ry agencies will pay particular attention to preclinical safety and toxicity studies and assessments of unexpected adverse events during clinical trials and after licensure. The case for licensure may be established through traditional clinical efficacy trials, but declining case counts or an urgent need for intervention may necessitate a different pathway. Alternatives include using efficacy data from studies in animals combined with human immunogenicity data or bridging to an as-yet-undefined immune correlate of protection. Human challenge studies have been proposed in order to augment information from efficacy trials, assist in exploring immune correlates of protection, or generate efficacy data if natural transmission substantially declines. In the absence of a clear understanding of the frequency of adverse neurologic outcomes or the persistence of ZIKV in biologic fluids, however, human ZIKV challenge is ethically complex.

Other flavivirus vaccines have been licensed, including those against yellow fever (live attenuated), Japanese encephalitis (inac-

tivated, live chimeric, live attenuated), tickborne encephalitis (inactivated), and dengue (live chimeric). Some have validated surrogates of protection, and all are based on neutralizing antibody. A neutralizing antibody titer of 1 in 10 is the surrogate of protection for the Japanese and tickborne encephalitis vaccines; for yellow fever, the titer is between 1 in 10 and 1 in 50. Preclinical ZIKV studies suggest that a titer of 1 in 10 for mice and approximately 1 in 100 for nonhuman primates protected against ZIKV challenge.[1,2] If these figures translate to humans, developing a ZIKV vaccine is very feasible.

The time required to develop a safe, efficacious ZIKV vaccine will be determined by prior experience with the selected technology, the continuation of outbreaks, and the required scale-up of manufacturing. Ultimately, developing, licensing, and deploying a vaccine capable of affecting the current epidemic will require seamless coordination among developers, regulatory agencies, the WHO, and national health authorities, along with a robust monetary commit-

ment from governments and funding agencies.

The views expressed in this article do not necessarily represent those of the U.S. Army or the Department of Defense.

Disclosure forms provided by the authors are available at NEJM.org.

From the Walter Reed Army Institute of Research, Silver Spring, MD (S.J.T.); Sanofi Pasteur, Lyon, France (M.L., N.A.C.J.); and the Sealy Center for Vaccine Development and Department of Pathology, University of Texas Medical Branch, Galveston (A.D.T.B.).

1. Larocca RA, Abbink P, Peron JP, et al. Vaccine protection against Zika virus from Brazil. Nature 2016;536:474-8.
2. Abbink P, Larocca RA, De La Barrera RA, et al. Protective efficacy of multiple vaccine platforms against Zika virus challenge in rhesus monkeys. Science 2016 August 4 (Epub ahead of print).
3. Nicastri E, Castilletti C, Liuzzi G, Iannetta M, Capobianchi MR, Ippolito G. Persistent detection of Zika virus RNA in semen for six months after symptom onset in a traveller returning from Haiti to Italy, February 2016. Euro Surveill 2016;21932).
4. Johansson MA, Mier-y-Teran-Romero L, Reefhuis J, Gilboa SM, Hills SL. Zika and the risk of microcephaly. N Engl J Med 2016;375: 1-4.
5. Cao-Lormeau VM, Blake A, Mons S, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. Lancet 2016; 387:1531-9.

Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

Ziad Obermeyer, M.D., and Ezekiel J. Emanuel, M.D., Ph.D.

B y now, it's almost old news: big data will transform medicine. It's essential to remember, however, that data by themselves are useless. To be useful, data must be analyzed, interpreted, and acted on. Thus, it is algorithms —

not data sets — that will prove transformative. We believe, therefore, that attention has to shift to new statistical tools from the field of machine learning that will be critical for anyone practicing medicine in the 21st century.

First, it's important to understand what machine learning is not. Most computer-based algorithms in medicine are "expert systems" — rule sets encoding knowledge on a given topic, which are applied to draw conclusions

about specific clinical scenarios, such as detecting drug interactions or judging the appropriateness of obtaining imaging. Expert systems work the way an ideal medical student would: they take general principles about medicine and apply them to new patients.

Machine learning, conversely, approaches problems as a doctor progressing through residency might: by learning rules from data. Starting with patient-level observations, algorithms sift through vast numbers of variables, looking for combinations that reliably predict outcomes. In one sense, this process is similar to that of traditional regression models: there are outcomes, covariates, and statistical functions linking the two. But where machine learning shines is in handling enormous numbers of predictors — sometimes, remarkably, more predictors than observations — and combining them in nonlinear and highly interactive ways.[1] This capacity allows us to use new kinds of data, whose sheer volume or complexity would previously have made analyzing them unimaginable.

Consider a chest radiograph. Some radiographic features might predict an important outcome, such as death. In a standard statistical model, we might use the radiograph's interpretation — "normal," "atelectasis," "effusion" — as a variable. But instead, why not let the data speak for themselves? Leveraging dramatic advances in computational power, digital pixel matrixes underlying radiographs become millions of individual variables. Algorithms then go to work, clustering pixels into lines and shapes and ultimately learning contours of frac-

ture lines, parenchymal opacities, and more. Even traditional insurance claims data can take on a new life: diagnostic codes trace an intricate, dynamic picture of patients' medical histories, far richer than the static variables for coex-

isting conditions used in standard statistical models.

Of course, letting the data speak for themselves can be problematic. Algorithms might "overfit" predictions to spurious correlations in the data, or multiple collinear, correlated predictors could produce unstable estimates. Either possibility can lead to overly optimistic estimates of the accuracy of a model and exaggerated claims about real-world performance. These concerns are serious and must be addressed by testing models on truly independent validation data sets, from different populations or periods that played no role in model development. In this way, problems in the model-fitting stage, whatever their cause, will show up as poor performance in the validation stage. This principle is so important that in many data-science competitions, validation data are released only after teams upload their final algorithms built on another publicly available data set.

