In Reply to Prashar and to Savage: I thank Prashar and Savage for their interest in my commentary, and I agree with their thoughtful comments. They both discussed the “black box” nature of artificial intelligence (AI) and deep learning (DL) in medicine. By “black box,” it is meant that such models may be highly sophisticated, with a massive quantity of parameters. Given this model complexity, interpretability of what aspects of the data are being used to drive predictions is difficult to confidently assess.1

The issue of interpretability connects to the data used for learning the model parameters, which goes beyond the complexity of the model itself. Specifically, there is a need to ensure that the training data do not have biases that will be transferred to the model (e.g., inadequacies of data from certain subgroups of people). It is essential to ensure that a sufficient quantity of data is available to capture the complete heterogeneity across all patients for whom these models are applied. A challenge on this front concerns what is meant by a “sufficient quantity,” as traditional “power” calculations for statistical models are unlikely to apply to complex DL models.

Attention must also be placed on validation. Processes for carefully validating AI technology in medicine are essential and difficult, as often validation must be based on observational data2 (randomized trials may not always be ethical or practical). Such validation is essential even when the “black box” issue is at least partially removed. Validation of AI technology has many of the same issues associated with the data used for model learning (e.g., the need to assure that the validation data are representative of all patients for whom these technologies will be deployed).

Like all technologies introduced into medicine, care must be placed on fundamental understanding of how AI makes predictions (interpretability) and on validation. Because AI is such a data-driven framework, careful attention to the data used for learning and validation must be applied, and the need for this care must be embedded into medical education.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Lawrence Carin, PhD
Provost, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; lawrence.carin@kaust.edu.sa; ORCID: https://orcid.org/0000-0001-6277-7948.

References

Recommendations for Integrating the Fundamentals of Machine Learning into Medical Curricula

To the Editor: We agree with James and colleagues that principles of machine learning (ML) need to be integrated into medical curricula.1 As noted, these concepts may be readily incorporated into evidence-based medicine and doctoring courses. However, we would like to emphasize the importance of presenting an introduction to ML in both large and small groups. Large lectures would create a foundation for students, regardless of prior experience, but would allow for discussion of the nuances of real-world usage and allow case-based study of ML.

In this vein, although ML may allow physicians to provide better care for their patients, it is imperative that ML is not taught as absolute truth. ML tools are constantly changing programs that derive their starting point from the programmers and evolve based on user data. As with any tool crafted by human beings, ML is susceptible to error and bias. Physicians and trainees should be taught to continue to use their clinical judgment while using ML applications and to dynamically improve tools through feedback. A diversity of voices, backgrounds, and data in ML development can help broaden the reach of ML as well as potentially help combat systematic racial and social biases that permeate even the most objective of data-based tools.2

Strong relationships between coders and physicians will allow for more efficient adoption and usage of ML tools, so we suggest that medical professionals and trainees learn directly from the experts who pioneer this technology.3 Medical schools are increasingly including interdisciplinary learning experiences in curricula, with the consensus that knowledge is best gleaned from those who practice it. This interaction would have the added benefit of allowing developers to learn from physicians and students to create more effective tools for future patient care. It would also award students experience in communicating goals and specifications with engineers, abilities they will draw upon as future physicians using ML.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Newton V. Nagirimadugu
Third-year medical student, George Washington University School of Medicine and Health Sciences, Washington, DC; nvnagirimadugu@gwu.edu; ORCID: https://orcid.org/0000-0001-8396-2001.

Sudhamsh Tippireddy
Third-year medical student, University of Chicago Pritzker School of Medicine, Chicago, Illinois.

References