



## HEMATOPOIESIS AND STEM CELLS

Comment on Matek et al, page 1917

# Can the computer see what the human sees?

Aziz Nazha | Amazon Web Services

**In this issue of *Blood*, Matek et al<sup>1</sup> developed a computer vision model to differentiate between bone marrow cell morphologies on a large, expert annotated dataset.**

Machine learning (ML) is a branch of computer science in which an algorithm generates predictive models by learning from training data without explicitly being programmed.<sup>2,3</sup> Significant advances in ML algorithms have been achieved in recent years, especially in the field of computer vision (using ML algorithms to detect objects and patterns in images).<sup>2,3</sup> These advances allowed Google search to accurately label complex images, enable smartphones to recognize faces, and paved the way for the development of self-driving cars. Computer vision algorithms have also been used in healthcare to label radiological and pathological images.<sup>4,5</sup>

Computer vision uses deep neural networks that represent mathematical modeling of the neuron in the human brain.<sup>2</sup> These networks have 3 components: an input layer that takes source material from either text, structured data, or image; an output layer that provides the desired output; and multiple hidden layers that connect the input and output layers. The most common subtype of neural network used in computer vision is convolutional neural network (CNN). CNN can be used to classify images (identify if the image has a dog or a cat), segment images (draw a line around the desired prediction), or detect objects in the image by drawing a box around them. These methods can be used separately or combined in 1 intelligent system.

In this study, the authors trained a CNN model on 171 374 bone marrow cytology images from 961 patients diagnosed with variety of hematological diseases and evaluated at the Munich Leukemia Laboratory between 2011 and 2013. To prepare the slides to be used by the ML model, the authors first digitally scanned the whole slide images ( $\times 40$  oil immersion with original dimensions of  $2452 \times 2056$  pixels) for the morphological cell analysis. From the examined regions, diagnostically relevant cell images were annotated into 21 different classes (these classes include myeloid and lymphoid malignancies and nonmalignant conditions). The images were then annotated by experienced hematopathologists and sized at  $250 \times 250$  pixels. It is common to downsize the images before training them with a CNN model given that larger images can be difficult to train and require a lot of computational power. The authors acknowledged the imbalance between the classes given that some of these classes are less prevalent than others. To overcome that, the authors oversampled the imbalanced classes using image augmentation (a common technique to improve the ML model performance). The authors then used a CNN model (ResNext-50) to train the algorithm for recognizing each class. The ML model showed accurate performance for the classes with a higher number in the training samples, whereas the performance was lower on the rarer classes. This observation is common when

training ML models on imbalanced datasets. Whereas humans can be trained on a small number of images to identify objects, ML models require a large amount of data to be able to produce accurate results. The authors acknowledged the difficulties in identifying individual morphologies, especially when they are closely related in the leukocyte-differentiating lineage. The authors then validated the model on an independent data set of 627 single-cell images from 30 slides of 10 patients. Approximately 39% of these images in the validation cohort were classified as an artifact or not identifiable categories, suggesting that the classes of these images cannot be predicted by the model. This discrepancy between the model performance in the original dataset and the validation cohort could be related to different techniques in staining and annotating the images, but also raise an important point when evaluating the performance of an ML model in healthcare. The reproducibility of the model in multiple different clinical settings is an important part of the successful implementation/adaptation of this model in the clinical workflow and should be required before implementing these algorithms in any hospital or laboratory.

The authors also completed another important step in building an ML model by performing an explainability analysis (trying to identify the structure that the algorithm is looking at when making a prediction). Although the model was able to identify the structure that human looks at in common images, it was less robust in rarer images. Several studies have shown that CNNs in some cases are using irrelevant structures in radiological or pathological images to make a prediction of an outcome.<sup>6,7</sup> Explainability of a ML algorithm represent another important step of the successful adaptations of these models in clinical practice.

This study represents a significant step toward the integration of ML models in

clinical practice. The authors professionally annotated a very large dataset and used a state-of-the-art computer vision algorithm to train and validate the model. The dataset can be used for future research to improve the model performance and to integrate it in the workflow of hematopathologists.

In summary, artificial intelligence and ML algorithms are changing our lives. These technologies will have significant impact on healthcare in the next decade. Although these algorithms may not replace physicians and researchers, it will definitely aid them in providing better care/research that can improve patient lives. As Oren Harari once said: "The electric light did not come from the continuous improvement of candles." If we really want to have a significant impact on healthcare in the future, we need to start embracing the impact of these technologies and learn how to use and integrate them into our daily workflow.

