Machine learning throws away the buckets to better understand red blood cell quality

What is this research about?

In Canada, red blood cells for transfusion are stored in the refrigerator at 1-6°C for up to 42 days, after which they are discarded. During storage, the cells change as they metabolize and age. This leads to accumulated degradation of their function and safety, which is seen as the cells change shape from a smooth disc to spiky sphere then a smooth sphere. This red blood cell "storage lesion" can be analyzed by laboratory tests. For example, cell shape is usually measured by experts who prepare the cells, spread them on a slide, look at them using a microscope and categorize their shape according to standard definitions, which places each cell into one of six shape sub-classes or "buckets". These data are used to give a "morphology index" (MI) for the red blood cells. Using this traditional method, the loss of quality of red blood cells during storage has been very well characterized by researchers. However, the method is complex, time- and labour-intensive, prone to subjective bias, and limited by small sample sizes. The researchers aimed to address these limitations and come up with better methods to assess cell quality using label-free imaging and deep convoluted neural network learning algorithms.

IN BRIEF: Automated machine learning has potential to better predict red blood cell quality than current manually annotated approaches

What did the researchers do?

Fully supervised learning: The researchers analyzed red blood cells by imaging flow cytometry, a laboratory test that requires little preparation of the cells. Imaging flow cytometry takes pictures of single cells at rates of 100s or 1000s per second, yielding large numbers of images suitable for deep learning algorithms. The researchers imaged red blood cells every 3-7 days during their 42-day storage and used the pictures to train a neural network to classify images.

Weakly supervised learning: Seeking to improve on the first strategy, the researchers took a different approach in which the neural network learns about the shape of red blood cells independent of the six visual "buckets" defined by experts. Instead, they trained the network based on the storage age of the red blood cell sample being analyzed – a factor that is always known and is correlated with red blood cell quality. This approach eliminates the subjective annotation developed by humans and allowed the machine to independently determine features that correspond to morphology and storage age. The researchers then trained the neural network to estimate the storage age of more than 1 million red blood cell images.

What did the researchers find?

Fully supervised learning: Five researchers annotated over 52,000 red blood cells spanning many units, creating the largest freely available dataset of its kind. Using these images, the fully supervised model was able to approximate humans in categorizing red blood cell morphology. The trained neural network achieved >76% agreement with an expert. This can be compared to ~79% intra-expert agreement usually associated with this assay when conducted by humans.

Weakly supervised learning: This approach was not particularly accurate in predicting storage age from single cell images. However, it did reveal a chronological progression of morphological changes that might predict red blood cell quality and the storage age of red blood cells without human-curated annotations. A new measure of red blood cell quality was derived based on this progression: Self-learned Morphological Index (SMI). SMI was found to better correlate to a biochemical red blood cell assay (hemolysis) than the standard morphology index.

How can you use this research?

This study shows that imaging flow cytometry and deep learning can evaluate red blood cell morphology and quality. The simple label-free cell preparation for imaging flow cytometry offers advantages over traditional techniques. The neural network-based approach opens up many possibilities. For example, a single network could be used across many facilities and instruments allowing universal, objective, and standardized analysis. To achieve this, training on a greater variety of images would be needed to make the method more robust.

Overall, having an improved method to assess quality of blood products for transfusion would help advance blood transfusion research to identify donor factors and manufacturing methods that produce higher quality products, potentially leading to better patient outcomes. This approach could also have applications to other biological systems involving chronological progression, such as cancer.

About the research team: This research was conducted by a large international and interdisciplinary team led by **Anne Carpenter** of the Broad Institute of Harvard and MIT, **Michael Kolios** of Ryerson University's Department of Physics, and **Jason Acker,** senior scientist at Canadian Blood Services and a professor of laboratory medicine and pathology at the University of Alberta. , From Dr. Acker's group, senior research assistant T**racey Turner**, and postdoctoral fellow **Olga Mykhailova** are authors.

This Research Unit is derived from the following publication(s):
[1] Doan M, Sebastian JA, Caicedo, JC, Siegert S. Roch A, Turner TR, Mykhailova O, Pinto RN, McQuin C, Goodman A, Parsons MJ, Wolkenhauer O, Hennig H, Singh S, Wilson A, Acker JP, Rees P, Kolios MC, Carpenter AE. Objective assessment of stored blood quality by deep learning. *PNAS*, 2020. DOI: 10.1073/pnas.2001227117.

Acknowledgements: This research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR) Collaborative Health Research Projects grant to MC Kolios and JP Acker. The research was also funded by the US National Science Foundation / UK Biotechnology and Biological Sciences Research Council to AC Carpenter. O Mykhailova is supported by a Canadian Blood Services' Postdoctoral Fellowship. Canadian Blood Services, funded by the federal government (Health Canada) and the provincial and territorial ministries of health, supports JP Acker and his laboratory. The views herein do not necessarily reflect the views of Canadian Blood Services or the federal, provincial or territorial governments of Canada. Canadian Blood Services is grateful to blood donors for making this research possible.

Keywords: red blood cell, quality, quality control, storage lesion, machine learning, algorithm, neural network.

Want to know more? Contact Dr. Jason Acker at jason.acker@blood.ca.