Privacy Implications of AI-Enabled Predictive Analytics in Clinical Diagnostics, and How to Mitigate Them

By Dessislava S. Fessenko, LL.M, MSc.BE, CIPP/E

Abstract: AI-enabled predictive analytics is widely deployed in clinical care settings for healthcare monitoring, diagnostics and risk management. The technology may offer valuable insights into individual and population health patterns, trends and outcomes. Predictive analytics may, however, also tangibly affect individual patient privacy and the right thereto. On the one hand, predictive analytics may undermine a patient's state of privacy by constructing or modifying their health identity independent of the patient themselves. On the other hand, the use of predictive analytics may violate the patient's right to privacy if the patient has no control over the use or output of the technology. These repercussions ultimately erode patient autonomy and agency. This paper discusses these implications in further detail, and proposes possible measures for their mitigation. They involve the incorporation in the AI systems of accuracy-enhancing statistical models and methods, more privacy-conscious institutional policies and practices, and effective choice for patients to accept or refuse diagnostics and treatment drawing on AI-enabled predictive analytics.

Privacy Implications of AI-Enabled Predictive Analytics in Clinical Diagnostics, and How to Mitigate Them

I. Introduction

The use of artificial intelligence (AI)-enabled predictive analytics in clinical diagnostics holds a significant potential for early detection and timely treatment of various serious health conditions. The technology, however, has its privacy implications as well. On the one hand, it may undermine a patient's state of privacy by constructing or modifying their health identity independent of the patient themselves. On the other hand, the use of predictive analytics may violate the patient's right to privacy if the patient has no control over the use or output of the technology. These repercussions ultimately erode patient autonomy and agency. This paper discusses these implications in further detail, and proposes possible measures for their mitigation. The measures include: (i) the incorporation in the AI systems of accuracy-enhancing statistical models and methods, (ii) more patient-friendly privacy policies and practices that adequately inform patients of the use, essence and possible implications of predictive analytics for their diagnostics and treatment, and (iii) effective choice for the patient to consent to or reject the use of predictive analytics for their diagnostics and treatment.

II. The Concept of Privacy

To orient the discussion, it is worth recounting the concept of "privacy" first. Beauchamp and Childress distinguish privacy from the right to privacy [1, p. 338]. The former signifies a "*state or condition.*" The latter entails control over being in that state, i.e. one's authority or powers to

allow or restrict access to one's privacy. The distinction matters because, as Beauchamp and Childress note, one may be in the state of privacy without formally having control over (i.e. the right to) it. And vice versa, I would add: one may formally and ostensibly have control – by exercising one's right to – privacy and yet not be in the state thereof. This paper will be concerned with both the state of and the right to privacy, as the use of AI-enabled predictive analytics in clinical diagnostics may affect them both.

The state of privacy is considered to denote the condition of limited access to the self [1, p. 338]. The various dimensions of privacy, however, suggest that the concept is potentially broader. Some dimensions are indeed purely corporeal, e.g. physical and proprietary privacy, which entail the integrity of one's body, biological materials, images, and private spaces [2]. Other dimensions are more intangible, so to speak. For example, informational privacy involves limited access to information about one's personality and affairs (e.g. emotions, thoughts, secrets) [3]. Decisional privacy emphasizes the solitude of one's personal choices [1, p. 338; 3, p. 5]. Relational privacy concerns the intimacy of one's personal relations and of their influence on one's decision-making (e.g. shared with the family) [1, p. 338].

With the advent of digital technologies, however, two new dimensions of privacy have emerged and proved pertinent. The first one is one's unfettered ability to *construct* one's digital identity, i.e. to present in the digital domain (e.g. digital records, social media) the type and amount of digital data about oneself that the person wishes and when the person wishes [4; 3, p. 8; 5, p. 218]. The second additional dimension is *the integrity* of one's digital identity. This means the absence of external interferences with the construction of one's identity whereby third parties craft or modify one's digital identity by, for instance, inferring and modeling one's features and behavior from data about others through predictive analytics [3,6,5,7] (the mechanics of which

will be explained in the next section). Thus, the concept of privacy has expanded to mean a state of unbound identity (self). Respectively, the right to privacy entails one's freedom to re/construct one's identity without external interventions [3, p. 5; 4, p. 367; 7, p. 477; 8].

