Hello and welcome to this Science/AAAS webinar. I'm Sean Sanders, commercial editor and webinar editor at Science.

Today's webinar, talks, and discussion will revolve around high-content analysis or HCA, a technique that has become widely adopted in both academic research and drug discovery laboratories. Image acquisition and analysis technologies have evolved rapidly in the last decade and are now being applied to investigating a great variety of biological systems from 2D and 3D cell cultures to small organisms and tissue sections.

This webinar will discuss the pros and cons of the most commonly used imaging methods across a range of applications and highlight recent advances, which promise to increase the power of HCA to answer more complex research questions and to yield deeper insights to advance the drug discovery process.

I'm delighted to welcome to studio today three preeminent scientists who will be helping us delve into this topic. To my left is Dr. Hakim Djaballah from Memorial Sloan-Kettering Cancer Center in New York. Next to him is Dr. D. Lansing Taylor from the University of Pittsburgh in Pennsylvania. Finally, we have Dr. Nick Thomas from GE Healthcare's campus in Cardiff, Wales. It's a pleasure to have you all here today. Thanks for coming.

Dr. Hakim Djaballah: Thank you, Sean.

Dr. D. Lansing Taylor: My pleasure.

Sean Sanders: As usual, each of our speakers is going to give a short presentation, after which we will have a Q&A session, during which the panel will address questions submitted by you, the live audience.

Note that you can resize or hide any of the windows in your viewing console. The widgets at the bottom of the console control what you see. Click on these to see the speaker bios or additional information.
about the technologies related to today's discussion, or to download a PDF of the slides.

If you're joining us live, you can submit a question to the panel at any time by typing it into box on the bottom left of your viewing console and clicking the submit button. If you can't see this box, just click the red Q&A widget at the bottom of the screen. Please remember to keep your questions short and to the point as this will give them the best chance of being put to our panel.

You can also log in to your Facebook, Twitter, or LinkedIn accounts during the webinar and post updates or send tweets about the event, just click the relevant widgets on the bottom of the screen. For tweets, you can add the hash tag, #sciencewebinar.

Finally, thank you to GE Healthcare for their sponsorship of today's webinar.

---

**Slide 2**

Now, it gives me great pleasure to introduce our first speaker for this webinar, Dr. Hakim Djaballah. Dr. Djaballah was recruited to the Memorial Sloan-Kettering Cancer Center in 2003 to set up and direct the High Throughput Screening Core Facility, a drug discovery laboratory involved in both chemical and RNAi screening. Dr. Djaballah has significant industrial experience in preclinical drug discovery, gained over the years in the pharmaceutical and biotechnology industry. He has been involved in developing assays and screening targets in various therapeutic areas, including antibacterials, antivirals, antifungals, diabetes, CNS, cardiovascular, oncology, and inflammation. He was the recipient of the 2007 Robots and Vision User Recognition Award sponsored by the Robotic Industries Association and the Automated Imaging Association. Dr. Djaballah, welcome to you.

---

**Slide 3**

**Dr. Hakim Djaballah:** Great. Thank you, Sean. It’s my pleasure to be here and I thought what I will focus today is on work that’s currently being done in my laboratory using different aspects of high-content analysis.

---

**Slide 4**

Before I get into that, I want just to take you briefly through memory lane and conceptually put forward the idea that high-content analysis was actually performed back in the ’40s and the ’50s. And here, an example of a true high-content assay being performed in
the 1950s where a cell-based assay is being carried out to what looks like microtiter plates instead of micro and looking at pH change whether the cell is alive or dead in the presence of compound.

The middle picture reflects one way of looking at the data through a light box and seeing the gray versus pinkish colors of the cells in the right panel where you do confirmation using a simple light microscope. So, you can imagine all these techniques being put into a box where the box contains a microscope of sort and a camera of sort and that alleviate the sheer use of multiple individuals performing the work, especially when you're gearing towards screening hundreds of thousands or millions of compounds.

**Slide 5**

My lab has three imaging systems, the IN Cell 1000, 2000, and 3000 and they do different things. The 1000 can take four fields that are fixed within a 384-well with an acquisition time of 16 seconds. The 3000 is our favorite instrument. It's a laser confocal microscope that can take up to nine fields, beautiful, crisp high-resolution images and the penalty of an acquisition time of 31 seconds. More recently, the IN Cell 2000 with a large camera where we can do full acquisition of a whole well at 4x magnification with an acquisition time of 4 seconds. So, that puts this into perspectives with the following examples I want to just share with you today.

