And They Said It Couldn't Be Done: Predicting Known Driver Mutations From H&E Slides

Received: 27 November 2018 Accepted: 16 January 2019 Published: 06 May 2019

Summary

In the study by Coudray *et al*.[1] titled "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," the authors use a commercially available convolution neural network (CNN) platform (Google's Inception v3) to accurately classify different types of lung cancer and predict known and potential cancer driver mutations from hematoxylin and eosin (H&E) slides. Whole slide images from The Cancer Genome Atlas (TCGA) database were first divided into 512 × 512 pixel tiles and used to train the CNN to identify the two major histologic subtypes of non–small cell lung cancer (NSCLC), lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC). The performance of this algorithm when tested on a held-out dataset from TCGA was shown to be 0.97 of the area under the curve. The CNN classification model was subsequently validated on two independent cohorts of NSCLC. The validation accuracy remained high, regardless of tissue preparation for these cohorts (i.e., frozen; formalin-fixed, paraffin-embedded [FFPE]; and biopsies).

The authors also compared the deep-learning model classification to that of three pathologists who manually scored the test-set images and further compared the model to classifications per TCGA. The deep-learning model had comparable overall agreement (i.e., not statistically different) to the pathologists' classification using TCGA as the ground truth. When comparing each pathologist with another, the agreements ranged from 0.52 to 0.78, whereas the deep-learning model ranged from 0.64 to 0.77, indicating that the model had comparable inter-reader agreement with individual pathologists. The deep-learning model was also comparable to molecular profiling methods used to distinguish LUAD from LUSC. The authors also developed an "automatic tumor selection" model that was trained using pathologist-selected tumor areas and tested on different tissue preparations (frozen, FFPE, and biopsies). In each instance, automatic selection performed comparably to manual selection methods.

Importantly, the authors subsequently trained their model to predict ten specific driver mutations in NSCLC, including *STK11*, *EGFR*, *SETBP1*, *TP53*, *FAT1*, and *KRAS*, with accuracies of 0.856, 0.826, 0.775, 0.760, 0.750, and 0.733, respectively, when validated on the TCGA test set. The CNN model used to predict the *EGFR* mutation was subsequently tested in an independent cohort of lung resection specimens from New York University (NYU) Langone Medical Center that contained both wild-type and known *EGFR* mutation status. The CNN model predicted *EGFR* mutation status in the NYU cohort with ~69% accuracy, demonstrating better-than-chance estimates. The lower accuracy compared with validations in TCGA was attributed to the differences in *EGFR* mutation determination methods, namely the sequencing model used in TCGA and the immunohistochemistry (IHC)/sequencing model used by NYU.

Comments

There are two aspects of this report that are worth highlighting. The first is the perception that differentiating LUAD from LUSC will benefit from the use of artificial intelligence (AI), and the second is the observation that specific mutations can be predicted from H&E using AI. With respect to the former, it is not clear that AI will benefit lung cancer classification as much as the authors suggest. In the current treatment paradigm for newly diagnosed NSCLC, distinguishing histologic subtypes is critical both for triggering downstream molecular testing[2] and for defining appropriate chemotherapeutic regimens given with or without immunotherapy.[3] Histologic classification of NSCLC by microscopic review of H&E-stained slides relies on the identification of classic morphologic features. For poorly differentiated tumors, an IHC panel (i.e., TTF-1, p63, and CK5/6) can facilitate accurate classification.[4-6] Diagnoses are typically rendered within 1–2 days such that the added value of an AI algorithm for improving the speed of initial diagnosis is unclear.

In contrast, the utility of AI may be of great interest in defining the heterogeneity of tumor cell differentiation. NSCLC may show elements of both LUAD and LUSC histology, leading to a diagnosis of adenosquamous carcinoma when the less dominant histology constitutes at least 10% of the tumor.[7] While uncommon (0.4% to 4% NSCLC[8,9]), adenosquamous carcinoma is an aggressive tumor with inferior prognosis relative to LUAD or LUSC, and thus may be treated with chemotherapy even if early stage, underscoring the importance of accurate diagnosis.[10] Of note, in Supplementary Figure 1, Coudray *et al*. showed an example of a tumor classified overall as LUAD; however, a subset of the tiles were classified as LUSC by the algorithm.[1] If AI approaches can accurately identify small foci of divergent differentiation within a tumor, this could enable defining the true frequency and significance of this phenomenon with regard to association with a specific mutational profile, with stromal features including level and type of immune infiltrate, and potentially with response to immunotherapy. Further, such data have potential to provide

important information related to clonal heterogeneity, which may help inform new target and therapeutic approaches as well.

The second important aspect of the Coudray *et al*. study is the finding that AI can predict driver mutations from H&E images.[1] Next-generation sequencing (NGS) is gaining a foothold in diagnostic medicine as the list of potentially actionable driver mutations grows and the cost of NGS decreases.[11] Recent US Food and Drug Administration clearances for NGS panels developed by commercial laboratories and by academic cancer centers have paved the way for NGS to be used routinely in the assessment of molecular cancer subtypes, several of which are associated with specific therapeutic approaches.[12,13] The potential for assessing tumor mutational burden as a potential predictive biomarker for immunotherapy may further drive NGS demand.[14,15]

In relation to the significant advances in NGS, there is a perceived comparative paucity of technological advances in tissue-based H&E-stained slides. However, several important studies published over the past decade have reminded the scientific community that there are tremendous amounts of information in tissue sections that can be used to predict molecular test status and/or predict patient outcomes with accuracy similar to molecular methods.[16-18] While it is becoming clearer that deep learning applied to tissue-based pathology can predict outcomes, there has been little attempt to directly connect specific driver mutations to morphological patterns within cancer subtypes. It is relevant to note that the genetic predictions in the Coudray *et al*. study are within NSCLC subtypes. Other studies have linked cancer subtypes to unique molecular profiles,[19] but this is very different from identifying specific mutational status within morphological subtypes. The Coudray *et al*. report is the most comprehensive and well-validated method to date that has demonstrated such connections are feasible.

