Discussion Paper:
Social accountability for students in a machine learning era

L. Z. J. Williams[1] & R. Grainger[2, 3, 4]

Abstract

Over the last 30 years, there have been repeated calls to integrate health informatics into undergraduate health professional curricula, in recognition of the integral role computing plays in medicine. The rise of big data sets in health, and the application of advanced computer algorithms to interrogate these, is yet another call for health professionals to receive appropriate training in these technologies.

Machine learning (ML) algorithms can learn tasks or make decisions without a requirement for specific behaviours to be pre-programmed. High-impact literature has described ML approaches to clinical problems such as achieving more accurate and timely diagnoses, increasing precision of prognosis and guiding treatment. Despite the promise of ML in healthcare, there are risks of adverse outcomes, unanticipated consequences, misuse and even abuse of ML technologies. For health professionals to advocate for patients and hold those developing ML algorithms in healthcare accountable, they must feel comfortable discussing the fundamental concepts and limitations of ML in healthcare.

Healthcare professionals are uniquely positioned to identify problems that could be solved by ML and related technologies. Yet, there is inadequate coverage of ML, or of the wider field of health informatics, in most medical curricula. To create future health professionals who can advocate for positive change and ensure that patients remain at the centre of ML applications in healthcare, we must provide future health professionals with an understanding of how ML will change healthcare delivery and the doctor–patient dynamic, as well as new ethical challenges that arise with the digital healthcare revolution.

[1] Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
[2] Department of Medicine, University of Otago, Wellington, New Zealand
[3] Department of Pathology and Molecular Medicine, University of Otago, Wellington, New Zealand
[4] Department of the Dean, University of Otago, Wellington, New Zealand

Correspondence
Logan Zane John Williams
Becket House
1 Lambeth Palace Road
London SE1 7EU
United Kingdom
Tel: +44 73 9154 3961
Email: logan.williams@kcl.ac.uk

Keywords: machine learning; social accountability; medical education; curriculum

Introduction

Over the last 30 years, there have been repeated calls to integrate health informatics into undergraduate health professional curricula, in recognition of the integral role computing plays in medicine. Health informatics is "the study and application of methods to improve the management of patient data, medical knowledge, population data and other information relevant to patient care and community health" (Wyatt & Liu, 2002, p. 1). Recent advances in information technology now make these calls urgent. In 2019, the National Health Service (NHS) (2019) in the United Kingdom urged that health professional curricula educate future graduates on "the possibilities of digital healthcare technologies and the ethical and patient safety considerations" (p. 17) and recommended genomics, data analytics and artificial intelligence (AI) as essential learning in undergraduate health professional curricula. Healthcare now has big data sets that require advanced computer algorithms for analysis, and health professionals need the skills to critically assess the outputs produced from these approaches. The potential for patient harm from the inappropriate use of computer algorithms in healthcare also makes this a training priority for health professionals.

In this paper, we outline the promise and pitfalls of AI and machine learning (ML) in healthcare, which underpin our argument that to equip graduates with skills for the future of medicine and to meet social accountability responsibilities, health professional curricula must include critical elements of health informatics to address the increasing importance of ML and other technologies in healthcare.

What are artificial intelligence and machine learning?

The digitisation of modern life is creating large data sets, commonly referred to as "big data". In addition to their size, big data sets are diverse and rapidly generated, meaning that their manipulation and management are often logistically challenging. AI, the development of intelligent systems using computational techniques, is usually required for insights into and utility from big data (NHS, 2019). Everyday examples of AI include Google's search algorithms, which provide customised search results, and Uber's "route-based pricing" system, which adjusts prices for sociological factors. These tasks often require ML algorithms, which are algorithms that can learn tasks or make decisions without a requirement for specific behaviours to be pre-programmed (NHS, 2019). These algorithms lie on a spectrum between completely human-specified and completely machine-specified, each with advantages and disadvantages. ML algorithms are already providing insights and knowledge from health data, with the aim of achieving better health outcomes.

Emerging machine learning successes in healthcare

High-impact literature has described ML approaches to clinical problems such as achieving more accurate and timely diagnoses, increasing precision of prognosis and guiding treatment. Recently, an ML algorithm trained on optical coherence tomography scans exceeded experts' ability to recognise macular pathology (Topol, 2019). Another ML algorithm for sepsis management was recently trained and tested on routinely collected data from almost 100,000 patients in the United States to recommend doses of intravenous fluid and vasopressors (Topol, 2019). The authors found that human prescribing that differed from ML algorithm recommendations was associated with a dose-dependent increase in mortality (Topol, 2019). With increasing sources of routinely-collected structured health data, digital images and diagnostic techniques, there are endless opportunities for ML to improve healthcare and reduce costs. The hope is that intelligent use of technologies such as ML-driven decision-support systems will not only improve patient outcomes but also reduce the administrative burdens of health professionals, allowing more time for meaningful interactions with patients.

