Tidsskriftet Tidsskrift for Den norske legeforening

Artificial intelligence in clinical practice – an industry without raw materials

PERSPECTIVES

VIBEKE BINZ VALLEVIK

vibeke.binz@dnv.com
Vibeke Binz Vallevik, MSc, project manager in the BigMed Research Project at Oslo University Hospital and principal researcher in digital health at *Det Norske Veritas*.
The author has completed the ICMJE form and declares no conflicts of interest with the exception of her own employer.

PÅL H. BREKKE

Pål H. Brekke, PhD, specialist in internal medicine and heart diseases, senior consultant and researcher at the Department of Cardiology at Oslo University Hospital, Rikshospitalet.
The author has completed the ICMJE form and declares no conflicts of interest.

Research on health data is unfortunately guided by the available data rather than the clinical problems we need to solve. Clinically-related data are locked away in silos. As a result, both patients and research are losing out.

Randomly selected faces generated by artificial intelligence via the website thispersondoesnotexist.com. None of the faces belong to real people – they are automatically composed from many different images. Montage: Journal of the Norwegian Medical Association

The year is 2017, and a proud new project manager has just arrived at a conference in Cambridge to present Norway's biggest initiative in precision medicine at the same session as giants such as Genomics England and the USA's All of Us Research Program. This is the BigMed initiative (1) – a national project funded by the Research Council of Norway aimed at raising awareness of the barriers to the clinical implementation of precision medicine and paving the way for the analysis of big data in clinical practice. During the presentation we are asked, 'How did you select these particular patient groups?' The honest yet embarrassing answer is that we assumed that these datasets would be the easiest to access.

Unfortunately, years later, we now know that this assumption was naive. The process took three years, a paradox for a project that had four years to lay the foundation for big data analysis in clinical practice. BigMed's conclusion is that it can be surprisingly difficult to gain access to health data for analytical purposes. As a result of the experience gained in this project, Norway has taken a first step towards simplifying access by amending health legislation, but many challenges remain. Gaining access to health data in Norway still requires an inordinate amount of time and energy.

The precautionary principle hampers development

Artificial intelligence has significant potential to improve the health sector through increased efficiency in administrative health systems, medical research, clinical decision support and public health activities (2, p. 14). With its 22 partners and approximately 100 researchers, the BigMed project was launched as a national lighthouse project, developing and testing methodology and tools for using artificial intelligence in three clinical areas. Since data-driven decision support requires a significant body of data, we applied for access to pseudo-anonymised data from a large number of patients. The application process turned into a game of snakes-and-ladders, going one step forward then two steps back. In fact, this was a multi-purpose project spanning health research, methodology development, quality assurance and efficiency improvements – or all these at once. Since the Norwegian laws that govern health data are fragmented and have different statutory purposes, our project spanned more than one law.

We were bounced around like a ping pong ball between various authorities with different interpretations of the regulations – between the Norwegian Data Protection Authority, the Regional Committees for Medical and Health Research Ethics (REC), the National Committee for Medical and Health Research Ethics (NEM), Oslo University Hospital's own legal department, the hospital's data protection officer and the hospital's IT security officer as well as the hospital IT provider *Sykehuspartner* for risk analyses.

> ***«We were bounced around like a ping pong ball between various authorities with different interpretations of the regulations»***

The principle of data minimisation in the General Data Protection Regulation (GDPR) – limiting the amount of personal data collected – makes the use of machine learning and artificial intelligence challenging. Large datasets are needed, but their utility may not be clear at the outset of the research, as the rationale for using artificial intelligence is often to find trends in the data that we are unable to identify with our own eyes. For BigMed, this resulted in us minimising the data to two small datasets just to get started. As a result of this minimisation, we had to exclude most of the clinically useful data in these projects. After three years had passed, there was no time left for the planned research.

The BigMed experience shows that there is considerable confusion about the interpretation of the legislation, particularly when it comes to the GDPR. The precautionary principle in the case of data protection together with a lack of overall understanding of the legislation governing health data makes for unduly complex and bureaucratic processes. It would be interesting to calculate what this costs us in terms of both delays and resource use, not to mention the lost innovation potential from activities that never get off the ground.

