Tidsskriftet
Tidsskrift for Den norske legeforening

A critical juncture for the integration of artificial intelligence

DAMOUN NASSEHI

damoun.nassehi@uib.no
Damoun Nassehi, GP and associate professor at the Department of Global Public Health and Primary Care, University of Bergen.
The author has completed the ICMJE form and declares the following conflicts of interest: He is a former member of the Norwegian Board of Technology and has invested in global index funds that include shares in technology companies involved in artificial intelligence.

MICHAEL RIEGLER

Michael A. Riegler, professor at OsloMet – Oslo Metropolitan University and researcher at SimulaMet.
The author has completed the ICMJE form and declares no conflicts of interest.

A comprehensive approach is needed to ensure that artificial intelligence benefits both patients and healthcare professionals.
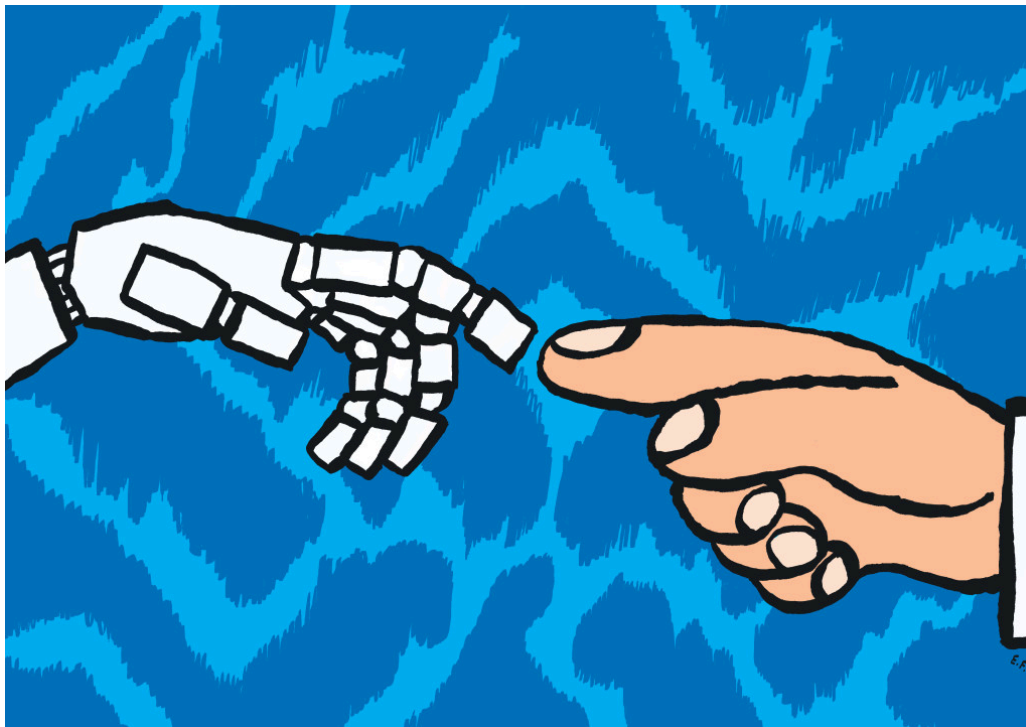
Illustration: Espen Friberg

Artificial Intelligence (AI) is transforming healthcare globally, and Norway is no exception. A study published in the *International Journal of Environmental Research and Public Health* showed that the number of articles on AI in medicine has more than doubled from 2015 to 2022 [1]. This growth reflects not only the potential of AI to transform healthcare, but also the complex challenges associated with implementing the technology. Technological advances have made it possible to use AI for everything from image analysis to predicting health outcomes, but without a solid framework, significant ethical and practical problems may arise.

In Norway, we are now at a critical juncture. We must develop a strategy that ensures that AI integration in healthcare not only focuses on efficiency gains but also considers our core values and ethical principles. This is especially important in light of numerous examples of problematic research and exaggerated claims within the field. For instance, a controversial study claimed that diseases could be diagnosed based on images of the tongue. This study proved to have serious methodological shortcomings, both in medical methodology and AI technology [2].

In this commentary, we base our discussion on the ongoing international debate on artificial intelligence in medicine [3], but our primary focus is to discuss the unique challenges and opportunities that arise when implementing AI in Norwegian healthcare.

> *«A central challenge in developing AI systems for healthcare is how we can balance efficiency with core values such as equality and respect for patient autonomy»*

Balancing values in the Norwegian context

A central challenge in developing AI systems for healthcare is how to balance efficiency with core values such as equality and respect for patient autonomy. Norway has a strong tradition of solidarity in healthcare, where everyone, regardless of socioeconomic background, has the right to equal access to health services. This must be reflected in the AI systems implemented in Norwegian healthcare. An AI model that improves diagnostics by reducing waiting times must also safeguard the patient's right to be seen and heard as an individual.