III. AI-Enabled Predictive Analytics

Predictive analytics is a form of automated data processing that predicts possible outcomes based on diverse aggregate digital data [9,10]. To this end, bespoke software programs are used that deploy AI techniques, i.e. a combination of statistical methods (e.g. machine learning) and algorithms (i.e. mathematical processes) [11, p. 2]. The software processes vast amounts of (often anonymized) input data about human features (e.g. age, sex), conditions (e.g. illnesses) and/or behavior (e.g. health habits), and identifies statistically likely patterns and correlation in the data (e.g. between initial symptom and ultimate health outcome). When presented with new data about a particular individual, the software matches it against the established patterns/correlations, and infers – i.e. presumes – this individual's possible traits (e.g. socioeconomic status), condition (e.g. "at risk of pneumonia") or conduct (e.g. frequent smoker) [7, p. 477; 11, p. 676]. These presumptions (also called "statistical inferences" or "causal inferences") may or may not be entirely accurate and/or relevant [11, p. 4] Based on the presumptions, further estimations – e.g. classifications as "high-risk" or predictions about health outcomes – regarding the individual are made that drive decision-making about oneself (e.g. regarding health risk monitoring/management) [12, p. 867; 13].

The data used for predictive analytics in clinical diagnostics may include genetic and genomic, clinical, insurance claims, and socioeconomic information [9,13,14]. The data often concerns various aspects of patients' health, lifestyle and social persona, e.g. vital signs, mental health, socioeconomic status, "*marital and living status*" [9, p. 1124; 14, p. 60]. The data may

come from different sources, e.g. patients' electronic health records ("**EHR**"), insurance claims, social media application, etc. [9, p. 1124; 13; 14, p. 60]. As medical databases (e.g. EHR) are increasingly digitized, such data become also more accessible and widely used for predictive analytics [10].

IV. The Impact of AI-enabled Predictive Analytics on Individual Patient Privacy

AI-enabled predictive analytics is widely deployed in clinical care settings for healthcare monitoring, diagnostics and risk management. For example, predictive analytics is used to identify high-costs patients in order for their healthcare and associated costs to be managed more efficiently [9, p. 1124]. In this context, scholars have suggested that predictive analytics should be used to "*identify and address behavioral health problems*", such as depression, as well [9, p. 1124].

Healthcare providers also rely on predictive analytics for their high-risk care management programs [13]. Such a predictive analytics software forecast patients' future healthcare needs from historic data about medical expenditures incurred (e.g. insurance claims). As a result, patients with lower medical expenditures were considered to have lesser health needs and to be at lower health risks. Subsequent technical audits, however, revealed that these predictions misrepresented the actual health status of some groups of patients (e.g. Black) because the predictions did not reflect these patients' limited access to healthcare and more serious underlying conditions (e.g. chronic diseases often reported in the patients' EHR) [13, pp. 2-3].

Another predictive analytics software was used to forecast patients' probability of death from pneumonia so that high-risk patients could be admitted to and treated in hospital, and low-risk patients to be treated at home [15]. Initially, the software classified patients with asthma as low-

risk because it inferred no correlation between their underlying condition and instances of death. Historically, once such patients presented with a pulmonary infection, they were often directly admitted to hospital in order to manage the risk of complications and death. Hence, little or no instances of readmission appeared in the statistical data and no statistical correlation seemed to exist.

The applications of AI-enabled predictive analytics could undoubtedly benefit patients if predictive analytics could indeed help diagnose serious conditions earlier and coordinate healthcare better. Empirical research testify to such potential of AI to identify, for example, rare diseases based on data contained in large repositories of clinical data [16,17]. However, the use of predictive analytics in clinical diagnostics could have its privacy implications as well, concerning both individual patient privacy (i.e. the state) and the patient's rights thereto.

AI-enabled predictive analytics may in effect undermine the patient's *state* of privacy by constructing or modifying their health identity independent of the patient themselves [14, p. 65; 11; 16]. This happens by virtue of the very logic and mechanics of predictive analytics. Predictive analytics leverages generalized knowledge to infer and predict individual patient features, conditions and/or behavior. In the process, predictive analytics hypothesizes about and models them based on data about others, i.e. potentially modifies (somewhat) the patient's informational identity along others' common traits, states and conduct. Individual patient privacy could then be deemed violated because of this external intervention in the construction of the patient's informational health identity [5, p. 220; 7, p. 477]. Specifically, it is not the patient themselves but the predictive analytics software that re/crafts the patient's informational health identity. Moreover, predictive analytics could do so from not entirely relevant and/or accurate generalized benchmarks, which is essentially tantamount to "assigning" the patient an altered informational

identity for the intents and purposes for which predictive analytics is used [7, p. 479]. In the examples above, such an "assignment" of identity to the patient are effectively their risk scores and health needs profiles. Through them, the patients are classified into certain categories, which essentially defines them as more or less in need of healthcare. In this sense, the patient's identity is reduced to, and essentially substituted by, this qualification without the patient's active participation in the process.