The value of the raw data

Artificial intelligence has significant potential to help us identify new correlations and gain new insights from large, complex datasets. The BigMed project concludes that this requires access to the raw material, i.e. the raw data, not summaries of the conclusions (3, p. 22–23). The raw data produced in the health sector contains considerable reserves of knowledge and insight, but much of it is not available for reuse. For example, Oslo University Hospital's current practice is to save a scanned image of an ECG printout rather than the raw data describing the curves. This is not due to a lack of space. The raw data can be used for analyses and requires only 200 kB storage. Although the scanned PDF copy cannot be reused, it requires almost 240 times more storage space. The development of a national Health Analysis Platform (HAP) aims to improve access to the large national health registers and quality registers, and in recent years politicians have regarded this as meeting the need for health data. Register data are interesting but constitute a mere fraction of the health data we produce. Existing patient record systems and specialist systems contain millions of patient pathways, diagnostic images, observations, findings and treatments that can deepen our understanding of disease and treatment processes, and at the same time add fresh knowledge about how the health service actually functions. A more detailed representation of patient data can provide us with more answers than the few selected measuring points that are reported to the national registers (4).

Clinically-related data can form the basis of informed decisions in a data-driven health service. The need for real-time clinical data and a clinically-related infrastructure to make them accessible has become evident to the wider public during the COVID-19 pandemic.

The EU wants to turn us into a data-driven society (5). Such an innovation requires access to raw data, not just register data. Today, large hospitals that lack the infrastructure to extract data for the development of artificial intelligence in internal projects will find it difficult to prioritise industrial partners in clinical trials. At Norway's largest hospital, a dozen or so important clinical trial projects are waiting in a data access queue with no start date in sight.

In Norway, significant efforts and multiple investments have so far focused on the centralised quality registers in HAP, but have failed to target clinical data at the hospitals. What implications for society does the inaccessibility of the raw data have in terms of lost knowledge?

Need for changes in Norwegian health legislation

BigMed quickly identified the importance of focusing on the legal barriers in order to pave the way for the use of big data and artificial intelligence in clinical practice (6, p. 47). By uniting the research community and investing in legal research, we contributed to a change in Norwegian law as a direct consequence of raising these issues. In June 2021, the Norwegian parliament amended the Health Personnel Act (*Helsepersonnelloven)* and the Health Records Act (*Pasientjournalloven*) to allow data access to clinical decision support based on artificial intelligence (7).

Despite this being a cause for celebration, the amendments still represent an exception from the fundamental principle that health data should not be reused. In other words, although this is now legally permitted, in practice it remains difficult because the use of health data is governed by different laws. Is it possible to build a data-driven health system on such a foundation? We do not believe so. A fragmented body of regulations that is interpreted differently by the various authorities throughout Norway makes it even more challenging for healthcare personnel and researchers to assess whether they are breaking any laws during the course of their daily work. We should therefore avoid new exemptions that lead to greater complexity and a wider scope of interpretation. We should instead be bold enough to change our perspective so that the data collected by the public health service *will* be used to provide patients with better, evidence-based treatment.

Several countries have put in place strategies for using artificial intelligence and a roadmap for implementation. Finland introduced its strategy as early as 2017 (8) and has adopted a number of measures (9) – legal measures (10) and infrastructure investments, as well as an organisation that facilitates better access to data. The effects of these changes are now benefiting users at Helsinki University Hospital, where both clinical data and production data are available for predefined time periods for research, development and improvement of the health services .

What do patients want?

The primary task of the health service is to provide healthcare. As such, its value – patient safety and high-quality treatment – should be a key factor in discussions on the use of health data. Nonetheless, we find that the debate is often skewed towards protecting data rather than using it to protect our patients.

At present, a clinician in Norway cannot turn to data systems for help or make use of the collective knowledge obtained by the hospital through the treatment of hundreds or thousands of similar patients. The therapist can only use the data stored in their own head. In order to offer the best possible personalised treatment, we have to learn from everything we do as we go along.

> ***«We find that the debate is often skewed towards protecting data rather than using it»***

We owe it to our patients to better manage the data collected by the public health system. As patients, the majority of us want our treatment to be based on the complete, collective expertise and experience of the health service in relation to our condition.

Experience of the recruitment of patients for health research shows that the majority want knowledge about the condition to be shared so that the next patient can benefit from it. In 2019, the Norwegian Cancer Society's survey under the auspices of BigMed showed that 96 % of respondents agreed that data about them should be used to help other patients. The Directorate of e-health's survey of citizens showed that more than 90 % thought it was important that health data is used to improve existing knowledge and generate new knowledge (11, p. 18). Patients are often surprised that they have to consent to their data being collected and analysed – they assume that therapists and researchers automatically have access. Consequently, patients must be included in the public debate on what should be part of the social contract when we receive healthcare in Norway.

