18. Hayashi T, Fleming MV, Stetler-Stevenson WG, Liotta LA, Moss J, Ferrans VJ, *et al.* Immunohistochemical study of matrix metalloproteinases (MMPs) and their tissue inhibitors (TIMPs) in pulmonary lymphangioleiomyomatosis (LAM). *Hum Pathol* 1997;28:1071–1078.
19. Nascimento ECTD, Baldi BG, Mariani AW, Annoni R, Kairalla RA, Pimenta SP, *et al.* Immunohistological features related to functional impairment in lymphangioleiomyomatosis. *Respir Res* 2018;19:83.
20. Atochina-Vasserman EN, Guo CJ, Abramova E, Golden TN, Sims M, James ML, *et al.* Surfactant dysfunction and lung inflammation in the female mouse model of lymphangioleiomyomatosis. *Am J Respir Cell Mol Biol* 2015;53:96–104.
21. Taveira-DaSilva AM, Steagall WK, Rabel A, Hathaway O, Harari S, Cassandro R, *et al.* Reversible airflow obstruction in lymphangioleiomyomatosis. *Chest* 2009;136:1596–1603.
22. Barnes PJ. Effect of beta-agonists on inflammatory cells. *J Allergy Clin Immunol* 1999;104:S10–S17.
23. Soltani A, Ewe YP, Lim ZS, Sohal SS, Reid D, Weston S, *et al.* Mast cells in COPD airways: relationship to bronchodilator responsiveness and angiogenesis. *Eur Respir J* 2012;39:1361–1367.
24. Niida Y, Stemmer-Rachamimov AO, Logrip M, Tapon D, Perez R, Kwiatkowski DJ, *et al.* Survey of somatic mutations in tuberous sclerosis complex (TSC) hamartomas suggests different genetic mechanisms for pathogenesis of TSC lesions. *Am J Hum Genet* 2001; 69:493–503.
25. Niida Y, Lawrence-Smith N, Banwell A, Hammer E, Lewis J, Beauchamp RL, *et al.* Analysis of both TSC1 and TSC2 for germline mutations in 126 unrelated patients with tuberous sclerosis. *Hum Mutat* 1999;14: 412–422.
26. Beauchamp RL, Banwell A, McNamara P, Jacobsen M, Higgins E, Northrup H, *et al.* Exon scanning of the entire TSC2 gene for germline mutations in 40 unrelated patients with tuberous sclerosis. *Hum Mutat* 1998;12:408–416.
27. Badri KR, Gao L, Hyjek E, Schuger N, Schuger L, Qin W, *et al.* Exonic mutations of TSC2/TSC1 are common but not seen in all sporadic pulmonary lymphangioleiomyomatosis. *Am J Respir Crit Care Med* 2013; 187:663–665.
28. Johnson SR, Clelland CA, Ronan J, Tattersfield AE, Knox AJ. The TSC-2 product tuberin is expressed in lymphangioleiomyomatosis and angiomyolipoma. *Histopathology* 2002;40:458–463.

Electronic Health Records and Machine Learning for Early Detection of Lung Cancer and Other Conditions

Thinking about the Path Ahead

In this issue of the *Journal*, Gould and colleagues (pp. 445–453) describe an innovative use of machine learning with stored patient electronic health record (EHR) data to develop a risk model assessing the short-term risk of non-small-cell lung cancer (1). The study authors identified two primary applications of the approach in clinical practice, namely providing patients and providers with a tool to assist in personalized decision-making and identifying persons for outreach and for potential eligibility for lung cancer screening with low-dose computed tomography (LDCT).

The model developed by the study authors, denoted by "MES," used demographic information, smoking history, clinical data, and laboratory data that were available in EHRs in their health maintenance organization's (HMO's) data warehouse. The MES model showed better prediction of lung cancer diagnosis within 3–12 months than the current standard LDCT eligibility criteria (which are based only on age, pack-years, and years since quitting) as well as better prediction than a well-known risk model based on detailed smoking history and demographics (the PLCOm2012 model) (1, 2).

Uptake of LDCT screening following the initial B recommendation of the U.S. Preventive Services Task Force (USPSTF) in 2013 has been slow and limited, with currently (before coronavirus disease [COVID-19]) only an estimated 5–10% of eligible individuals undergoing LDCT screening (3). Therefore, there is a critical need to use strategies to substantially increase this rate. Although with shared decision-making not all eligible individuals will choose to be screened, a rate of 50% or higher is desirable and potentially attainable.

Around half of patients in this HMO had missing data on pack-years, meaning that final determination of USPSTF eligibility could not be made based on EHRs alone (1). Use of the MES model could help identify, among those with missing data, those patients more likely to meet the USPSTF criteria. In addition, by helping estimate risk, it could assist with shared decision-making and potentially encourage individuals to choose to be screened.

The USPSTF recently updated their lung cancer screening recommendation, increasing the eligible pool by lowering the age and minimum pack-year requirements (4, 5) Their recommendation states that there was insufficient evidence to "assess whether or not risk prediction model–based screening would improve outcomes" (5). An argument against using current risk models (e.g., PLCOm2012) is that they incorporate age as a risk factor and thus skew eligible individuals toward older individuals. Although older people are at a higher risk of lung cancer, they also represent fewer potential life-years saved by screening, so using risk-based criteria may increase the number of lives saved but not necessarily the years of life saved (5). In the MES model, age was one of the top 10 most informative features, suggesting that this same issue applies (1).

