




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# Advances in High-Content Imaging and Informatics: A Joint Special Collection with Society for Biomolecular Imaging and Informatics and SLAS

SLAS Discovery  
2021, Vol. 26(9) 1073–1075  
© Society for Laboratory  
Automation and Screening 2021  
DOI: 10.1177/24725552211042299  
[journals.sagepub.com/home/jbx](http://journals.sagepub.com/home/jbx)  


Justin D. Boyd<sup>1</sup> and Ann F. Hoffman<sup>2</sup>

This Special Collection of *SLAS Discovery* contains four original research articles and one perspective. The topics are diverse, ranging from a COVID-related screen, to two examples of enabling platforms for screening, to an assessment of deep learning technologies for label-free segmentation, and to a perspective on industrializing drug discovery. As in the previous SBI2 Special Issue of *SLAS Discovery*, most of these articles were recruited from the SBI2 membership following the SBI2 virtual conference last September.

Without a doubt, the COVID-19 pandemic was the single most impactful global event of 2020 and continues with nearly 200 million documented cases globally, including more than 4 million deaths (source: Johns Hopkins University Center for Systems Science and Engineering). Contributing to the high morbidity and mortality rate of SARS-CoV-2, the coronavirus causing COVID-19 is the associated widespread lung injury caused by the immune response to infection leading to pulmonary fibrosis, among other maladies. With the abundance of COVID-associated lung injuries, the need to identify new drugs and/or different mechanisms contributing to pulmonary fibrosis is great. In this Special Collection of *SLAS Discovery*, Marwick et al.<sup>1</sup> describe their effort to develop a phenotypic-based drug screen using fibrotic readouts from primary human pulmonary fibroblast in a high-content screen. The authors describe a screen of 2743 compounds that yielded hits inhibiting fibrogenesis in a model of pulmonary fibrosis. Their methods could be applied broadly to different pulmonary fibrosis-associated indications, including those caused by COVID infections. The authors remind us that there is no clear mechanism causing this type of injury, and there are only a couple of drugs available to treat pulmonary fibrosis. Marwick et al. take a relatively simple, yet clever, approach to this high-content screen: they focused on quantifying several extracellular matrix (ECM) readouts, specifically collagen I+III, collagen IV, and fibronectin, for the primary screen. This required the removal of cells in order to unambiguously quantify the ECM. A caveat to this approach is the inevitable identification of false positives due to toxicity. The clever addition of live-cell imaging to the screening process ensured an easy removal of false

positives caused by toxicity. The authors then followed up the screen to validate the hits with three additional assays, including the assessment of alpha-smooth muscle actin expression, fibroblast proliferation assay, and migration assay. A live-cell toxicity counterscreen of hits was also used to evaluate the potency of hits on primary fibrogenic readouts versus toxicity. Readers will be encouraged to find that one of the hits, RepSox, targets ALK5, a transforming growth factor- $\beta$  receptor inhibitor. An added bonus of this work is the inclusion of an informatics-based target pathway prediction based on the compound structure of the hits, resulting in starting points from both compounds and putative pathways for future follow-up studies.

Additional common COVID-associated conditions are acute respiratory distress syndrome (ARDS) and acute lung injury, both characterized by rapid onset of pulmonary edema and subsequent lung failure. A disruption of the epithelial and endothelial barriers in the lungs is thought to be the cause of the pulmonary edema. There are no current drugs on the market to address this mechanism. Dubrovskiy et al.<sup>2</sup> describe a novel approach to screen for endothelial barrier function using image cytometry. Prior to this platform, complicated systems with transwell and transendothelial resistance platforms were commonly used to study the effects of perturbations on barrier function. These are both relatively expensive and specialized assays with challenges of scalability. Moreover, these systems do not enable one to investigate barrier function at the single-cell level, making it difficult to assess the heterogeneity or homogeneity of endothelial barrier function response. In contrast, the approach that Dubrovskiy et al. takes is the Express

<sup>1</sup>Vaxxinity, Inc., Exploration Park, Florida, USA

<sup>2</sup>GlaxoSmithKline, Collegeville, PA, USA

## Corresponding Authors:

Justin D. Boyd, Vaxxinity, Inc., Space Life Sciences Lab-Space Florida, 505 Odyssey Way, Exploration Park, FL 32853, USA.  
Email: [justin@vaxxinity.com](mailto:justin@vaxxinity.com)

Ann F. Hoffman, GlaxoSmithKline, NCE / Drug Design and Selection, 1250 S Collegeville Road, Collegeville, PA 19426, USA.  
Email: [ann.f.hoffman@gsk.com](mailto:ann.f.hoffman@gsk.com)