We need an ambitious goal

When both healthcare staff and patients find that the regulation of health data poses an obstacle to evidence-based improvement, it is time to act. We must shift the discussion from *whether* we should use the health data to *how* we should use it in a safe and secure manner that protects the patient. The government's Hurdal Platform (12, p. 59) states that the government will facilitate better use of health data in health research and strengthen the foundation for evidence-based health and care services. This is admirable but not enough on its own. We call for a more ambitious vision that can unite us. There are many pieces to this puzzle that must be solved in the various parts of our health system. Therefore, we need a clear, common goal for the reuse of health data that we can all work towards. This can facilitate an overarching regulation of health data rather than fragmentation arising from the different statutory purposes.

We propose a new goal for the Norwegian health sector: that by 2025, a total of 95 % of all health data collected by Norwegian hospitals should be available for analysis and the production of new knowledge.

Only then can we discuss what changes are needed to reach this goal – in policies, laws and regulations, as well in clear strategies, plans and practical implementation in the health service.

REFERENCES

1. 2022. BigMed. www.bigmed.no Accessed 31.3.2022.

2. Teknologirådet. Kunstig Intelligens i klinikken, Seks trender for fremtidens helsetjeneste. 2021. https://teknologiradet.no/wp-content/uploads/sites/105/2022/02/Kunstig-intelligens-i-klinikken.pdf Accessed 31.3.2022.

3. Vallevik VB, Zaka A, Ray-Sannerud B et al. Reflections on the clinical implementation of precision medicine – Experiences from BigMed, a Norwegian ICT Lighthouse project. Oslo universitetssykehus, 2021. https://bigmed.no/assets/bigmed_reflections-on-the_2021_v1.0.pdf Accessed 31.3.2022.

4. El-Galaly A, Grazal C, Kappel A et al. Can Machine-learning Algorithms Predict Early Revision TKA in the Danish Knee Arthroplasty Registry? Clin Orthop Relat Res 2020; 478: 2088–101. [PubMed][CrossRef]

5. European Commission. Communication from the commission to the european parliament, the council, the european economic and social committee of the regions: A European strategy for data. COM/2020/66 final. Brussels, 19.2.2020. https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf Accessed 31.3.2022.

6. BigMed. Big Data Management for the precise treatment of three patient groups. Oslo universitetssykehus, 2018. https://bigmed.no/assets/Reports/Big_data_management_for_the_precise_treatment_of_three_patient_groups.pdf Accessed 31.3.2022.

7. Lovvedtak 142 (2020–2021). Vedtak til lov om endringer i helsepersonelloven og pasientjournalloven (bruk av helseopplysninger for å lette samarbeid, læring og bruk av kunstig intelligens i helse- og omsorgstjenesten mv.) https://lovdata.no/static/LOVVED/lovved-202021-142.pdf Accessed 31.3.2022.

8. Ministry of Economic Affairs and Employment in Finland. Finland's age of Artificial Intelligence – Turning Finland into a leading country in the application of artificial intelligence. Objective and recommendations for measures. 47/2017. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y Accessed 31.3.2022.

9. OECD. OECD.AI policy dashboard: Policy initiatives for Finland. https://oecd.ai/en/dashboards/policy-initiatives?conceptUris=http:%2F%2Fkim.oecd.org%2FTaxonomy%2FGeographicalAreas%23Finland Accessed 26.1.2022.

10. Ministry of Social affairs and health in Finland. Act on the secondary use of health and social data (552/2019). https://stm.fi/en/secondary-use-of-health-and-social-data Accessed 31.3.2022.

11. Direktoratet for e-helse. Helsedataprogrammets innbyggerundersøkelse 2019. https://www.ehelse.no/publikasjoner/helsedataprogrammets-innbyggerundersokelse-2019 Accessed 31.3.2022.

12. Hurdalsplattformen. For en regjering utgått fra Arbeiderpartiet og Senterpartiet. 2021–2025. https://www.regjeringen.no/contentassets/cb0adb6c6fee428caa81bd5b339501b0/no/pdfs/hurdalsplattformen.pdf Accessed 31.3.2022.

Publisert: 28 June 2022. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.22.0084
Received 28.1.2022, first revision submitted 6.4.2022, accepted 28.4.2022.