1

Artificial Intelligence and Public Health

Opportunities Abound

Sheldon H. Jacobson and Janet A. Jokela

Artificial intelligence (AI) has become the poster child for data science and advanced analytics. Pick up any newspaper or magazine and one finds the promise that AI can bring to solve society's most vexing problems. AI is a critical driver in the next economic revolution as part of the information age expansion, revealing hidden insights and unleashing the power of big data.

Medicine is an attractive target for AI (Maddox et al. 2019). Given the human decision-making element that drives medicine, the potential for AI to enhance diagnosis and treatment protocols is immense. Indeed, any healthcare decision for which there is a plethora of data and which requires human intervention and judgment is a ripe target to benefit from AI.

What exactly constitutes AI? The term "AI" has actually existed for more than 60 years, dating back to 1956 (Bringsjord and Govindarajulu 2018). Since the introduction of the modern computer, researchers have asked whether computers can be designed and programmed to mimic human intelligence, decision making, and reasoning. The classic science fiction movie *2001: A Space Odyssey* portrays HAL as a computer with human decision-making capabilities. Science fiction as such has been inspired by the AI vision, years prior to the actual realization of AI capabilities. More recently, the AlphaZero artificial intelligence program achieved a level of "superhuman performance" in the games of Go, shogi, and chess by reinforcement learning from self-play, an example of the potential power of deep learning (Silver et al. 2018).

At its core, AI is the process by which human decision making and reasoning can be achieved by algorithms programmed on a computer. The advantage of using a computer for such applications is the speed by which they can be executed and the data that can be stored and accessed by such algorithms. Specific types of AI include computer vision (i.e., interpreting images, such as how human sight functions), optimization (i.e., decision making associated with complicated or complex man-made or natural

3

systems), and learning (i.e., using data as input to improve decision making). Several high-profile applications of AI have drawn widespread attention. These include autonomous vehicles (land, water, and air, including drones), high-speed trading of financial products, and consumer marketing. Machine learning is a specific type of AI, whereby algorithms use data to train neural network models so that when such models take in real-time data, they are able to make decisions and use such data to further enhance their learning phase. Deep learning is a more advanced form of machine learning that uses layers of neural networks and massive data sets to achieve a similar objective.

Public health is a ripe area for AI to have an impact, especially in the era of COVID-19 (McCall 2020). AI possesses the potential to provide valuable pandemic-related insights and guidance. For example, efficient and appropriate healthcare system operations to ensure financial stability and enable such systems to fulfill their mission are critical. Smart strategies to estimate and mitigate risk are necessary. The high disparity rates in COVID-19 morbidity and mortality associated with race, ethnicity, and socioeconomic status expose pressing needs for investigation and sustainable solutions. Rising rates of depression, suicide, and mental health concerns present opportunities to appropriately target and deploy innovative solutions, such as telemedicine interventions.

Public health is the science of protecting and improving the health of people and their communities. This work is achieved by promoting healthy lifestyles, researching disease and injury prevention, and detecting, preventing, and responding to infectious diseases (CDC Foundation 2019). At its core, it is concerned with the well-being of a population. The American Public Health Association (APHA) lists 33 targeted topics and issues that they consider within the rubric of public health (American Public Health Association 2019). The Centers for Disease Control and Prevention (CDC) Organizational Chart (Centers for Disease Control and Prevention 2019a, 2019b) partitions public health into one institute (Occupational Health and Safety) and four directorates (Public Health Service and Implementation Science, Public Health Science and Surveillance, Non-infectious Diseases, Infectious Diseases), providing a broader classification that includes the APHA topics and issues. The CDC Strategic Priorities are the following: (1) improve health security at home and around the world; (2) prevent the leading causes of illness, injury, disability, and death; and (3) strengthen public health and healthcare collaboration.

To meet these priorities, the CDC activities include detecting, responding to, and stopping new and emerging health threats; preventing injuries, illnesses, and premature deaths; and discovering new ways to protect and improve the

public's health through science and advanced technology. These activities involve tracking the causes and rates of death in a population, as well as surveillance of population health data. Examples of issues that are germane to such public health activities include monitoring infectious diseases (such as seasonal influenza and childhood diseases), obesity, smoking trends, death rates and causes, suicide rates, and drug use (both elicit and legal, including opioids), to name but a few. The emergence of the SARS-CoV2, the virus that causes COVID-19, has projected a bright light on public health data and practices. Surveillance testing and contract tracing, core public health approaches, have drawn attention both in the media and with the general public.