AI-enabled predictive analytics may also undermine the patient's *right* to privacy when the patient has no control over or even a say in the use or output of predictive analytics. This might be the case when the patient is not aware of and has not consented to the use of predictive analytics, and/or when the patient cannot review and object to the predictions and classifications generated through predictive analytics. In the examples above, the use of predictive analytics to infer and predict patients' mental health problems would essentially result in a breach of their privacy rights if this happens without the patients' consent or procedural recourse.[1] Recent scholarship suggests that such practices are not insulated incidents. Cohen et al. comment that "*patients are generally unaware if their physicians are using computerized decision aids to guide treatment*." [17, p. 1143]. Bates et al. testify to patients' overall unwillingness to have their data linked and processed through predictive analytics [9, p. 1129]. Wachter and Mittelstadt elicit the challenges to exercising individual privacy rights to information and objection against predictive analytics [8]. The protection of individual privacy rights does not therefore appear watertight as far as the use of predictive analytics in clinical settings is concerned.

[1] For example, in the European Union, under Article 9 in conjunction with Article 22 of the European Union General Data Protection Regulation [19].

V. Recommendations for Mitigating the Privacy Impactions of AI-enabled Predictive Analytics

The impact of AI-enabled predictive analytics on individual privacy and the right thereto raises two broader questions. The first one is institutional and concerns the actual ability of the current generation of data protection laws to (re)institute privacy given the essence, mechanics, increasing use of and benefits from AI-enabled predictive analytics in socially significant domains, such as healthcare. As scholars have highlighted, AI challenges the very philosophy and regulatory design of current data protection regulations and potentially necessitates their re-think [21,22]. To what extent and in what way(s) then? While falling outside the scope of this paper, this question is worth highlighting as it indicates the significant interplay between technology and innovation, privacy, and policy- and law-making, and, ultimately, the role of the latter for advancing (any of) the former two aspects.

The second question is practical. It concerns the specific measures that healthcare institutions and providers should take in order to mitigate the impact of AI-enabled predictive analytics on patients' privacy and right thereto. As highlighted in previous sections, the question goes beyond merely preserving patients' unbound health identity but has also to do with how the digital (re)construction of that identity in fact enables or hinders patients to access, decide about and receive healthcare corresponding to their actual needs. In this sense, the privacy implications of predictive analytics cut to the core also of patients' autonomy and agency. Hence, any mitigation measures should seek to respect and restore them. The impact of AI-enabled predictive analytics on individual privacy and the right thereto could therefore be mitigated in two main ways.

The implications on privacy itself could be possibly dampened if the AI systems used for predictive analytics yield more accurate and statistically reliable inferences. This appears

possible with the help of new statistical models, methods and techniques. For example, high-performance generalized additive models -- novel machine learning methods -- have demonstrated better accuracy in healthcare settings than traditional machine learning techniques (e.g. decision tree, logistic regression or naive-Bayes methods) [15]. When taking part in the design or assessment of AI systems used for predictive analytics, bioethicists could specifically inquire about and explore with AI developers the possible deployment and the expected utility of such methods for enhancing accuracy and thus preserving individual patient privacy as unaltered health identity.

The implications of predictive analytics for individual privacy rights could potentially be resolved also through more patient-friendly privacy policies and practices of healthcare providers. They should inform patients of the use of predictive analytics and explicate its overall essence and general mode of operation, intended applications and purposes. Patients should specifically be made aware at least of: (i) the fact that inferences, classifications and predictions about them would be drawn from various generalized data, (ii) the likely or at least potential level of in/accuracy of these inferences, classifications and predictions with regards to the specific patient, and (iii) the likely or at least possible consequences of relying on them, e.g. inadequate or only proximate assessment of health risks, of possible diagnoses or treatment options. In cases when AI-enabled predictive analytics has proven to be reliable, e.g. diagnosing rare diseases, patients should be informed also of the potential of predictive analytics to identify or at least help hypothesize about viable diagnoses or treatments. Providing patients with this information would be a recognition and an act of respect for their ability and need to self-govern, i.e. for their autonomy. In combination, these measures are also a fundamental prerequisite to patients' informed decision-making and overall agency as the measures ensure essential details

needed by patients to start understanding and appreciating (the plausibility of) their diagnoses and treatment options, and forming an informed view about the latter.

Patients should then be given the opportunity to consent to or reject the use of predictive analytics for their diagnostics. If they choose to reject, the healthcare provider should still be able to provide the healthcare sought based on other diagnostics tools. If the patient consents to the use of predictive analytics, they should be presented with the substantively material classifications/predictions and given the opportunity to reject them or request their rectification in order to reflect the patients' actual health status. This third category of measures recognizes the need for and grants patients an actual and effective choice of how to exercise their right to privacy. Thus, the measures are also an act of respect for patients' autonomy and an enabler of their agency.