An illustrative example is an elderly patient visiting a doctor with diffuse symptoms. The patient has a great need to share their story, express their concerns, and feel understood. Meanwhile, the doctor is under pressure to make a diagnosis and develop a treatment plan. In such a situation, an AI model could help optimise patient flow and the diagnostic process by analysing large amounts of data faster than the doctor alone could. Nevertheless, there is a risk that the model may overlook nuances in the patient's story or misinterpret important non-verbal cues that could have been important for the correct diagnosis.

This example illustrates the need for AI systems in healthcare not only to focus on algorithmic efficiency but also to preserve the human dimension of patient care. The technology must serve as a tool that strengthens the doctors in their role, rather than replacing the interpersonal aspects. In Norway, we have the opportunity to develop a national consensus on which values should be prioritised in the development of AI for the healthcare sector. This requires a broad and inclusive process, where healthcare professionals, patients, technologists, and ethical experts are all heard. The goal must be to establish a framework that reflects our societal values while leveraging the potential of AI technology to improve healthcare services.

Addressing bias and Norwegian adaptation

One of the biggest challenges with implementing AI is the risk of bias in the models developed. Bias in AI systems is not just a technical issue but also an ethical and clinical problem that can have direct consequences for patient care. An example of this is race-based adjustments in pulmonary function tests (PFTs), where it has long been standard practice to apply different 'normal' values for lung function based on the patient's race. This has raised concerns that such adjustments may underestimate the severity of lung diseases in Black patients, potentially leading to delayed diagnosis and treatment (4).

In Norway, we must be particularly aware of how bias can affect the treatment of vulnerable groups, such as the Sami population, immigrants and patients from various socioeconomic backgrounds. Bias in AI models could reinforce existing health disparities if the dataset used to train the models does not

represent the entire population. For example, if an AI model is developed based on data from a population in which certain groups are underrepresented, the results may be skewed, leading to poorer treatment for these groups.

To counteract bias and ensure fair AI implementation in Norwegian healthcare, we propose several measures.

First, we must ensure representative data collection. Datasets used to train AI models must include all population groups in Norway, including ethnic minorities, different age groups, genders and socioeconomic backgrounds. Specific data collection initiatives should be launched to include underrepresented groups.

Second, regular review is necessary. AI models must be regularly reviewed to identify and correct biases. This means establishing a national system to monitor and evaluate the performance of AI models in clinical practice.

In terms of ethical considerations, interdisciplinary panels comprising clinicians, ethicists, patient representatives and AI experts should assess potential biases and ethical implications before implementing the models.

Moreover, AI models developed outside Norway should be evaluated and adapted to Norwegian conditions before use. This involves not only linguistic translation but also adjustments to Norwegian clinical guidelines, healthcare priorities, and cultural norms.

Lastly, healthcare professionals and patients should be informed about the limitations of AI models, including potential biases, so that they can make informed decisions.

By implementing these measures, we can develop an AI system that is fair and robust, enhancing both the quality of healthcare services and trust in AI among patients and healthcare professionals alike.

Interdisciplinary collaboration

To successfully integrate AI into the Norwegian healthcare system, it is essential to facilitate interdisciplinary collaboration. AI systems used in clinical practice must be technologically advanced, but also ethically sound and clinically relevant. This requires collaboration between technologists, clinicians, ethicists, lawyers and patient representatives.

A model for interdisciplinary collaboration could be to establish multidisciplinary teams involving all relevant stakeholders in the development and implementation of AI, from the conceptual phase to clinical use. These teams can work on everything from technological challenges to ensuring that systems comply with the legal and ethical standards required in healthcare. In parallel, advisory ethical committees should be established to regularly assess ethical questions related to AI.

To ensure that AI solutions are genuinely relevant to Norwegian clinical practice, it is also necessary to educate and employ professionals with expertise in both medicine and AI. This can help ensure that AI systems are adapted to

Norwegian clinical needs, and that healthcare personnel are capable of using the technology responsibly.

Furthermore, we must include patients and patient representatives in the development process for AI systems. This will help ensure that the technology meets real needs and respects patient autonomy.

Long-term implications and a value-driven framework

The integration of AI into Norwegian healthcare will undoubtedly have profound and long-term consequences for both patient care and the relationship between healthcare professionals and patients. As AI takes on a more central role in decision-making processes, we must carefully consider how to maintain patients' trust in the healthcare system.