For HMOs, which do not rely on fee-for-service reimbursement from the U.S. Centers for Medicare and Medicaid Services or private insurers, there is some leeway in deciding whom to screen for lung cancer. Accordingly, they could use either standard risk models such as PLCOm2012 or models developed from EHRs—such as MES—to broaden their eligibility criteria for screening. However, most healthcare

Originally Published in Press as DOI: 10.1164/rccm.202104-1009ED on June 7, 2021

providers and systems operate under fee-for-service models in which they are dependent on the policies of outside insurers and are thus constrained by existing coverage rules. It is not clear how feasible it would be for healthcare organizations to work with insurers (including the U.S. Centers for Medicare and Medicaid Services) to broaden coverage based on risk models.

More generally, in terms of the use of EHR data and machine learning for potential clinical interventions, there has been considerable research in this field in the last decade or so. Examples of conditions for which models were developed include heart failure, high-risk surgery, asthma, and cancer (6–9). However, this work—as with the work by Gould and colleagues—has generally been retrospective, in that it assessed the predictive ability of algorithms for known prior outcomes.

There is less research on how these models perform when implemented prospectively in clinical practice with actual resultant interventions. For many conditions for which prediction is possible, the benefit is not clear. In cancer screening, it is generally accepted that the benefit of early detection must be proven with a randomized controlled trial (RCT). The early detection of cancer does not necessarily reduce mortality from the cancer of interest, and there are also harms associated with cancer screening.

In thinking about using EHR-based models for patient care, several issues should be considered. First, there is the magnitude of increased risk identified and the degree of potential mitigation available through interventions (10). For LDCT lung cancer screening, even among eligible individuals, the risk of lung cancer death is only moderately high (about 2% within 6 yr), and the mitigation of that risk is modest (15–20% reduction) (5, 11). However, LDCT screening has been proven at least in RCTs to have a benefit in reducing lung cancer mortality. For the early detection of other conditions for which EHR-based prediction models have been developed, there is no such strong, RCT-based evidence.

There is no guaranteed benefit of early detection *per se*, and there are harms—including anxiety—and complications of work-ups. There are also costs of developing and maintaining EHR-based models, as well as practical issues. For example, how often would the model inputs need to be updated, at every new visit or lab test? Has the model been validated in various subpopulations, some of which may have been underrepresented in building the model? Also, there could be regulatory issues, depending on the extent to which the model was actually directing clinical interventions.

Clearly, it would not be feasible to perform an RCT of every predictive model developed from EHRs. However, it would be informative to see the results of several RCTs of such predictive models and their resultant interventions as proof of concept. For such RCTs, pragmatic trials could potentially be performed, with issues such as the need for patient consent decided on by institutional review boards.

There are ethical concerns as well. Do patients want artificial intelligence or machine learning systems assessing their EHRs in order to identify conditions that may be present subclinically or for which they may be at a high risk? Should patient consent be obtained for such searches? Furthermore, if the algorithm has limited explainability, as many do, how do physicians present the findings to patients? All of these are questions that are currently being debated in the medical artificial intelligence field (12, 13).

The first phase of research on using EHRs coupled with machine learning in clinical care settings has successfully demonstrated—as the with the model by Gould and colleagues—that such systems can successfully predict disease risk or disease onset better than standard approaches. Now it seems time to proceed to the more difficult second phase, which is to show that the prospective use of such systems to guide patient care actually has a net benefit. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Paul Pinsky, Ph.D.
Division of Cancer Prevention
National Cancer Institute
Bethesda, Maryland

References

1. Gould MK, Huang BZ, Tammemagi MC, Kinar Y, Shiff R. Machine learning for early lung cancer identification using routine clinical and laboratory data. *Am J Resp Crit Care Med* 2021;204:445–453.
2. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, *et al.* Selection criteria for lung-cancer screening. *N Engl J Med* 2013; 368:728–736.
3. Pham D, Bhandari S, Pinkston C, Oechsli M, Kloecker G. Lung cancer screening registry reveals low-dose CT screening remains heavily underutilized. *Clin Lung Cancer* 2019;21:e206–e211.
4. Moyer VA; U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014;160:330–338.
5. Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, Caughey AB, *et al.*; US Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *JAMA* 2021; 325:962–970.
6. Lu MT, Raghu VK, Mayrhofer T, Aerts H, Hoffmann U. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening CT: development and validation of a prediction model. *Ann Intern Med* 2020;173:704–713.
7. Vijayakrishnan R, Steinhubl SR, Ng K, Sun J, Byrd RJ, Daar Z, *et al.* Prevalence of heart failure signs and symptoms in a large primary care population identified through the use of text and data mining of the electronic health record. *J Card Fail* 2014;20:459–464.
8. Corey KM, Kashyap S, Lorenzi E, Lagoo-Deenadayalan SA, Heller K, Whalen K, *et al.* Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): a retrospective, single-site study. *PLoS Med* 2018;15:e1002701.
9. Juhn Y, Liu H. Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. *J Allergy Clin Immunol* 2020;145:463–469.
10. Pinsky P. Artificial intelligence and data mining to assess lung cancer risk: challenges and opportunities. *Ann Intern Med* 2020;173:760–761.
11. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, *et al.*; National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.
12. Vellido A. Societal issues concerning the application of artificial intelligence in medicine. *Kidney Dis* 2019;5:11–17.
13. McKernan LC, Clayton EW, Walsh CG. Protecting life while preserving liberty: ethical recommendations for suicide prevention with artificial intelligence. *Front Psychiatry* 2018;9:650.