Another key issue is the quantity and quality of input data.

Machine learning algorithms are highly data hungry, often requiring millions of observations to reach acceptable performance levels.[2] In addition, biases in data collection can substantially affect both performance and gen-

Machine learning has become ubiquitous and indispensable for solving complex problems in most sciences. The same methods will open up vast new possibilities in medicine.

eralizability. Lactate might be a good predictor of the risk of death, for example, but only a small, nonrepresentative sample of patients have their lactate levels checked. Private companies spend enormous resources to amass high-quality, unbiased data to feed their algorithms, and existing data in electronic health records (EHRs) or claims databases need careful curation and processing to become usable.

Finally, machine learning does not solve any of the fundamental problems of causal inference in observational data sets. Algorithms may be good at predicting outcomes, but predictors are not causes.[3] The usual common-sense caveats about confusing correlation with causation apply; indeed, they become even more important as researchers begin including millions of variables in statistical models.

Machine learning has become ubiquitous and indispensable for solving complex problems in most sciences. In astronomy, algorithms sift through millions of

images from telescope surveys to classify galaxies and find supernovas. In biomedicine, machine learning can predict protein structure and function from genetic sequences and discern optimal diets from patients' clinical and microbiome profiles. The same methods will open up vast new possibilities in medicine. A striking example: algorithms can read cortical activity directly from the brain, transmitting signals from a paralyzed human's motor cortex to hand muscles and restoring motor control.[4] These advances would have been unimaginable without machine learning to process real-time, high-resolution physiological data.

Increasingly, the ability to transform data into knowledge will disrupt at least three areas of medicine. First, machine learning will dramatically improve the ability of health professionals to establish a prognosis. Current prognostic models (e.g., the Acute Physiology and Chronic Health Evaluation [APACHE] score and the Sequential Organ Failure Assessment [SOFA] score) are restricted to only a handful of variables, because humans must enter and tally the scores. But data could instead be drawn directly from EHRs or claims databases, allowing models to use thousands of rich predictor variables. Does doing so lead to better predictions? Early evidence from our own ongoing work, using machine learning to predict death in patients with metastatic cancer, provides some indication: we can precisely identify large patient subgroups with mortality rates approaching 100% and others with rates as low as 10%. Predictions are driven by

fine-grained information cutting across multiple organ systems: infections, uncontrolled symptoms, wheelchair use, and more. Better estimates could transform advance care planning for patients with serious illnesses, who face many agonizing decisions that depend on duration of survival. We predict that prognostic algorithms will come into use in the next 5 years — although prospective validation will take several more years of data collection.

Second, machine learning will displace much of the work of radiologists and anatomical pathologists. These physicians focus largely on interpreting digitized images, which can easily be fed directly to algorithms instead. Massive imaging data sets, combined with recent advances in computer vision, will drive rapid improvements in performance, and machine accuracy will soon exceed that of humans. Indeed, radiology is already partway there: algorithms can replace a second radiologist reading mammograms[5] and will soon exceed human accuracy. The patient-safety movement will increasingly advocate the use of algorithms over humans — after all, algorithms need no sleep, and their vigilance is the same at 2 a.m. as at 9 a.m. Algorithms will also monitor and interpret streaming physiological data, replacing aspects of anesthesiology and critical care. The time scale for these disruptions is years, not decades.

Third, machine learning will improve diagnostic accuracy. A recent Institute of Medicine report highlighted the alarming frequency of diagnostic errors and the lack of interventions to reduce

them. Algorithms will soon generate differential diagnoses, suggest high-value tests, and reduce overuse of testing. This disruption will happen more slowly, over the next decade, for three reasons: first, the standard for diagnosis is unclear in many conditions (e.g., sepsis, rheumatoid arthritis) — unlike binary judgments in radiology or pathology (e.g., malignant or benign) — making it harder to train algorithms. Second, high-value EHR data are often stored in unstructured formats that are inaccessible to algorithms without layers of preprocessing. Finally, models need to be built and validated individually for each diagnosis.

Clinical medicine has always required doctors to handle enormous amounts of data, from macro-level physiology and behavior to laboratory and imaging studies and, increasingly, "omic" data. The ability to manage this complexity has always set good doctors apart from the rest. Machine learning will become an indispensable tool for clinicians seeking to truly understand their patients. As patients' conditions and medical technologies become more complex, the role of machine learning will grow, and clinical medicine will be challenged to grow with it. As in other industries, this challenge will create winners and losers in medicine. But we are optimistic that patients, whose lives and medical histories shape the algorithms, will emerge as the biggest winners as machine learning transforms clinical medicine.

From the Department of Emergency Medicine, Harvard Medical School and Brigham and Women's Hospital, and the Department of Health Care Policy, Harvard Medical School, Boston (Z.O.); and the Department of Medical Ethics and Health Policy, Perelman School of Medicine, and the Department of Health Care Management, Wharton School, University of Pennsylvania, Philadelphia (E.J.E.).

1. Mullainathan S, Spiess J. Machine learning: an applied econometric approach. J Econ Perspect (in press).

2. Halevy A, Norvig P, Pereira F. The unreasonable effectiveness of data. IEEE Intell Syst 2009;24(2):8-12.

3. Kleinberg J, Ludwig J, Mullainathan S, Obermeyer Z. Prediction policy problems. Am Econ Rev 2015;105:491-5.

4. Bouton CE, Shaikhouni A, Annetta NV, et al. Restoring cortical control of functional movement in a human with quadriplegia. Nature 2016;533:247-50.

5. Gilbert FJ, Astley SM, Gillan MGC, et al. Single reading with computer-aided detection for screening mammography. N Engl J Med 2008;359:1675-84.