**[0:05:22]**

**Slide 6**

The first one has been a problem that we tried to solve in the past five or six years and that is how do you a real live assay looking at caspase activation inside cells, looking at an induction of apoptotic effect. Not necessarily inducing caspases and then no apoptotic effect happens, but following all the way to the cell death.

And we borrowed the dye from a company called Biotium called NucView that has been successfully used in flow cytometry where the dye resides mainly in the cytosol. And upon of activation of caspases through DEVD sequence, the dye makes it to the nucleus and intercalate with the DNA and it goes from a quiet fluorescence to bright green fluorescence.

**Slide 7**

To validate the system, our tools in the lab are an isogenic pair of HeLA cells, the N10 and the B5 and the only difference between them is the B5 over-expresses Bcl-XL compared to the N10. And
when you induce apoptosis with Doxorubucin, you wipe out most of the cells or Bcl-XL is not there. You have some residual cells and that really highlights the heterogeneity of cell population in any given system that you study. You have to watch out for this one because outliers can skew your data. And in the presence of Bcl-XL, we attenuate somehow the induction of apoptosis.

Slide 8

The beauty about this, we can quantify it at 96 hours. We can see that here the example is using siRNA. Cell death siRNA will induce cell death and we can see at 96 hours, we see a huge induction of apoptosis versus no siRNA or a sequence against GFP. In the presence of Bcl-XL, we completely attenuate that. So this is one way to validate your system.

Slide 9

And we can confirm this by doing a nuclei count and this is live measure of nuclei count. We don't fix and stain here. This is live using the DRAQ8 dye where the cell can sustain the DRAQ8 for about 20 minutes or so. And you can see by 96 hours in the presence of the siRNA cell death, we completely wipe out the cells both in the presence or the absence of the Bcl-XL. And this kind of start highlighting the advantages of high-content analysis where you can get multi-parametric. Sometimes, the induction of apoptosis does not necessarily result in death.

Slide 10

We further validated this using another system that's part of an ongoing program we have at Sloan-Kettering and that's non-small cell lung cancer where we have been fortunate in having great tyrosine kinase inhibitors such as Tarceva and Iressa, etc. But as soon as the cell that you're targeting has a KRAS mutation, you're in trouble.

Here, what we distinguished between the cell line 2030 versus 3255, we can live assay measure the induction of apoptosis in the presence of Tarceva. Within 6 to 18 hours, we can start huge induction of apoptosis and by 96 hours, all the cells are gone. This is on the right-hand panel. On the left panel, when we deal with a cell line that has a KRAS mutation, in this case, it's a G12C mutation, we hardly induced any apoptosis, that's all a background noise and at 96 hours, most of the cells are healthy, they're still there. This has constituted a very nice way for us now to adopt this technology into HTS and we have run multiple screens looking at the induction of apoptosis as a
function of time. So having a really good microscope that has a fast acquisition time as well as an automated system, robotics or such that can move the plates back in and out will allow you to do that successfully.

**Slide 11**

My second example within the same terrain of live cells is the idea conceptually of how would you screen for EFGR inhibitors within the cell. Many people have looked at reporter technologies, either using the DiscoveRX technology or the Promega technologies, the luciferase readout or beta-gal readout that typically those have a huge problem because of the engineering of the cell lines. The reporter technology is great, it's fantastic, but it has to be within concept and if it's not within concept and the context that you're looking for, it may lead to read the artifacts.

Here, we decided to work slightly differently and this project has been in collaboration with Sigma-Aldrich and the idea is why don't we build domain-based biosensors, in this case, labeling Grb-2 for the SH2 domain with GFP and introducing the construct into an A549 cell. And the idea, therefore, if you add EGF to the cells, you should see activation of EGFR, dimerization, internalization and such. And the lower right panels shows just that in this data from Dmitry Malkov at Sigma.

And why we are very interested in this, on the left-hand panel, the lower panel, work of William Pao and [Vincent] Miller has identified really interesting point mutations in patient populations that don't respond very well to inhibitors of EGFR. And we thought this would give us a great opportunity to start asking the question of whether we can find better inhibitors for EGFR not necessarily kinase inhibitors, but they can be upstream or downstream of the signaling.