Permeability Test (XPerT), in which biotinylated gelatin coats the bottom of wells in a multiwell assay plate (96-well imaging plate). Primary human pulmonary artery endothelial cells were then plated on top of the gelatin and allowed to grow in high density to form tight junctions. Cells were treated with various stimuli (thrombin, tumor necrosis factor- $\alpha$ , and lipopolysaccharide) for different lengths of time. After exposure to stimuli, cells were rinsed and (very) quickly exposed to an FITC-conjugated avidin. After aspiration and fixation, wells were imaged in brightfield and fluorescence. Total FITC-avidin signal was then quantified as a proxy readout for membrane permeability. Key to this assay was whole-well labeling with single-cell resolution. Such an assay could easily be scaled up (384- and 1536-well plates) for high-content screening. It would be interesting to explore multiplexing this assay with other markers for screening. Furthermore, by adding markers, this assay is primed for high-dimensional profiling and advanced analytics for clustering hits from a screen or genetic perturbations for target validation studies. Nonetheless, such an assay should be quite easy to deploy compared with the transwell or TER options. Both could be used to validate the readouts of orthogonal secondary assays. We expect this assay could broadly impact the ARDS field, with immediate and urgent need in the COVID context.

Another article in this Special Collection focuses on platform development for screening drugs to treat prostate cancer. It is almost universally accepted that three-dimensional (3D) cultures systems recapitulate *in vivo* biology better than two-dimensional monolayer cultures. However, the ability to screen in 3D for high-content readouts is an inherent challenge. Over the past several years, there has been a significant improvement in assays for enabling organoid/spheroid screening, thanks to advancements in 3D culture at scale and analytical tools to quantify organoid phenotypes. Choo et al.<sup>3</sup> describe an elegant workflow to establish a biologically relevant screening platform of prostate cancer organoids. The authors note that prostate organoids are particularly difficult to establish, making established patient-derived xenografts (PDX) and subsequent organoids more precious. Having a screening assay that maximizes the insight into cell health maximizes the stewardship of these resources. The authors harvest cells from patients with diverse clinical profiles and establish stable organoids through a relatively complex and time-consuming workflow from patient tissue to PDX in mice to organoids. Key to this workflow was automated cell and assay preparation at the organoid step, with close monitoring of cell cultures using live-cell imaging. The assay itself capitalizes on label-free segmentation of organoids for quantifying phenotypes. Multiple readouts for organoid morphology, including area, shape, and texture measurements, were the key quantitative features for analysis. The authors note the inherent heterogeneity among and within organoids derived from different patient cells, with the organoid radius being

especially sensitive in distinguishing different patient cells. However, overall morphologies at the population level were equivalent for four of five patient-derived lines. Fluorescence labeling with Hoechst was used to confirm the brightfield morphological analysis. The authors proceeded to examine a PARP inhibitor effects on organoid morphology, noting Hoechst labeling was more sensitive to dose-dependent PARP treatment when binning organoids by size. These results might reflect the limitation of conventional brightfield segmentation compared with fluorescence labeling. Nonetheless, having a workflow that enables scaling 3D cultures for screening offers tremendous value to the prostate cancer field. It is likely that either specific marker labeling or cell dyes could be applied to this workflow to add content for screening. The addition, for instance, of DRAQ5 alone could be informative and allow for higher magnification and resolution while providing deep penetration into organoids. Alternatively, better brightfield analysis, such as those that use artificial intelligence (AI) to segment images, could enable more subtle distinction among the morphologies of the organoids sensitive to treatments.

It just so happens that the final original research article of this Special Collection of *SLAS Discovery* examines several deep learning, an AI, approaches to segment brightfield images—nuclei segmentation in this case. Ali et al.<sup>4</sup> evaluate five convolutional neural networks (CNNs) for segmenting nuclei from different (albeit conventional monolayer) cell lines. The authors considered four to be the state-of-the-art very deep CNN and offer a fifth that they developed in response to the performance of the others. The models that the authors evaluated were U-Net, U-Net++, DeepLabv3+, and Tiramisu. They describe these models as “end-to-end trained encoder-decoder networks with a down-sampling contraction path, an up-sampling expansion path, and a bottle neck to connect them.” The authors go into details about the specific network components, such as the expansion and contraction paths, convolution layers, skip connections and filters, for each model. As a result of the models’ architectures, there is a wide range of trainable parameters among these models (ranging from approximately 1 million to 40 million parameters). The models were used to train and test for brightfield and fluorescence-labeled nuclear segmentation for pixel-wise and object-wise scoring for segmentation. Not surprisingly, the authors observed higher performance with the fluorescence-labeled images compared with the brightfield images (>95% for all models and >81% for all models for fluorescence and brightfield, respectively). The authors investigated different approaches to improving brightfield segmentation, including factors associated with the cell culture such as density and gross cellular morphology (clustering, cell area, etc.). Other contributions to segmentation errors were due to artifacts/contamination, registry errors between fluorescence and brightfield images, and low contrast in brightfield images. Interestingly, the authors note that there is a point of diminishing returns on the number of