The common thread that permeates public health activities is data, and the tools used to glean information from such data are statistical methods. Given that numerous AI methods require large input data sets, the marriage of public health and AI appears to be a natural fit. Indeed, AI has already been applied to public health problems, though the name AI may not have been explicitly attached to such activities.

1.1 AI Successes in Public Health

Prior to the widespread emphasis on machine learning in AI, numerous data-driven studies were conducted that used optimization models to address public health issues. AI is often demonstrated through the design of an expert system or decision support framework for capturing and enhancing the decision-making process. Operations research, a field that uses models to improve the decision-making process, can be classified within the AI family of methodologies, though it has traditionally not been so formally described.

One public health area that attracted significant attention is broadly designated as pediatric immunization formulary design. Weniger et al. (1998) introduced an optimization model for designing pediatric vaccine formularies based on a variety of cost components, including the price and packaging for the vaccines, and how the vaccine antigens fit into the Recommended Childhood Immunization Schedule (RCIS). This model was then used as the driver in a decision support system for stocking pediatric formularies (Jacobson and Sewell 2008).

Given the complexity of the RCIS, children often fall behind in receiving their vaccines, creating population risk with decreased herd immunity. Engineer et al. (2009) present an optimization framework for designing catch-up schedules for pediatric childhood immunization. The framework is embedded into a

decision support system, making it easier for clinicians to assess which vaccines should be administered to a child at any point when they present at a clinic with an incomplete or unknown immunization history.

Population disease screening is another area in which decision support systems have been effectively implemented to enhance the decision-making process. Long and Brandeau (2009) provide an overview of operations research that can be effective in modeling and containing the spread of infectious diseases. Brandeau et al. (1991) outline an analytic decision model for setting intervention screening policies for the human immunodeficiency virus (HIV). Atkinson et al. (2007) outline a strategy for reducing the spread of dengue fever by proposing a model for capturing the growth of the underlying mosquito population. Models as such have important implications that impact the decision-making process for population health. Lee et al. (2006) provide a decision support system to help manage the operation of dispensing clinics in the event of biological threats and infectious disease outbreaks.

1.2 Recent Interest in AI in Public Health

The visibility of both AI and healthcare has been a natural driver for identifying areas of interest for AI in public health. Stead (2018) and Shah (2019) provide thoughtful editorials on the growth of AI in healthcare and its potential to improve clinical care. Hinton (2018) and Naylor (2018) make a similar case for deep learning as a vehicle to improve clinical healthcare practice. Maddox et al. (2019) pose several critical questions that provide indicators for when and how to use AI in healthcare. Beam and Kohane (2018) provide an overview of machine learning and the relationship between data size and the human-to-machine decision-making effort, providing a structured classification for numerous applications in health and industry.

The common thread that permeates such discussions is the proliferation of electronic health records, the vehicle to access enormous amounts of data as inputs for AI algorithms. However, thought-provoking papers as such are not new. Patel et al. (2009) discusses the coming wave of AI applications in medicine, though few recognized the potential for machine learning and deep learning across such a wide swath of healthcare.

Clearly, interest of AI in public health has begun to grow, with applications sprouting up addressing numerous public health issues. The Canadian Institute for Advanced Research (CIFAR) and the Canadian Institutes of Health Research's Institute of Population and Public Health (CIHR-IPPH) co-hosted a workshop in January 2018, "Application of Artificial Intelligence

Approaches to Tackle Public Health Challenges," to bring together public health leaders to summarize the state of AI in public health in Canada, and to identify opportunities where AI may influence public health processes and decision making. The outcome of the workshop provided a roadmap of opportunities for AI in public health in Canada, many of which are applicable in the United States and other countries. Given that the workshop organizers recognized the paucity of such activities to date, the workshop provided a much-needed impetus to begin dialogues between the various stakeholders, as well as initiate the process of introducing and energizing AI applications to address critical public health issues.

AI and epidemiology have a natural connection because of their reliance on data. Thiebault and Thiessard (2018) provide an overview of papers published in the area of public health and epidemiology informatics. The objective of their analysis was to identify the most significant papers published in 2017. Of particular note is that their initial scan using PubMed and Web of Science yielded 843 articles, indicating the substantial interest in the area. Of these papers, one of the two best used an artificial neural network to anonymize patient notes in electronic health records (Demoncourt et al. 2017).

AI has also been proposed to improve health in resource-poor settings (Wahl et al. 2018, Shah 2019). Given the challenges faced in such environments, they explore the types of AI that may be beneficial in such environments, and the types of problems that are most amenable for AI application within a resource-challenged setting.