**Slide 12**

We developed again an assay. These are images from the IN Cell 2000 at 40x magnification where without any EGF; most of the signal is diffused around the cells. Again, heterogeneous cell population, you always get outliers as you can see in three or four cells in there. And the presence of EGF as a function of time, you will induce granule formation and we can easily form this measure and follow the size of granule, both as a function of cell density as well as EGF and a function of time.

**Slide 13**
Once we have optimized an assay, the readout looks something like this. Here we just used gefitinib as a proof of concept that we can inhibit granule formation with IC50s of 350 nanomolar and this only works in the presence of EGF. If you don't have EGF in there, there is no signal to measure and that's kind of as to the validity of the assay we are proposing. We can further validate most of the EGFR inhibitors by picking them up as inhibitors of the granule formation in the lower panel as well as they only kill the two cell lines in the far right, the E3255 and 4011 that are addicted to EGFR signaling and very sensitive to EGFR inhibitors.

Slide 14

We performed the pilot screen against the LOPAC library from Sigma and the plot here looks at the inhibition of granule formation on the Y axis and the viability on the X axis. Once we eliminate those specific cytotoxic compounds, we kind of identified inhibitors, which is interesting for us. But what was interesting from this screen is the identification of these activators. So, we identified a bunch of activators there and I'll get to those in a minute. You know, within our hypothesis, they didn't really make sense because we wanted only to identify inhibitors.

The panel of the six images on the right-hand side really depicts the beauty of high-content screening; a picture says a thousand words. And these are all inhibitors so I just put there six. As you can see, the top two are artifacts. The one top on the left is a complete artifact and it's a beautiful phenotype where this particular compound completely shrinks the cytosol of the cell. The beauty of high-content analysis is you can see that. The same way as the one to the left causes membrane ruffling and that membrane ruffling makes it very hard to see the actual granules. So, in that context, we see them as inhibitors, but they are actually artifacts. The two lower panels on the left-hand side really depict these activators that they peaked some of our interests and we're currently continuing working on those activators.

Slide 15

My second example is the beauty of confocal microscopy in screening, in high-throughput screening. And here we move from 2D to 3D cell culture formation. Here, we're looking at a model system where we oncogenetically transformed NIH3T3 cells with an oncogene then they start piling up. They start forming the cluster, the cocoon, the 3D. And that difference these stacks on the IN Cell 3000, we can see them at actually different levels so they form the
cocoon. And the concept of this assay is not to kill of them, but to find small molecules that will reverse the transformed phenotype. As it's becoming apparent, more and more of the therapies we have in the clinic don't work very well because the tumors are actually cocoons and part of these cocoons are very resilient to any chemotherapeutic agents you have.

The advantage of the whole well imaging here gives us the beautiful perspectives of how the cocoon grows in wells, in 384 wells. This is a small environment; it's 2 mm x 2 mm. And as you can see through all these different examples, they like to grow on the outer of the wells. Never we see them in the middle.

In the bottom rows of wells, there are six images there, that's what happens when the cocoons get too big, they detach and when you fix and stain the cells, you just wash them off. So here, to develop an assay that would look at 3D cell formation without any additional matrix for these cells to grow because that potentially can constitute a lot of problems in terms of image acquisition and actually analysis. But here without that, you need to get the right cell density that will allow you to see and measure these cocoons.

Having done all that, now we need to develop methodology to analyze the data. In the panel here, the normal phenotype as you can see the nuclei stained by Hoechst are dispersed. In the presence of the transformed phenotype when we have a large cocoon, they are very, very condensed around each other and they form a very good aggregation. When we do segmentation on the right-hand panel, it's actually black and white. So we can use this to define what we call the nuclear enrichment factor.

In the nuclear enrichment factor, the only thing it does is whether your nuclei that you have in the well are dispersed or they are clustered.

And using this, we end up with an assay that looks like. This is looking at the nuclear enrichment factor and the reduction of the nuclear enrichment factor in the presence of several drugs currently used in the clinic. Where it's a bright red is where we have reversed the phenotype and the bottom compounds of the panel is your
general cytosolic compounds, proteasome inhibitor, HSP90 inhibitor, etc. On the top panel, we only picked up inhibitors that affect PDGFR since our oncogene has PDGFR activity associated with it, but none of the EGFR inhibitors affected the PDGFR.

And here, highlights an example of where we can develop this kind of assays and we can straight away take them to the screening environment first looking at the approved drugs, looking for low hanging fruits and then later going for large screening campaigns identifying novel molecules.

