Machine learning and the future of Medicare fraud detection

To the Editor: Increasing attention has been paid in recent years to the problem of fraud and abuse in the US health care system, which is estimated by the US Federal Bureau of Investigation to account for 3% to 10% of overall spending.[1] As the largest payer in the current system via Medicare, the US government has a strong interest in combating fraud. To address this problem, the Centers for Medicare & Medicaid Services has begun adopting increasingly sophisticated machine learning methods as part of its Fraud Prevention System to proactively identify perpetrators.[2] We highlight some of the methods being used and how they may affect future practice.

Machine learning applies statistical algorithms to data to uncover hidden patterns and make predictions about the future. These tools can be broadly characterized as either supervised or unsupervised. Supervised methods use well-understood and labeled data sets to create models that can be applied prospectively.[3] In the context of fraud detection, a supervised algorithm could use the characteristics of known instances of fraud to label new claims as legitimate or possibly fraudulent.[2] Unsupervised methods, on the other hand, are used to explore the structure of unlabeled data.[3] When considering health care claims data, investigators are faced with the challenge that, because not all cases of fraud are identified, the labeling applied to historical claims data cannot be considered to be completely accurate. Nevertheless, unsupervised methods such as clustering empower investigators to uncover and identify suspicious patterns in submitted claims that warrant further investigation.[2] Previously, such investigations relied on the use of descriptive statistics to identify claims patterns that were several standard deviations above average in terms of the number of services performed or charges submitted.[1,4] Advanced clustering methods can uncover far more subtle deviations from normal practice patterns that would previously have flown under the radar.

It is the latter of the 2, the unsupervised methods, that may prove to have the largest impact on future practice. Outlier detection methods work by finding normative patterns in claims data and flagging instances that are deviant. For example, providers that bill codes that are outside of their normal scope of practice may end up in an auditor's crosshairs if such usage is not seen in a large number of their peers. Another type of practice that may get flagged involves the use of codes that vary based on the number of services performed within a single visit, such as the codes for destruction of premalignant lesions (17000/17003/17004). If a provider bills only code 17004 during a given year without ever billing 17000 or 17003, this behavior may be flagged as anomalous in comparison with the provider's peers. Even if this specific instance is not an audit trigger, such deviations from standard practice may end up affecting the provider's rating with the Centers for Medicare & Medicaid Services, making future payments subject to increased scrutiny.[2] It is therefore important to properly document any practice that requires deviations from practice norms whenever possible. In the future of Medicare fraud investigations, standing out from the crowd may end up putting a target on your back.

Dan Gordon, BS, and Daniel M. Siegel, MD, MS

From the Department of Dermatology, State University of New York Downstate Health Sciences University, Brooklyn

Funding sources: None.

Conflicts of interest: None disclosed.

IRB approval status: Not applicable.

Reprints not available from the authors.

Correspondence to: Dan Gordon, BS, SUNY Downstate Medical College, Department of Dermatology, 450 Clarkson Ave, Brooklyn, NY 11203

E-mail: dan.gordon@downstate.edu

REFERENCES

1. Ekin T. *Statistics and Health Care Fraud: How to Save Billions*. Boca Raton, FL: Taylor & Francis; 2019
2. Centers for Medicare and Medicaid Services. Report to congress: Fraud Prevention System—second implementation year. Available from: https://www.cms.gov/About-CMS/Components/CPI/Widgets/Fraud_Prevention_System_2ndYear.pdf; June 2014. Accessed February 1, 2020.
3. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: With Applications in R*. 1st ed. New York, NY: Springer; 2013.
4. Krishnan A, Xu T, Hutfless S, et al. Outlier practice patterns in Mohs micrographic surgery: defining the problem and a proposed solution. *JAMA Dermatol*. 2017;153(6):565-570.

https://doi.org/10.1016/j.jaad.2020.03.059