1.3 AI Opportunities in Public Health

Public health analytics are driven by data. Therefore, statisticians and epidemiologists have found public health a rich domain for analysis. This very surfeit of data makes AI an attractive target for application. The following provides an overview of several public health issues and the opportunities for AI to bring new insights into their understanding and solution.

The COVID-19 pandemic, as a global public health crisis, has disrupted every aspect of society. Given the rapid response and decision making required to meet real-time challenges, AI can support such processes (McCall 2020). Areas that may benefit from AI during the COVID-19 response include the creation and reallocation of resources to respond to the demands of care (such as equipment, personnel, and personal protective equipment); the appropriate allocation of healthcare resources to balance COVID-19 and non-COVID-19 care, including preventive medicine; designing population-based testing

strategies that balance costs and societal benefits; reducing morbidity and mortality rates across all sociodemographic strata; and determining the optimal use of telemedicine.

Obesity has become a major worldwide public health challenge. Since the 1960s, the United States has seen a steady increase in the adult (20 years and older) obesity rate, reaching 39.8 percent in 2015–2016, according to the CDC (Hales et al. 2017). Obesity has also become more prevalent in other countries around the globe, making it a global public health epidemic. There is an abundance of data collected and available to root out the causes of obesity. These include patient medical information, societal and economic factors, food sources, and environmental issues. Extensive analysis of such data has provided insights into associations with obesity, yet little knowledge has been gleaned that can be put into practice to reverse population obesity trends. Clearly, the causes of obesity are complex and may not lend themselves to traditional statistical methods. This is where AI may be a useful target of opportunity (Zeevi et al. 2015). Given the abundance of data, coupled with the multiple association vectors that have been observed, machine learning models can be employed to identify hidden connections across multiple factors that may help identify potential drivers for obesity, as well as appropriate population and patient-centric interventions to reverse current obesity trends (Scheinker et al. 2019).

Smoking combustible cigarettes continues to be a public health concern. Numerous studies have demonstrated the detrimental impact of smoking on individual health, including higher rates of cardiovascular disease and numerous types of cancers. Secondhand smoke has also been associated with higher rates of lung cancer. There have been numerous policy changes implemented to limit locations where people are permitted to smoke. The footprint for acceptable smoking locations continues to shrink, with data analysis used to support such policies. Using noncombustible cigarettes, informally termed vaping, represents the new wave of smoking, creating opportunities for data analysis to assess the public health impact of such items (Visweswaran et al. 2020). However, the health problems associated with combustible cigarettes differ from those of non-combustible cigarettes. As such, AI may provide a useful tool to glean public health impacts for such cigarettes and provide insights that can be used to guide policies on their use.

Mass killings have been a persistent plague in society in the United States for the past several decades. Although the rate of mass killings (defined by the FBI as four of more deaths in a single incident) has remained steady for more than a decade (King and Jacobson 2017), the scourge of fear in society due to such events has become a public health challenge. After each such incident,

data are collected and analyzed to create a profile of the perpetrators, yet each incident continues to baffle law enforcement officials and policy makers. With over three-quarters of mass killings involving some form of firearm, calls for greater restrictions on gun accessibility and stronger background checks have grown louder but have also been met with resistance by gun advocates who argue for the deterrence value of firearms. Given the visceral reactions on both sides of this issue, AI may provide an avenue to adjudicate this debate, provide a better understanding of unforeseen patterns hidden within the plethora of data, and glean insights into how such events can be prevented or deterred.

Alcoholism and drug addiction (illicit and prescription, including opioids) have been persistent and pernicious public health challenges for several decades. The annual cost to society as a whole has been measured to be over US$500 billion in the United States alone (NIDA 2017). Significant progress has been made to address these issues, including both professional and self-help organizations. More recently the legalization of marijuana has drawn attention and concerns that such a policy will have deleterious effects on society. For these issues, there is a plethora of data available on factors that lead to alcoholism and drug addiction, yet there are no easy answers in eliminating these scourges, or, more realistically, in halting the progression earlier in their destructive cycle. AI may provide an avenue to identify driving factors that could be addressed using traditional treatment plans, or to help to identify new methods that have yet to be deployed, but may be more effective in reducing the negative impact of such diseases on individuals and society as a whole (Mak et al. 2019).