**Slide 20**

Now, if put all this into perspective and looking at this retrospectively, in my opinion, high-content screening really started in the '40s before I was born and I think even before my parents met each other. And where it started making an impact in the whole of the industry in terms of microscopy imaging is really in 2000-2001. You know, we shouldn't forget the efforts that molecular devices had in terms of live cell screening looking at calcium flux using FLIPRs and that was in the mid to late '90s.

So, we've had a substantial evolution of this and by 2003, I think, the laser-scanning confocal microscopy opened so many doors to us. I don't have time today to go through some fantastic examples of screens that we have done. And by 2009, the whole well imaging I think both for cells in 2D, 3D, as well as for other model systems zebrafish, etc, would be very beneficial because that takes you away from the whole business of image stitching that gets a little bit tiresome.

Now, in consequence of all that, the technology is great, we can use it. At least my lab does it on a routine basis. It is becoming accepted both in academia and in industry. It causes a huge headache for us in the data explosion.

**Slide 21**

Now, my next slide kind of takes you through this tidal wave that at least at Sloan-Kettering we're going through it and it's not cheap, it's really expensive and very demanding. Low-content screening, you know, back in the '90s, you can screen a million compounds and at best you get files in the megabyte size. You know, the human genome, the whole human genome by the late '90s, early 2000s was in mega to gigabyte size.
When you look at high-content screening from the late '90s to now, we are in the tera and pentalbyte size and these were the value of a picture says a thousand words, but there is actually cost associated with it. That's not the cost to perform the screen, that's the cost to maintain the data and to move the data about. There are no computational solutions today. There are great computational solutions from Teradata and those companies that are very good at moving data from one warehouse to the other. But as an academic lab or an industrial lab, it's very, very expensive to keep a growth of terabytes per month regardless of what computational power you have.

In my predictions, the thing that will bypass that would be the ability I hope soon to be able to do predictive pathways by perturbing systems and trying to get the data on the fly rather than doing an endpoint analysis both from a system biology approach as well as from screening.

Now, my last slide here is perhaps send the question back to the audience today is whether high-content analysis has really introduced a paradigm shift in HTS. And if we do an assessment, perhaps you would agree with me that we are already there. I think the past five to six years have seen a huge shift both in the number of publications coming out using high-content analysis as well as adopters within the academia and industry doing it.

Traditionally, screening was done against targets that are recombinant or purified. We have limited successes with those. I mean we have to be honest in a certain perspective and say that HTS today, you know, has yielded certain things, but not really what we had in mind back then.

Cell-based reporter assays have been used in HTS. Some examples have been great, have been very successful. Luciferase technology has been fantastic, but live cell-based assays started to be performed in HTS and if you have the whole infrastructure around it, you'll be very, very successful in actually doing it.

So the low content aspects of this have been sufficient in industry thus far, but moving forward at least in complex systems such as oncology, you have to get into high-content screening because the gain of information there is just priceless. Automated microscopy
was introduced in the '90s. It has become the norm in most academic and pharma.

Now, a picture says a thousand words, but there's a caveat to that and you have to deal with the screening data explosion and its management.

And in conclusion, I think from our perspective, we have now generated enough data that says that high-content screening for things like RNAi screening will be the best readout. You know, we've done those screens using low-content assays. Others have published using luciferase as the reporter technology and we tend to see differences in data or at least nomination of actives within RNAi screens.

And I'll just stop there, Sean, and...

Sean Sanders: Great. Thanks so much, Dr. Djaballah.

Our second speaker for today is going to be Dr. D. Lansing Taylor. Dr. Taylor received his B.S. from the University of Maryland and his Ph.D. in cell biology from the State University of New York in Albany. He was a postdoctoral fellow in biophysics at the Woods Hole Marine Biological Laboratory. Dr. Taylor started his academic career at Harvard University before moving to Carnegie Mellon University as a Professor of Biological Sciences, where he started the Center for Fluorescence Research in the Biomedical Sciences. Dr. Taylor has founded a number of life science companies including Cellomics, Inc., which was sold in 2003 to Fisher Scientific and, most recently, Cernostics, a tissue systems biology diagnostics company. He is currently director of the University of Pittsburgh Drug Discovery Institute and professor of the Department of Computational and Systems Biology in the School of Medicine.

Dr. Taylor, welcome to you.

