

## Towards artificial intelligence-driven pathology assessment for hematological malignancies

Olivier Elemento, PhD

Caryl and Israel Englander Institute for Precision Medicine  
Weill Cornell Medicine  
New York, NY, 10021

**In this issue of Blood Cancer Discovery, Brück *et al* applied unsupervised and supervised machine learning to bone marrow histopathology images from Myelodysplastic Syndrome (MDS) patients. Their study provides new insights into the pathobiology of MDS and paves the way for increased use of artificial intelligence for the assessment and diagnosis of hematological malignancies.**

Myelodysplastic syndrome (MDS) is a group of hematological disorders characterized by ineffective and dysplastic hematopoiesis in the bone marrow. For decades, the diagnosis of MDS has relied almost solely on visual morphologic analysis of stained bone marrow samples, involving assessment of cellular dysplasia, blast count, cellularity, monocytes, ring sideroblasts and a few other features. MDS have been further classified into subtypes based on the World Health Organization (WHO) classification scheme, which combines morphology with blood count and cytogenetics <sup>1</sup>. The number and complexity of features used for diagnosis and the subjectivity of their assessment can make a precise diagnosis difficult to achieve in some cases, even for the highly trained hematopathologist. As a result, bone marrow assessments are known to be associated with substantial inter-pathologist variability. More recently, genomic analysis of MDS has allowed for further characterization of MDS cells. A number of frequently mutated genes were identified, e.g. TET2, ASXL1, SF3B1, SRSF2, DNMT3A and RUNX1 <sup>2</sup>. With a few exceptions such as the megakaryocyte dysmorphology in 5q-deleted MDS or presence of ring sideroblasts in *SF3B1*-mutated MDS, the relationship between morphology and mutations remains poorly characterized. This is unfortunate, since the concordance between normally correlated features can reinforce the confidence in each individual feature and may contribute to enhancing diagnostic accuracy.

Can machine learning (ML, a form of artificial intelligence or AI) and automated image analysis provide additional insights for MDS? This is the problem that Brück *et al.* sought to address in this issue of Blood Cancer Discovery <sup>3</sup>. Histopathology slides are now routinely digitized and therefore amenable to automated or semi-automated image analysis. Large datasets combining these images with genomic and clinical information are being collected at a number of medical centers. In other malignancies, the application of ML to histopathology slides has been shown to provide diagnostic and prognostic accuracy on par with trained diagnosticians. For example, ML can diagnose skin malignancies from digital images with performance rivaling trained dermatologists <sup>4</sup>. ML can also distinguish accurately between select subtypes of cancers <sup>5</sup>, predict prognosis in mesothelioma <sup>6</sup> and even predict the presence of specific mutations e.g. EGFR in lung adenocarcinomas <sup>7</sup>.

Brück *et al* leveraged some relatively recent key developments in the field of machine learning applied to images. One of them is called transfer learning. Transfer learning refers to the re-utilization of deep neural networks that have already been trained to recognize a multitude of non-medical images and objects from huge image databases such as the well-known ImageNet

database<sup>8</sup>. By mimicking the way the human visual cortex works, these networks have learned to recognize basic image features such as shapes and textures. These features are encoded within the networks during training, a lengthy process that takes supercomputers and gigantic databases such as ImageNet to achieve successfully. Once a network is trained, new images, including pathology images, can be entered as input and propagated through the network. The network features that are activated (much in the same way that neurons are activated) during this process can be recorded. Thus a pathology image can be described by all the internal features that are activated in a pre-trained neural network.

There remains many obstacles to applying these pre-trained neural networks to pathology images. One is that entire slides of variable and sometimes large sizes cannot typically be directly used as input to trained network. Instead, digitized slides have to be split into smaller and uniformly sized tiles. This is a limitation as tiles need to somehow be combined but also an opportunity to capture intra-tissue heterogeneity. Moreover, the pre-trained network features are automatically generated and not selected based on prior knowledge of bone marrow morphology. To make them interpretable for a trained hematopathologist, they must be decoded.

Brück *et al* analyzed bone marrow samples from 143 MDS patients; 51 patients with myelodysplastic/myeloproliferative neoplasm (MDS/MPN) and 11 control subjects. In total 500 tiles were analyzed. Upon extracting features using the widely used VGG16 and Xception convolutional neural networks (CNNs), Brück *et al* performed an unsupervised analysis by mapping each tile onto a two-dimensional space using the Uniform Manifold Approximation and Projection (UMAP) technique, most frequently used for analyzing single cell transcriptomic data. In this representation, tiles with similar image feature profiles are located close to each on the two-dimensional space. Five clusters emerged and visual analysis of tiles within each cluster shed light on their content. These clusters captured high abundance of lipid droplets, red blood cells, stroma, hypocellular and hypercellular tiles. This served as validation that the deep learning-extracted features do in fact reflect biologically relevant features, a non-negligible observation given the fact that the CNNs were trained on ImageNet, a database that contains few if any histopathological images. Brück *et al* then averaged feature profiles for tiles belonging to the same sample, then reprojected the samples in two dimensions using UMAP, followed by unsupervised clustering. Five clusters emerged again. Healthy subjects formed a distinct cluster and were not only homogenous but also highly distinguishable from the four MDS clusters. Each of the four clusters were enriched in specific MDS WHO subtypes. However, MDS sample groupings only partially overlapped with the WHO subtypes. This is not unexpected since WHO subtypes are defined not only by bone marrow cytomorphology but also by cytogenetics, blast proportion and blood cell counts.