The societal threat of bioterrorism attacks presents major public health challenges. In the event of the intentional release of an airborne pathogen, appropriate preparation and rapid detection and response are the three factors that require attention to minimize the impact during such events. Given the breadth of possible pathogens, and their public health impact, large amounts of data exist to understand what can be done to mitigate their impacts on populations across multiple scenarios (Zaric et al. 2008). Given the extreme nature of such events, traditional methods of response may be challenged because of weaknesses in supply chains and transportation networks. AI may be effective in parsing through such data and providing insights on how public health may optimally prepare for and respond to such events.

1.4 AI Challenges in Public Health

Many of the challenges to using AI in public health apply across other fields. For example, developing domain-specific AI tools often requires a deep

knowledge of the domain and a thorough understanding of AI methods, or, more realistically, a collaborative and interdisciplinary environment whereby such individuals can work together. Moreover, given the intense competition for AI skills across numerous fields, public health will need to compete for such talent in such a highly competitive human-capital environment. Such competition in general has driven up the cost of AI talent, possibly taking it out of the reach of many public health settings. As AI becomes even more attractive across a wider swath of societal applications, such competition will only become more intense in the coming years. It is also possible, however, that the COVID-19 pandemic will inspire altruistic AI talent to pursue public health opportunities to tackle this most vexing problem.

Given the need for data in AI algorithms, the issues of privacy become of paramount concern. Electronic health records facilitate the availability of data. However, the Health Insurance Portability and Accountability Act (HIPAA) privacy rule places significant barriers in using such data without prior patient consent. Although HIPAA gives patients control over their health records, it creates a headwind for analysts to use such data (even when anonymized) to drive AI algorithms. A dialogue on such issues with electronic health record designers like EPIC and Cerner is paramount to overcoming such challenges. Possible solutions include requesting patients to release their anonymized data or redesigning electronic health records to allow personal information to be hidden while making anonymized health information accessible for research and analysis. Indeed, AI may be the driving force to revamp HIPAA laws to facilitate such advances.

Along these same lines, AI requires data that may contain inherent biases. Such biases would percolate through any algorithm, resulting in biased decisions. This is particularly critical in issues pertaining to human health, and public health in particular, where decisions can impact the lives and well-being of populations (Obermeyer et al. 2019). Aswani and Olfat (2018) discuss the issue of fairness in AI, and how inherent biases in decision making can be detected and mitigated. Parallel to such issues, Israni and Verghese (2019) discuss not only whether AI can improve healthcare delivery but also whether it can deliver more humanistic medicine. Indeed, the human–computer nexus will be increasingly scrutinized as AI works its way into healthcare systems.

1.5 Summary and Conclusions

Public health problems offer a panacea of opportunities for AI. The enormous amounts of data that are available make it well suited for this domain of

application. At the same time, as with any healthcare data analysis, issues of privacy must be addressed to ensure that individuals within a population are appropriately protected. Clearly, the benefits outweigh the risks in using AI for surveillance, assessment, and policy formulation within public health. The greatest challenge may be the will to advance AI and the fortitude to overcome existing barriers so that AI can provide the greatest benefits in public health.

References

American Public Health Association, 2019, Topics and issues. www.apha.org/topics-and-issues.

Aswani, A, Olfat, M., 2018, Designing algorithms to increase fairness in artificial intelligence, *SIAM News*, **51**(10), 1,3.

Atkinson, MP, Su, Z., Alphey, N., et al. 2007, Analyzing the control of mosquito-borne diseases by a dominant lethal genetic system, *Proceedings of the National Academy of Sciences*, **104**(22), 9540–9545.

Beam, A.L., Kohane, A.S., 2018, Big data and machine learning in health care, *JAMA*, **319**(13), 1317–1318

Brandeau, M.L., Lee, H.L., Owens, D.K., Sox, C.H., Wachter, R.M, 1991, A policy model of human immunodeficiency virus screening and intervention, *Interfaces*, **21**(3), 5–25.

Bringsjord, S., Govindarajulu, N.S., 2018, Artificial intelligence. In: *The Stanford Encyclopedia of Philosophy* (Fall 2018 edition), E.N. Zalta (ed.). https://plato.stanford.edu/archives/fall2018/entries/artificial-intelligence.

CDC Foundation, 2019, Public health in action. www.cdcfoundation.org/what-public-health.

Centers for Disease Control and Prevention, 2019a, Organizational chart. www.cdc.gov/about/pdf/organization/cdc-org-chart.pdf.