*Conflict-of-interest disclosure:* Speaker bureaus for Incyte Corporation and Novartis; data monitoring committee for MEI Pharma; advisory board/consulting (pharmaceutical/

biotechnology) for AbbVie; and stock in Amazon. ■

## REFERENCES

1. Matek C, Krappe S, Muzenmayer C, et al. Highly accurate differentiation of bone marrow cell morphologies using deep neural networks on a large image dataset. *Blood*. 2021; 138(20):1917-1927.
2. Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies. *Lancet Haematol*. 2020;7(7):e541-e550.
3. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24-29.
4. Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *J Am Coll Radiol*. 2018;15(3 Pt B):504-508.
5. Komura D, Ishikawa S. Machine learning approaches for pathologic diagnosis. *Virchows Arch*. 2019;475(2):131-138.
6. DeGrave AJ, Janizek JD, Lee SI. AI for radiographic COVID-19 detection selects shortcuts over signal. *Nat Mach Intell*. 2021; 3(7):610-619.
7. Geirhos R, Jacobsen JH, Michaelis C, et al. Shortcut learning in deep neural networks. *Nat Mach Intell*. 2020;2(11):665-673.

DOI 10.1182/blood.2021013259

© 2021 by The American Society of Hematology

## CLINICAL TRIALS AND OBSERVATIONS

Comment on Kulasekararaj et al, page 1928

# Control of hemolysis in patients with PNH

**Lucio Luzzatto** | Muhimbili University of Health and Allied Sciences; University of Florence

**Paroxysmal nocturnal hemoglobinuria (PNH) is characterized by the triad of intravascular hemolysis (IVH), a tendency to thrombosis, and an element of bone marrow failure. With the introduction of the anti-C5 antibody eculizumab (ECU),<sup>1</sup> IVH and associated symptoms are abrogated, and the risk of thrombosis is greatly reduced; however, most patients remain anemic, and some remain transfusion dependent.<sup>2</sup> In this issue of *Blood*, Kulasekararaj et al<sup>3</sup> report that danicopan (DNC), an inhibitor of complement factor D, added to ECU, corrects or at least ameliorates the anemia.**

This important clinical result is also of great interest with respect to pathophysiology (see figure panels A-B). When a patient with PNH is on ECU, red cells no longer suffer complement-mediated lysis, but unlysed red cells undergo opsonization by C3 fragments (see figure, panel C) and are thus marked for removal by macrophages. In patients who are on ECU,

extravascular hemolysis (EVH) is a regular iatrogenic feature<sup>4</sup>; its clinical correlates are persistent anemia with reticulocytosis and hyperbilirubinemia, sometimes with visible jaundice.

Overcoming EVH has been a major stimulus for introducing agents that act upstream of C5 on the proximal

complement pathway. There are now 3 such agents: pegcetacoplan (APL2), which is already approved by the US Food and Drug Administration,<sup>5</sup> as well as iptacopan (IPT)<sup>6</sup> and DNC, both of which are in phase 3 trials. In this context, the findings by Kulasekararaj et al are most relevant: DNC, in patients who are already on ECU, causes an increase in hemoglobin, a decrease in reticulocytes, and a decrease in bilirubin, which is exactly what one would expect when, with IVH already controlled by ECU, EVH is curbed as well. Of course the patients feel better, and Figure 3 in the article by Kulasekararaj et al illustrates a gratifying abolition of blood transfusion requirements in all but 1 patient. The decrease in the proportion of C3d-coated red cells is not as great as one might have hoped; perhaps plasma levels are not always optimal, and the dosage may need adjustments. Patient-to-patient variability in this parameter might also depend on the complement receptor 1 genotype,<sup>7</sup> which will require testing.

Because the complement system is part of innate immunity, tampering with it should not be taken lightly. Congenital C3 deficiency is associated with a risk for serious infection from capsulated and gram-negative bacteria; iatrogenic C3 deficiency may be less complete, but the same risk can be expected. We note, however, that with normally developed acquired immunity, the defensive function of complement may be less crucial in adults than in children.<sup>8</sup>

Inhibition of complement at the level of C3 can also largely preempt the distal pathway; therefore, in principle, any of the 3 agents mentioned above might control hemolysis in PNH on its own. This is not to say that the 3 are equivalent: APL2 targets C3 itself, a proenzyme and one of the most abundant plasma proteins (1.2 g/L); IPT targets factor B (plasma concentration of 200 mg/L), another proenzyme of the alternative pathway C3 convertase; and DNC targets factor D (plasma concentration 1 mg/L), the serine protease that cleaves/activates factor B (see figure). All of the drug-target interactions have been solved at the molecular level, but target concentrations and characteristics of the 3 drugs differ: the choice among them may not prove easy. Some of us believe that the price of a drug should not be dictated simply by market forces<sup>9</sup>; however, the benchmark will be the current