Social accountability and machine learning

Despite the promise of ML in healthcare, there are risks of adverse outcomes, unanticipated consequences, misuse and even abuse of ML technologies. Health professionals are expected to demonstrate social accountability, defined as the "obligation to direct education, research and service activities towards addressing the priority health concerns of the community, region, and/or nation they have a mandate to serve" (Rourke, 2018, p. 1). For health professionals to advocate for patients, and hold those developing ML algorithms for healthcare accountable (which include internet and health insurance companies), current curricula need to extend the knowledge and critical appraisal skills of health professional graduates to include ML applications in healthcare. The current concerns around ML include the lack of regulation and prospective validation, risk of bias and data governance, which are outlined with examples in the following paragraphs.

Risks of machine learning and role of regulation and validation

The excitement surrounding ML exceeds its readiness for some healthcare settings, partly due to the lack of regulation and prospective validation in real-life contexts. For instance, IBM Watson for Oncology (WFO) algorithms have been used in hospitals worldwide to recommend cancer treatments. Currently, however, no prospective trials have determined if WFO improves patient outcomes above routine care. Moreover, there is wide variation in concordance of treatment recommendations varying from 49% for colon cancer to 96% for ovarian cancer and 12% for gastric cancer (Topol, 2019).

Outside the hospital, smartphones now provide direct-to-consumer ML apps. A high profile study of an ML algorithm for identifying malignant skin lesions from photographs found that the algorithm performed on par with expert dermatologists and suggested that apps could improve accessibility to care. However, another study

has found that three out of four smartphone apps incorrectly classified 30% or more of melanomas (Wolf et al., 2013). Moreover, the recent downgrading of the United States Food and Drug Administration regulatory requirements for apps makes market entry easier (Topol, 2019). Given the potential for harm, future health professionals need to protect patient safety and demand robust validation of these technologies, especially in local contexts.

Risk of machine learning driving commodification of data

As healthcare expenditure rises, there is an imperative to develop financially sustainable healthcare models that gain efficiencies while keeping patients safe. ML applications for reducing medication errors and increasing workflow efficiency could save the US healthcare system an estimated $150 billion USD annually by 2026 (Kalis et al., 2018). The development of these ML algorithms will require developers to access large data sets. While healthcare providers may supply data to third party developers for altruistic reasons, ethical and legal obligations must be considered. For example, to develop algorithms for early detection of acute kidney injury, the Royal Free Hospital in London gave Google's DeepMind team access to 1.6 million identifiable patient records without obtaining explicit consent for data release (Shah, 2017). Society already suffers a "data trust deficit", so lack of transparency and openness could limit public and health system acceptance of data sharing. Open discussion with the public about responsible and acceptable use of their data to benefit healthcare is essential.

Risks of data (in)equity

The way in which data are collected and presented is influenced by the healthcare team collecting that data. Moreover, the data are only representative of those who can access or afford treatment, which potentially impacts on the development of ML algorithms for diagnosis or treatment. Just as doctors are repeatedly shown to exhibit implicit biases in healthcare, cognitive biases affect approaches to big data. One prevalent cognitive bias is that the use of big data sets substitutes for thorough study design, data collection and analysis. Even when data are intentionally collected, a strong Euro- and male-centric approach to research design and data collection may leave women and minorities underrepresented. Genomic data explaining pathophysiology and identifying therapeutic targets have repeatedly underrepresented minorities (Topol, 2019). Facial recognition technology, which could be used for diagnosis or recognition of clinical deterioration, is potentially biased against those with darker skin tones, especially females, as these algorithms are predominantly trained on males with light skin tones (Topol, 2019). Future algorithms need to be developed on data sets that represent all members of the population on which these will ultimately be applied to, as failure to do so could widen health inequities.

Are our graduates ready for ML in healthcare?

Healthcare professionals are uniquely positioned to identify problems that could be solved by ML and related technologies. Yet, they must be confident navigating these technologies and communicating with field experts. Moreover, as ML becomes more routinely applied to healthcare, health professionals will need to critically appraise the benefits and risks of such applications. This oversight is critical, as these technologies may perpetuate health inequities due to biased data sets, be used unethically due to financial conflicts of interest or be implemented without sufficient evidence.