Brück *et al* then moved on to supervised analyses, seeking to train statistical models based on elastic net regularized regression to predict a variety of tumor and patient characteristics based on image features. Certain characteristics such as mutations in TET2, ASXL1, and STAG2, chromosome 7 monosomy, and 7q deletion turned out to be highly predictable from morphological features. This demonstrates that the links between morphological features and mutations is more extensive than previously thought in MDS. Brück *et al* were also able to predict the risk stratifying IPSS-R score (Revised International Prognostic Scoring System), overall survival and progression to Acute Myeloid Leukemia (AML) with high accuracy by solely employing features extracted from stained slides. Strikingly, progression to AML was best predicted using a model that combined histopathology features and conventional IPSS-R scores. A model was able to discern MDS patients from MDS/MPN patients with an accuracy of

0.81 (1.0 being the maximum achievable accuracy). Analysis of direct cell and image segmentation and classification showed as expected, that predicted MDS samples were hypoplastic with higher stromal involvement.

This is not the first time that machine learning has been used to assess MDS. For example, Nagata *et al* successfully used morphological features from pathology reports to predict mutational profiles<sup>9</sup>. Mori *et al* were able to accurately assess dysplasia in bone marrow smear using a convolutional neural network<sup>10</sup>. Brück *et al* demonstrate how much the field of AI has advanced in the past few years, especially the concept of transfer learning. One can now use off-the-shelf pre-trained deep neural networks and deploy them on a new medical image dataset of modest size (a few hundred images) without need for extensive retraining (which would take a much larger dataset to achieve).

Altogether Brück *et al*'s work represents an exciting step towards AI-driven pathological assessment in MDS and more generally in hematologic malignancies. Although not the central goal of Brück *et al*, their work clearly suggests that MDS can be distinguished from normal bone marrow with minimal to no ambiguity using ML. One would imagine that once extended to a larger dataset, validated with data from other independent centers, the AI behind Brück *et al* may be used to help improve or refine MDS diagnosis, especially in cases that are difficult to diagnose or classify using conventional methods. There may be additional applications for approaches like Brück *et al*. For example, the unbiased assessment that AI provides can also be used either in preclinical animal models or in clinical trials to assess the effect of investigational therapies on the bone marrow of MDS patients.

Despite the advances described in Brück *et al* and in many other recent papers applying AI to medical images, much work remains to be done to enable routine use of AI in pathology and other fields. AI models ideally need to be trained on diverse data from multiple medical centers in order to minimize biases associated with individual centers. Unfortunately, data sharing across medical centers remains infrequent and difficult due to competition between centers and sometimes unnecessary regulatory requirements. Prospective validation of AI models using randomized clinical trials remains too rare. One of the key barriers to adopting medical AI is the limited interpretability of AI models. By opting to use transfer learning instead of directly applying convolutional neural networks to pathology images, Brück *et al*'s approach has the advantage of generating features whose importance for classification can rather easily be assessed. That does not mean that they can readily be understood, even by a trained hematopathologist. As shown in Brück *et al*, these features need to be decoded. AI models also need to further integrate the complexity of tumors. For example, averaging feature profiles for the same sample as performed by Brück *et al* works but also loses information on potential intra-tumor heterogeneity. Such heterogeneity, if better captured, may help improve predictive models based on histopathology slides.

Will conventional morphological assessment and the ubiquitously applied WHO classification soon be replaced by artificial intelligence? Probably not, but Brück *et al* outlines an exciting future where cooperation between hematopathologists and AI improves the quality and depth of pathology assessments for MDS and other malignancies.

## References

1. Malcovati L, Hellstrom-Lindberg E, Bowen D, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood*. 2013;122(17):2943-2964.
2. Haferlach T. The Molecular Pathology of Myelodysplastic Syndrome. *Pathobiology*. 2019;86(1):24-29.
3. Brück O, Lallukka-Brück S, Hohtari H, et al. Machine Learning of Bone Marrow Histopathology Identifies Genetic and Clinical Determinants in MDS Patients. *Blood Cancer Discovery*. 2021;In press.
4. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
5. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images. *EBioMedicine*. 2018;27:317-328.
6. Courtiol P, Maussion C, Moarii M, et al. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med*. 2019;25(10):1519-1525.
7. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-1567.
8. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition; 2009.
9. Nagata Y, Zhao R, Awada H, et al. Machine learning demonstrates that somatic mutations imprint invariant morphologic features in myelodysplastic syndromes. *Blood*. 2020;136(20):2249-2262.
10. Mori J, Kaji S, Kawai H, et al. Assessment of dysplasia in bone marrow smear with convolutional neural network. *Sci Rep*. 2020;10(1):14734.

# BLOOD CANCER DISCOVERY

## Towards artificial intelligence-driven pathology assessment for hematological malignancies

Olivier Elemento

*Blood Cancer Discov* Published OnlineFirst March 22, 2021.

<b>Updated version</b>	Access the most recent version of this article at: doi: <a href="https://doi.org/10.1158/2643-3230.BCD-21-0048">10.1158/2643-3230.BCD-21-0048</a>
<b>Author Manuscript</b>	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

<b>E-mail alerts</b>	<a href="#">Sign up to receive free email-alerts</a> related to this article or journal.
<b>Reprints and Subscriptions</b>	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a> .
<b>Permissions</b>	To request permission to re-use all or part of this article, use this link <a href="http://bloodcancerdiscov.aacrjournals.org/content/early/2021/03/19/2643-3230.BCD-21-0048">http://bloodcancerdiscov.aacrjournals.org/content/early/2021/03/19/2643-3230.BCD-21-0048</a> . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.