A Governance Framework for Artificial Intelligence Tools in Medicine

Jordan Richardson, Richard Sharp, PhD

Mayo Clinic, Rochester MN

The development and implementation of artificial intelligence (AI) tools in medicine has been met with and compelled forward by considerable enthusiasm over the past few years. Much of this enthusiasm has heralded AI as shaping the future of diagnostics, decision support, learning health systems and more. Amongst this hype, some scholars have raised some early governance concerns about various types and applications of AI in medicine due to their diversity and complexity. However this paper is the first to provide a comprehensive framework for understanding regulatory considerations potent to the development and use of discrete AI tools. Our analysis focuses on five primary domain areas: Consent and Patient Protections, Roles of Healthcare Providers, Encoding Bias, Accountability, and Transparency. These areas emerged as those most likely to be put in jeopardy by the introduction of AI tools, and therefore the most crucial to be conscientious of while designing governance aimed at ensuring responsible development and implementation. Our evaluation of consent and patient protections details the possible harms that individuals and groups assume when their data is used for training AI algorithms, and how current data stewardship norms are insufficient for mitigating the harms inherent in research involving Big Data and Machine Learning. When discussing the shifting roles of healthcare providers, we emphasize that, even without alarmist ideas of a robot dystopia, introducing AI tools into clinic will necessarily alter the responsibilities of providers and the patient/provider relationship. Doing this in a controlled and intentional way is crucial for maintaining medical humanism and patient trust. Our analysis of bias looks past the axiom of "garbage in, garbage out" data sets and provides insight into ways that the machine learning technologies and their adaption to clinical settings can themselves encode and generate inequality, as well as the future implications should these concerns not be addressed. Finally, we consider the various different stakeholders in medical AI tools and the novel and complex notions of accountability required to have ethical and effective governance with this specific emerging and disruptive field, as well as addressing the complex tradeoffs between proprietary and regulatory interests inherent in transparency. For each of these domains, we raise larger epistemological and values based questions that are necessary to confront in any effective governance model as medicine expands forwards into the full potential of AI technologies. Our proposed framework is unique and critical because its comprehensiveness and grounded applicability illuminates a need for not only one, but multiple governance structures to effectively regulate the diversity of technologies throughout the different stages of development, early implementation, and long-tern oversight. We believe these considerations are crucial and pertinent to not only formal regulators, but any individual or entity involved in evaluating AI tools in a healthcare setting.