training parameters versus improvements in performance. Moreover, they address a key issue in the field: the size of the training set needed to maximize the performance, noting that the ground truth data set with thousands of annotated images and computation needed for CNN is expensive. Surprisingly, they show that only 16 images with data augmentation were enough to achieve 95% of the performance of the full data set of more than 2000 images for all models. Another notable observation was the comparison between performance with transferred learning of models with one cell line versus the performance of transferred learning of models with multiple cell lines. Bootstrapping the training sets for transferred learning was optimal when the models were training across a variety of cell lines versus a single. Based on the overall performances of the models, the authors offer their own model, PPU-Net, to reflect the components from the other that contributed to maximum performance while remaining somewhat practical to deploy. The reader will need to judge for himself or herself what CNN model makes the most sense for his or her applications. However, this article highlights key considerations worth taking.

The final article in this Special Collection of *SLAS Discovery* takes the application of AI to an industrial scale for addressing improved efficiency of drug discovery. Allen and Nilsson<sup>5</sup> propose the combination of automation, biological readouts (imaging in this case) and machine learning to generate a discovery engine based on “de-specialized” outputs to inform on disease biology. Rather than take a “fit for purpose” approach for individual projects, where a priori hypotheses specific to assay readouts around disease mechanism are the foundations of entries and screens for drug discovery pipelines, the authors propose taking a general approach in which a high-throughput process designed to be applied across indications can be used to drive drug discovery. Instead of relying on upstream scholarship and the “hero scientist” for portfolio substrate, the process itself can be used to generate annotation for human biology, which can be used to mine for perturbations relevant to disease/disease phenotypes. The authors suggest that such a platform could supercharge the drug discovery process and continue to improve efficiency by learning from itself and adapting as data density increases. This approach becomes the discovery engine to “map every more biological relationships within a high-dimensional model of human cellular biology.” The authors, who are employees of Recursion Pharmaceuticals, describe Recursion’s model to applying this approach. This brief yet thought-provoking perspective offers a glimpse into an operation designed to take a data-driven industrialized approach at drug discovery. The successes are still to be determined, and while the early applications of this approach have been in the context of repurposing drugs for genetic diseases, it is exciting to think about taking this de-specialized approach to more widespread, complex insidious diseases, such as neurodegenerative disease, diabetes, or cancer, to identify new opportunities for treating common diseases.

Moreover, it will be very interesting to see if this approach indeed works to reduce the cost and (more importantly) the speed of drug discovery.

The effects of the COVID pandemic have challenged all of us and our productivity yet inspired us to find emerging solutions as a result of it. This Special Edition of *SLAS Discovery* in collaboration with SBI2 reflects this reality in its content. We are proud to print two articles addressing unmet needs associated with COVID. In addition, we feel that the other articles represent key advances their respective fields, specifically in prostate cancer and deep learning. Finally, the perspective from the Recursion team offers some food for thought around our approach to drug discovery. While we adapt our lives to a new normal, is it also time for us to take a new approach to our discovery process? Can we think in terms of “unmet need” and “de-specialization” to build a pipeline? If there is a common denominator among these articles in this Special Edition of *SLAS*, it would be the improvement of efficiency to get to generate meaningful data for drug discovery: from using high-content screening data to inform on putative pathways associated with lung disease, to simple time-saving solutions for measuring endothelial barrier function with spatial context preserved, to maximizing the readouts of precious cells grown in biologically relevant systems, to identifying practical solutions and managing expectations around deep learning deployment, to reinventing the drug discovery engine for efficiency—an industrialization of the discovery process.

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### References

1. Marwick, J. A.; Elliott, R. J. R.; Longden J.; et al. Application of a High-Content Screening Assay Utilizing Primary Human Lung Fibroblast to Identify Antifibrotic Drugs for Rapid Repurposing in COVID-19 Patients. *SLAS Discovery*. **2021**, *26*, 1091–1106.
2. Dubrovskiy, O.; Hasten, E.; Dudek, S. M.; et al. Development of an Image-Based HCS Compatible Method for Endothelial Barrier Function Assessment. *SLAS Discovery*. **2021**, *26*, 1079–1090.
3. Choo, N.; Ramm, S.; Luu, J.; et al. High-Throughput Imaging for Drug Screening of 3D Prostate Cancer Organoids. *SLAS Discovery*. **2021**, *26*, 1107–1124.
4. Ali, M. A. S.; Misko, O.; Solumaa, S.; et al. Evaluating Very Deep Convolutional Neural Networks for Nucleus Segmentation from Brightfield Cell Microscopy Images. *SLAS Discovery*. **2021**, *26*, 1125–1137.
5. Allen, A.; Nilsson, L. The Drug Factory: Industrializing How New Drugs Are Found. *SLAS Discovery*. **2021**, *26*, 1076–1078.