An important part of this process is to develop a broad ethical framework that reflects the values important in Norwegian healthcare. This includes respect for patient autonomy, where patients remain active participants in their own treatment, even when AI becomes more integrated into decision-making processes. The principle of 'do no harm' must also be a fundamental guideline for the development and implementation of AI, ensuring that the technology does not lead to unintended negative consequences.

> *«AI systems should not replace the relationship between healthcare professionals and patients but rather enhance it»*

To safeguard the interpersonal dimensions of healthcare, we should also include care ethics in our ethical framework. AI systems should not replace the relationship between healthcare professionals and patients but rather strengthen it. This will be key to ensuring that patients feel equal and active participants in their own treatment, even when AI technology becomes part of the care process.

Transparency and explainability

An important aspect of AI implementation is to ensure that systems are transparent and explainable. For AI to be a reliable part of healthcare, healthcare professionals as well as patients must understand how the systems work and how they reach their conclusions. Without sufficient explainability, AI can create mistrust and uncertainty among both healthcare professionals and patients.

To achieve this, we propose implementing *model cards* as a standard practice in the Norwegian healthcare sector. These documents, which can be compared to safety data sheets in industry, provide a comprehensive overview of an AI model's performance, training data, methodology and potential limitations. By giving healthcare professionals access to this information, we enable informed

decisions about when and how AI models should be used in clinical practice. This not only contributes to increased transparency but also to a more responsible and ethical use of AI in healthcare.

Furthermore, we need to emphasise evaluating the reasoning behind AI models' decisions. Recent research has shown that even advanced models, such as GPT-4 V, can provide correct answers based on flawed reasoning (5). In a Norwegian healthcare context, where patient safety and trust in the healthcare system are paramount, it is critical to evaluate not only the accuracy of AI models' outputs but also the quality of their reasoning.

Another important factor is the communication of uncertainty (6). AI models in Norwegian healthcare should be capable of conveying the level of uncertainty in their predictions and recommendations. This will provide healthcare professionals with a more nuanced basis for decision-making and help maintain healthy scepticism towards AI-assisted decisions.

> *«Ethical evaluation should be an integrated part of the AI plan, and we propose establishing mechanisms for continuous evaluation and adjustment of AI systems in use»*

Evaluation of the Directorate of Health's AI plan

The Directorate of Health's AI plan is an important step towards the systematic use of AI in Norwegian healthcare. However, we believe that the plan could have gone further in defining a clear ethical framework for AI implementation. Ethical evaluation should be an integrated part of the AI plan, and we propose establishing mechanisms for continuous evaluation and adjustment of AI systems in use.

Furthermore, we recommend strengthening the plan with concrete measures to balance different considerations, such as efficiency and privacy, as well as equity in access and individualised treatment. This will ensure that AI solutions not only improve efficiency but also safeguard patients' rights and interests.

The implementation of AI in the Norwegian healthcare system represents both a significant opportunity and a considerable challenge. Successful integration requires more than just technological expertise—it requires a comprehensive ethical and interdisciplinary framework that ensures the technology is used fairly and responsibly. By developing AI solutions that respect Norwegian values and safeguard patient autonomy, we can ensure that healthcare becomes both more efficient and more patient-centred.

REFERENCES

1. Karpov OE, Pitsik EN, Kurkin SA et al. Analysis of Publication Activity and Research Trends in the Field of AI Medical Applications: Network Approach. Int J Environ Res Public Health 2023; 20: 5335. [PubMed][CrossRef]

2. Nassehi D. Kunstig intelligens og pseudovitenskap er en farlig kombinasjon. Dagens Medisin 20.8.2024. https://www.dagensmedisin.no/forskning-kunstig-intelligens-medisin/kunstig-intelligens-og-pseudovitenskap-er-en-farlig-kombinasjon/647826 Accessed 23.8.2024.

3. Yu K-H, Healey E, Leong T-Y et al. Medical Artificial Intelligence and Human Values. N Engl J Med 2024; 390: 1895–904. [PubMed][CrossRef]

4. Braun L, Fausto-Sterling A, Fullwiley D et al. Racial categories in medical practice: how useful are they? PLoS Med 2007; 4. doi: 10.1371/journal.pmed.0040271. [PubMed][CrossRef]

5. Jin Q, Chen F, Zhou Y et al. Hidden flaws behind expert-level accuracy of multimodal GPT-4 vision in medicine. NPJ Digit Med 2024; 7: 190. [PubMed][CrossRef]

6. Huang Y, Song J, Wang Z et al. Look Before You Leap: An Exploratory Study of Uncertainty Measurement for Large Language Models. arXiv 2023 doi: 10.48550/arXiv.2307.10236. [CrossRef]