Despite the growing ML literature and coverage of ML in mainstream news, there is inadequate coverage of ML, or of the wider field of health informatics, in most medical curricula. Most UK medical schools have health informatics in the curriculum, however only 57% assessed health informatics content, and only 41% of schools updated content regularly (Walpole et al., 2016). In Australia, medical school education leaders all agree on the theoretical importance of eHealth to current and future clinical practice but report no strong internal drivers for its inclusion in the curriculum (Edirippulige et al., 2018). The challenges of a full curriculum and lack of subject matter expertise need to be addressed to equip future health professionals with the knowledge and critical appraisal skills they need to ensure ML applications in healthcare improve patient outcomes and address inequities.

Where to from here?

To create future ready health professionals, we must introduce learning opportunities for ML into health professional curricula, with an emphasis on critical appraisal and equity. Curricula must be developed with input from all relevant stakeholders, including healthcare students, curricular design staff, professional and regulatory bodies, patient groups and field experts (Rourke, 2018). In designing an ML curriculum, it is important to distinguish between information that students need to know and information they need to be aware of (NHS, 2019). Given that an ML curriculum will be somewhat dependent on new healthcare technologies, and digital literacy will evolve as technology becomes increasingly integrated into daily life, we must also ensure curricula are continually renewed.

An ML curriculum might be tailored in order to meet student interest. A compulsory component might aim to foster an understanding of how genomics and other "omic" technologies, health informatics and AI will change healthcare delivery and the doctor–patient dynamic, as well as new ethical challenges that will arise with the digital healthcare revolution (NHS, 2019). It is important that this compulsory curriculum has an experiential component so students have the chance to see first-hand the impact of current state-of-the-art ML technologies on patient care. Optional curricular components for students with higher interest might include software coding, data analytics and management science. For example, a short course on software coding was highly valued by medical students, who felt that coding was a relevant skill for

doctors (Morton et al., 2019). In the UK, it has been suggested that expertise in ML and related fields could be achieved through intercalated degrees in engineering or computer science (NHS, 2019).

Conclusion

It is the responsibility of current and future health professionals to advocate for positive change and to ensure that patients remain at the centre of healthcare. It is essential that ML is used in a conscientious manner to create a cost-effective healthcare system that delivers equitable outcomes, and future health professionals should have the knowledge and skills to achieve and maintain this. By incorporating teaching on ML and other technologies into their curricula, schools for health professionals can drive the digital healthcare revolution.

Funding and conflicts of interest statement

References

Edirippulige, S., Brooks, P., Carati, C., Wade, V. A., Smith, A. C., Wickramasinghe, S., & Armfield, N. R. (2018). It's important, but not important enough: eHealth as a curriculum priority in medical education in Australia. *Journal of Telemedicine and Telecare*, *24*(10), 697–702. https://doi.org/10.1177/1357633X18793282

Kalis, B., Collier, M., & Fu, R. (2018). 10 promising AI applications in health care. *Harvard Business Review*. https://hbr.org/2018/05/10-promising-ai-applications-in-health-care

Morton, C. E., Smith, S. F., Lwin, T., George, M., & Williams, M. (2019). Computer programming: Should medical students be learning it? *JMIR Medical Education*, *5*(1), e11940–e11940. https://doi.org/10.2196/11940

National Health Service (NHS). (2019). *The Topol review: Preparing the healthcare workforce to deliver the digital future*. https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf

Rourke, J. (2018). Social accountability: A framework for medical schools to improve the health of the populations they serve. *Academic Medicine*, *93*(8), 1120–1124. https://doi.org/10.1097/ACM.0000000000002239

Shah, H. (2017). The DeepMind debacle demands dialogue on data. *Nature News*, *547*(7663), 259.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Walpole, S., Taylor, P., & Banerjee, A. (2016). Health informatics in UK medical education: An online survey of current practice. *JRSM Open*, *8*(1). https://doi.org/10.1177/2054270416682674

Wolf, J. A., Moreau, J. F., Akilov, O., Patton, T., English, J. C., III, Ho, J., & Ferris, L. K. (2013). Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA Dermatology*, *149*(4), 422–426. https://doi.org/10.1001/jamadermatol.2013.2382

Wyatt, J. C., & Liu, J. L. Y. (2002). Basic concepts in medical informatics. *Journal of Epidemiology and Community Health*, *56*(11), 808–812. https://doi.org/10.1136/jech.56.11.808