VI. Conclusion

Digital technologies may be a curse and a blessing. AI-enabled predictive analytics makes no exception. It may offer valuable insights into individual and population health patterns, trends and outcomes. Predictive analytics may also tangibly affect individual patient privacy and the right thereto, and thus ultimately undermine patient autonomy and agency. These implications could be mitigated if the limitations of predictive analytics are recognized and addressed through privacy-conscious institutional policies and practices, and through the incorporation in the AI systems of accuracy-enhancing statistical models and methods. Clinical ethicists and bioethicists overall could support this venture by raising awareness regarding the privacy implications of predictive analytics, helping assess the ethical compatibility of predictive analytics solutions, and crafting ethically sound privacy policies and protections.

References:

[1] Beauchamp TL, Childress JF. Principles of Biomedical Ethics. 8th edition. New York: Oxford University Press; 2019.

[2] Warren SD, Brandeis LD. The Right to Privacy. Harvard Law Review 1890;4:193–220. https://doi.org/10.2307/1321160.

[3] Wachter S. Privacy: Primus Inter Pares Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights. SSRN Journal 2017. https://doi.org/10.2139/ssrn.2903514.

[4] Hildebrandt M. Balance or Trade-off? Online Security Technologies and Fundamental Rights. Philos Technol 2013;26:357–79. https://doi.org/10.1007/s13347-013-0104-0.

[5] Loi M, Christen M. Two Concepts of Group Privacy. Philos Technol 2020;33:207–24. https://doi.org/10.1007/s13347-019-00351-0.

[6] Floridi L. The Informational Nature of Personal Identity. Minds & Machines 2011;21:549–66. https://doi.org/10.1007/s11023-011-9259-6.

[7] Mittelstadt B. From Individual to Group Privacy in Big Data Analytics. Philos Technol 2017;30:475–94. https://doi.org/10.1007/s13347-017-0253-7.

[8] Wachter S, Mittelstadt B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. LawArXiv; 2018. https://doi.org/10.31228/osf.io/mu2kf.

[9] Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big Data In Health Care: Using Analytics To Identify And Manage High-Risk And High-Cost Patients. Health Affairs 2014;33:1123–31. https://doi.org/10.1377/hlthaff.2014.0041.

[10] Escobar GJ, Turk BJ, Ragins A, Ha J, Hoberman B, LeVine SM, et al. Piloting electronic medical record–based early detection of inpatient deterioration in community hospitals. Journal of Hospital Medicine 2016;11. https://doi.org/10.1002/jhm.2652.

[11] Mühlhoff R. Predictive privacy: Collective data protection in the context of artificial intelligence and big data. Big Data & Society 2023;10:205395172311668. https://doi.org/10.1177/20539517231166886.

[12] Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring Fairness in Machine Learning to Advance Health Equity. Ann Intern Med 2018;169:866. https://doi.org/10.7326/M18-1990.

[13] Obermeyer Z, Mullainathan S. Dissecting Racial Bias in an Algorithm that Guides Health Decisions for 70 Million People. Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta GA USA: ACM; 2019, p. 89–89. https://doi.org/10.1145/3287560.3287593.

[14] Voigt K. Social Justice, Equality and Primary Care: (How) Can 'Big Data' Help? Philos Technol 2019;32:57–68. https://doi.org/10.1007/s13347-017-0270-6.

[15] Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney NSW Australia: ACM; 2015, p. 1721–30. https://doi.org/10.1145/2783258.2788613.

[16] Chen C, Lu MY, Williamson DFK, Chen TY, Schaumberg AJ, Mahmood F. Fast and scalable search of whole-slide images via self-supervised deep learning. Nat Biomed Eng 2022;6:1420–34. https://doi.org/10.1038/s41551-022-00929-8.

[17] Wojtara M, Rana E, Rahman T, Khanna P, Singh H. Artificial intelligence in rare disease diagnosis and treatment. Clin Transl Sci 2023;16:2106–11. https://doi.org/10.1111/cts.13619.

[18] Mühlhoff R. Predictive privacy: towards an applied ethics of data analytics. Ethics Inf Technol 2021;23:675–90. https://doi.org/10.1007/s10676-021-09606-x.

[19] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). vol. 119. 2016.
[20] Cohen IG, Amarasingham R, Shah A, Xie B, Lo B. The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care. Health Affairs 2014;33:1139–47. https://doi.org/10.1377/hlthaff.2014.0048.
[21] Mantelero A. Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI. vol. 36. The Hague: T.M.C. Asser Press; 2022. https://doi.org/10.1007/978-94-6265-531-7.
[22] Solove DJ. Artificial Intelligence and Privacy 2024. https://doi.org/10.2139/ssrn.4713111.