The practical implications of predicting driver mutation status from H&E should not be underestimated. The ability to make important treatment decisions regarding targeted therapy using a low-cost and accessible test, such as H&E, would be disruptive to the current NGS methods aiming to do the same. While this concept is exciting, we also need to be cognizant of the challenges related to improving predictive accuracy for stringent clinical validations, accessibility of digital pathology platforms, and the lack of standardization in H&E staining methods across laboratories, all of which are currently major obstacles to implementing H&E-based AI approaches broadly into medical practice.

Acknowledgment

Editorial assistance was provided by Jay Rathi, MA, of Spark Medica Inc., funded by Bristol-Myers Squibb.

Michael C. Montalto[1], Robin Edwards[1]

[1]Department of Translational Medicine, Bristol-Myers Squibb, Princeton, NJ 08543, USA

Address for correspondence: Dr. Michael C. Montalto, Translational Medicine, Bristol-Myers Squibb, Route 206 and Province Line Road, Princeton, NJ 08543, USA.
E-mail: michael.montalto@bms.com

References

1. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, *et al.* Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nat Med 2018;24:1559-67.
2. Lindeman NI, Cagle PT, Aisner DL, Arcila ME, Beasley MB, Bernicker EH, *et al.* Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: Guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. Arch Pathol Lab Med 2018;142:321-46.
3. Ettinger DS, Aisner DL, Wood DE, Akerley W, Bauman J, Chang JY, *et al.* NCCN guidelines insights: Non-small cell lung cancer, version 5.2018. J Natl Compr Canc Netw 2018;16:807-21.
4. Wu M, Wang B, Gil J, Sabo E, Miller L, Gan L, *et al.* P63 and TTF-1 immunostaining. A useful marker panel for distinguishing small cell carcinoma of lung from poorly differentiated squamous cell carcinoma of lung. Am J Clin Pathol 2003;119:696-702.
5. Gurda GT, Zhang L, Wang Y, Chen L, Geddes S, Cho WC, *et al.* Utility of five commonly used immunohistochemical markers TTF-1, napsin A, CK7, CK5/6 and P63 in primary and metastatic adenocarcinoma and squamous cell carcinoma of the lung: A retrospective study of 246 fine needle aspiration cases. Clin Transl Med 2015;4:16.
6. Koh J, Go H, Kim MY, Jeon YK, Chung JH, Chung DH. A comprehensive immunohistochemistry algorithm for the histological subtyping of small biopsies obtained from non-small cell lung cancers. Histopathology 2014;65:868-78.
7. Li C, Lu H. Adenosquamous carcinoma of the lung. Onco Targets Ther 2018;11:4829-35.
8. Ruffini E, Rena O, Oliaro A, Filosso PL, Bongiovanni M, Arslanian A, *et al.* Lung tumors with mixed histologic pattern. Clinico-pathologic characteristics and prognostic significance. Eur J Cardiothorac Surg 2002;22:701-7.
9. Travis WD, Brambrilla E, Müller-Hermelink HK, Harris CC, editors. WHO Classification of Tumours: Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart. Lyon, France: IARC Press; 2004.
10. Filosso PL, Ruffini E, Asioli S, Giobbe R, Macri L, Bruna MC, *et al.* Adenosquamous lung carcinomas: A histologic subtype with poor prognosis. Lung Cancer 2011;74:25-9.
11. Kamps R, Brandão RD, Bosch BJ, Paulussen AD, Xanthoulea S, Blok MJ, *et al.* Next-generation sequencing in oncology: Genetic diagnosis, risk prediction and cancer classification. Int J Mol Sci 2017;18. pii: E308.
12. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, *et al.* Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol 2013;31:1023-31.
13. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23:703-13.
14. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, *et al.* First-line nivolumab in stage IV or recurrent non–small-cell lung cancer. N Engl J Med 2017;376:2415-26.
15. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, *et al.* Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med 2018;378:2093-104.
16. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, van de Vijver MJ, *et al.* Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 2011;3:108ra113.
17. Whitney J, Corredor G, Janowczyk A, Ganesan S, Doyle S, Tomaszewski J, *et al.* Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer.

BMC Cancer 2018;18:610.
18. Rawat RR, Ruderman D, Macklin P, Rimm DL, Agus DB. Correlating nuclear morphometric patterns with estrogen receptor status in breast cancer pathologic specimens. NPJ Breast Cancer 2018;4:32.
19. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, *et al.* Comprehensive molecular portraits of invasive lobular breast cancer. Cell 2015;163:506-19.

Access this article online

Quick Response Code:

Website: www.jpathinformatics.org

DOI: 10.4103/jpi.jpi_91_18

How to cite this article: Montalto MC, Edwards R. And they said it couldn't be done: Predicting known driver mutations from H&E slides. J Pathol Inform 2019;10:17.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2019/10/1/17/257737