knowledge might be worrisome for medical professionals!!!

To what extent, for example, does a clinician need to disclose that they cannot fully interpret the diagnosis/treatment recommendations by the AI?

How much transparency is needed?

How does this interface with the so-called "right to explanation" under the EU's GDPR (discussed further)?

What about cases where the patient may be reluctant to allow the use of certain categories of data (e.g., genetic data and family history)?

How can we properly balance the privacy of patients with the safety and effectiveness of AI?

Ethical challenges

AI health apps and chatbots,
are also increasingly being used, ranging from

- *diet guidance*

to

- *health assessments,*
- *help to improve medication adherence, and*
- *analysis of data collected by wearable sensors*

Such apps raise questions for bioethicists about user agreements and their relationship to informed consent.

In contrast to the traditional informed consent process,
a user agreement is a contract that an individual agrees to without a face-to-face dialog.

Most people,

- *do not take the time to understand user agreements, routinely ignoring them*
- *frequent updates of the software make it even more difficult for individuals to follow what terms of service they have agreed to*

Ethical challenges

What information should be given to individuals using such apps and chatbots?

Do consumers sufficiently understand that the future use of the AI health app or chatbot may be conditional on accepting changes to the terms of use?

How closely should user agreements resemble informed consent documents?

What would an ethically responsible user agreement look like in this context?

Tackling these questions is tricky, and they become even more difficult to answer *when information from patient-facing AI health apps or chatbots is fed back into clinical decision-making.*

Ethical challenges

2. Safety and transparency

Safety is one of the biggest challenges for AI in healthcare.

For example,
IBM Watson for Oncology,
uses AI algorithms to

- assess information from patients' medical records, and
- help physicians explore cancer treatment options for their patients

It has recently come under criticism by giving "unsafe and incorrect" recommendations for cancer treatments.

The problem seems to be in the training of Watson for Oncology:
instead of using real patient data,
the software was only trained with a few "synthetic" cancer cases,
meaning they were devised by doctors at the Memorial Sloan Kettering (MSK) Cancer Center.

MSK has stated that errors only occurred as part of the system testing and thus no incorrect treatment recommendation has been given to a real patient.

Ethical challenges

This real-life example has put the field in a negative light.

It also shows that it is of uttermost importance that AIs are safe and effective.

But… how do we ensure that AIs keep their promises?

To realize the potential of AI,
AI developers, need to make sure two key things:
(1) the reliability and validity of the datasets, and
(2) transparency

The used datasets need to be reliable and valid.

The slogan "garbage in, garbage out" applies to AI in this area. The better the training data (labeled data) is, the better the AI will perform.

In addition, the algorithms often need further refinement to generate accurate results.

Another big issue is data sharing!!!

In cases where the AI needs to be extremely confident (e.g., self-driving cars), vast amounts of data and thus more data sharing will be necessary. However, there are also cases (e.g., a narrow sentiment AI-based off text) where less data will be required.

Ethical challenges

AI developers should be sufficiently transparent, for example,

- *about the kind of data used, and*
- *any shortcomings of the software (e.g., data bias)*

We should learn our lessons from examples such as Watson for Oncology, where IBM kept Watson's unsafe and incorrect treatment recommendations secret for over a year.

Transparency creates trust among stakeholders, particularly clinicians and patients, which is the key to a successful implementation of AI in clinical practice.

The recommendations of more "black-box" systems raise particular concerns.

It will be a challenge to determine how transparency can be achieved in this context.

Ethical challenges

3. Algorithmic fairness and biases

AI has the capability to
- "globalize" healthcare, and
- bring it to even remote áreas

AI bears a risk for biases and thus discrimination.

It is therefore vital that **AI makers are aware of this risk and minimize potential biases at every stage in the process of product development**.

In particular, they should consider the risk for biases when deciding:
(1) which ML technologies/procedures they want to use to train the algorithms and
(2) what datasets (including considering their quality and diversity) they want to use for the programming

Several real-world examples have demonstrated that algorithms can exhibit **biases** that can result in **injustice with regard to ethnic origins and skin color or gender.**

Ethical challenges

Biases can also occur regarding other features such as **age or disabilities**.

The explanations for such biases can, for example, result from

- the datasets themselves (which are not representative),
- how data scientists and ML systems choose and analyze the data, and
- the context in which the AI is used, etc.

In the health sector,
where phenotype- and sometimes genotype related information are involved,
biased AI could,
for instance,
lead to false diagnoses and render treatments ineffective for some subpopulations and thus jeopardize their safety.

Ethical challenges

For example,
imagine an AI-based clinical decision support (CDS) software that helps clinicians ***to find the best treatment for patients with skin cancer****.*

If the algorithm is trained on ***Caucasian*** *patients,*
the AI software will likely give less accurate or even inaccurate recommendations for subpopulations for which the training data was underinclusive such as ***African American****.*

Some of these biases ***may be resolved*** *due to*

- *increased data availability,*
- *attempts to better collect data from minority populations, and*
- *better specify for which populations, the algorithm is or is not appropriately used*

Ethical challenges

A remaining problem is that a variety of algorithms are sophisticated and nontransparent.

In addition, some companies developing software will resist disclosure and claim trade secrecy in their work.

*In cases of "black-box" algorithms, many scholars have argued that **explainability is necessary when an AI makes health recommendations**, especially also to detect biases.*

However, does this view really hold true?

*Some argue that **what matters** is not how the AI reaches its decision but that it is accurate, at least in terms of diagnosis.*

Ethical challenges

A related problem has to do with **where AI will be deployed**.

AI developed for top-notch experts in resource-rich settings, will not necessarily recommend treatments that are accurate, safe, and fair in low-resource settings.

One solution would be not to deploy the technology in such settings. **But such a "solution" only exacerbates preexisting inequalities**.

More thought must be given to regulatory obligations and resource support to make sure that this technology does improve not only the lives of the people living in high-income countries but also of those people living in low- and middle-income countries.