Project title: Statistical and machine learning analysis of images of platelet formation

Supervisor(s): William Astle (MRC BSU, NHSBT) and Cédric Ghevaert (Department of Haematology, Stem Cell Institute, NHSBT)

Abstract: Platelets are small blood cells, which are necessary for the blood to form clots. If a person cannot make enough platelets or if a person’s platelets do not work properly he or she is likely to suffer from regular bleeding and bruising, because routine damage to the surface of blood vessels will go unrepaired. To stop such bleeding, it may be necessary to treat patients with transfusions of platelets from blood donors. However, this treatment is not ideal, because patients form antibodies against the transfused platelets meaning they cannot be treated again with platelets of the same blood type. Consequently, blood services are interested in developing alternatives to donated platelets. One approach is to generate platelets artificially, in the hope that, in future, some difficult to treat patients might be given laboratory grown platelets. Another is to identify compounds that, when used as drugs, increase the rate of platelet production. In order to pursue either of these possibilities, we need to study how platelets are generated in healthy people.

We know that platelets are produced, by a series of intriguing biological processes, from their progenitor cell - the megakaryocyte - in the bone marrow. As a megakaryocyte matures its genome replicates, which can result in the nucleus containing up to 8 times the usual number of copies of each chromosome. The cytoplasm fills with granules which contain proteins used in clot formation and the cell membrane grows extensively. Long protrusions called proplatelets are formed from the main body of the cell, the tips of which ultimately break off to form mature platelets. At present we do not fully understand the biological mechanisms that lead to proplatelet formation. In order to do this, we need a reliable way to measure the tendency of a megakaryocyte to form proplatelets. One approach is to use microscopy imaging of megakaryocytes and count the number of proplatelets. Unfortunately, this is extremely labour intensive as it needs a scientist to identify megakaryocytes and count the number of pro-platelets, often in several hundred images for a single experiment.

The aim of this project is to develop the statistical models and machine learning algorithms necessary to automate the analysis of microscopy images of megakaryocytes. The student will develop methods to identify the regions of images corresponding to individual cells, segment them, classify them according to cell type and identify and count the proplatelets visible on the megakaryocytes. These analyses may require a variety of techniques including the application of convolutional neural networks, morphological transformations and graph based representations of images. The student will also develop statistical models for the count distribution of pro-platelets, seeking to explain variation at the level of the image, the individual and the experimental treatment and also by experimentally relevant covariates.

The project will be supervised jointly by NHS Blood and Transplant supported research groups in the MRC BSU and the Department of Haematology, respectively specialising in statistical/data science and the development of methods to generate artificial blood cells for transfusion.

All application queries regarding eligibility should be directed to phdstudy@mrc-bsu.cam.ac.uk