Centers for Disease Control and Prevention Foundation, 2019b, Who we are/what is public health. www.cdcfoundation.org/what-public-health

Demoncourt, F. Lee, J.Y., Uzuner, O., Szolovits, P., 2017, De-identification of patient notes with recurrent neural networks, *Journal of the American Medical Informatics Association*, **24**, 596–606.

Engineer, F.G., Keskinocak, P., Pickering, L.K., 2009, OR practice – catch-up scheduling for childhood vaccination, *Operations Research*, **57**(6), 1307–1319.

Hales, C.M., Carroll, M.D., Fryar, C.D., Ogden, C.L., 2017, Prevalence of obesity among adults and youth: United States, 2015–2016, NCHS Data Brief, No. 288, October 2017. www.cdc.gov/nchs/data/databriefs/db288.pdf.

Hinton, G., 2018, Deep learning – a technology with the potential to transform health care, *JAMA*, **320**(11), 1101–1102.

Israni S.T., Verghese, A., 2019, Humanizing artificial intelligence, *JAMA*, **321**(1), 29–30

Jacobson, S.H., Sewell, E.C., 2008, A web-based tool for designing vaccine formularies for childhood immunization in the United States, *JAMA*, **15**(5), 611–619.

King, D.M., Jacobson, S.H., 2017, Random acts of violence: examining probabilistic independence of mass killing events in the United States, *Victims and Violence*, **32**(6), 1014–1023.

Lee, E.K., Maheshwary, Mason, J., Glisson, S., 2006, Large-scale dispensing for emergency response to bioterrorism and infectious-disease outbreak, *Interfaces*, **36**(6), 591–607.

Long, E.F., Brandeau, M.L., 2009, OR's next top model: decision models for infectious disease control, *TutORials in Operations Research*, 123–138.

Maddox T.M., Rumsfeld J.S., Payne P.R.O., 2019, Questions for artificial intelligence in health care, *JAMA,* **321**(1), 31–32.

Mak, K.K., Lee, K., Park, C., 2019, Applications of machine learning in addiction studies: a systematic review, *Psychiatry Research*, **275**, 53–60.

McCall, B., 2020, COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread, *The Lancet Digital Health*, **2**(4), e166–e167.

Naylor, C.D., 2018, On the prospects for a (deep) learning health care system, *JAMA*, **320**(11), 1099–1100.

NIDA, 2017, Trends & statistics. www.drugabuse.gov/related-topics/trends-statistics.

Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S., 2019, Dissecting racial bias in an algorithm used to manage the health of populations, *Science*, **366**(6464), 447–453.

Patel, V.L., Shortliffe, E.H., Stefanelli, M., et al., 2009, The coming of age of artificial intelligence in medicine, *Artificial Intelligence in Medicine*, **46**(1), 5–17.

Scheinker, D., Valencia, A., Rodriguez, F., 2019, Identification of factors associated with variation in us county-level obesity prevalence rates using epidemiologic vs machine learning models, *JAMA Network Open*, **2**(4), e192884–e192884.

Shah NR., 2019, Healthcare in 2030: will artificial intelligence replace physicians? *Annals Internal Medicine*, **170**, 407–408. doi:10.7326/M19-0344

Silver, D.S., Hubert, T., Schrittwieser, J., et al., 2018, A general reinforcement learning algorithm that masters chess, shogi and go through self play, *Science* **362**, 1140–1144.

Stead W.W., 2018, Clinical implications and challenges of artificial intelligence and deep learning, *JAMA*, **320**(11), 1107–1108.

Thiebout, R., Thiessard, F., 2018, Artificial intelligence in public health and epidemiology, *IMIA Yearbook of Medical Informatics*, **27**(1):207–209.

Visweswaran, S., Colditz, J.B., O'Halloran, P., et al., 2020, Machine learning classifiers for twitter surveillance of vaping: comparative machine learning study, *Journal of Medical Internet Research*, **22**(8), e17478.

Wahl, B., Cossy-Gantner, A., Germann, S., Schwalbe, N.R., 2018, Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Global Health*, **3**, 1–7.

Weniger, B.G., Chen, R.T., Jacobson, S.H., et al., 1998, Addressing the challenges to immunization practice with an economic algorithm for vaccine selection, *Vaccine*, **16**(19), 1885–1897.

Zaric, G.S., Bravata, D.M., Cleophas Holty, J.E., et al., 2008, Modeling the logistics of response to anthrax bioterrorism, *Medical Decision Making*, **28**(3), 332–350.

Zeevi, D., Korem, T., Zmora, N., et al., 2015, Personalized nutrition by prediction of glycemic responses, *Cell*, **163**(5), 1079–1094.