Thank you, Sean. It's a pleasure to take part in today's webinar. So, my goal today will be to give a projection of where we're going at the Drug Discovery Institute and how it may have an impact at both academic research and in drug discovery.
I want to start by discussing a couple of issues surrounding high-content analysis and high-content screening and then describe a couple of projects at the University of Pittsburgh Drug Discovery Institute, and then look to the future of important things that need to be developed. One is understanding heterogeneity of response of cells to drugs and creating integrated pathway mapping and cellular models.

**Slide 26**

So to begin with, semi-automated microscopy started more than 20 years ago and the upper left is an example of that, a multi-mode microscope that we used at Carnegie Mellon. Key characteristic was that eyepieces were still part of the system, although there was a lot of computerization. So, the investigator had to interact with it in order to change from multicolor fluorescence to ratio imaging to transmitted light microscopy. Over the last decade, microscope companies and other life sciences companies have developed semi-automated microscopes to do things like super resolution, fluorescence lifetime imaging, multiprobe microscopy.

And the other side of the fence of high-content analysis is fully automated microscopy, which we coined the term high-content screening a number of years ago. The key here is there are no eyepieces, that this is a fully automated process.

**[0:25:12]**

And of course there's tradeoffs between using semi-automated systems and automated systems, but in fact, we believe there is a continuum here and most projects can use both.

**Slide 27**

High-content screening has evolved quite nicely over the last decade. The focus is still on having the best biology in each well. Standards still lag behind what has been done in flow cytometry, but some progress is being made. But there have been great developments in the instrumentation, new classes of reagents for both live cells and fixed endpoint assays.

Imaging software has become more sophisticated as has the data management and analysis programs. One of the things that's evolved nicely over the last few years is the use of machine learning and I think machine learning is going to play a critical role especially in drug discovery. The other thing is to integrate all of this direct cell based measurement into an integration with other omics, gene expression and proteomics using all of that information to build in
silico models, which can be, we hope, able to then suggest the next experiment to be done.

**Slide 28**

The tagline for the Drug Discovery Institute is novel chemistries and systems biology power discovery. And whenever possible, we try to understand the target molecules and its structure, but we also do a large number of phenotypic screens starting with cells, tissue engineered arrays, and experimental animals such as C. elegans and the zebrafish. In addition, we're beginning to translate some of that work to in vivo imaging in rodents.

A key element of what we're doing in the future that we see is the use and development of computational and systems biology tools in order to be able to extract the most information from these complex studies.

**Slide 29**

So, I'm going to give you a few examples. One is a zebrafish model of disease. Also, our evolving protein-protein interaction platform, our ability to begin looking at heterogeneity of drug effects, and then how do we use this data to do direct pathway mapping and modeling.

**Slide 30**

So live cell kinetics in both cells and the experimental animals is growing in importance. On the upper left, you can see a semi-automated microscope image of a one-day-old zebrafish embryo. Within this short period of time, most of the major organ systems have been developed. The fish is transparent, you can do transgenics, and so you can create models of disease.

On the right using an HCS platform, you can put one embryo in each well and with this engineered model you can do things. Like in this example, look for inhibitors of angiogenesis and when you find these so-called hits and validate them, you can go back to semi-automated microscopy in the lower left. And to look at hyper spectral information or in this case, high speed information, this is where we're looking directly at the rate of blood flow at high magnification and a high time resolution.

**Slide 31**

We're also building what we think is a very important class of target, that is protein-protein interactions. This has been a challenge for the
pharmaceutical industry because it's complicated. So, you start with roman number I, the selection of disease relevant protein-protein interactions. Then you select the optimal pairs based on known binding domains where you have the structure, and we're looking for hotspots between the two proteins. Then we do virtual screening using a program developed at the University of Pittsburgh called the AnchorQuery where we can look for an anchor amino acid in the pocket and then design outward and design and then synthesize compounds that best fit into that pocket. Then we can do validation by using protein-protein interaction biosensors and do high-content screening to look at both efficacy and early safety assessment.

**Slide 32**

Fixed end-point assays are still valuable. The key value of fixed end-point assay is the ability to do large-scale combinatorial profiles such as looking at a large number of different cells types with a large number of different compounds with dose-response curves where you literally can prepare the plates offline, stack them, and run them on high-content screening platforms.

[0:30:11] Presently, the standard is to use 2, 3, 4, or maybe 5 different fluorescence channels in an assay. And one of the things we're working on as a research project is extending this eight or more parameters that can be looked at in each cell using spectroscopic and chemical methods.

**Slide 33**

An essential challenge in biology is to understand how individual cells process information and respond to perturbations, for example drugs. However, much of our knowledge is based on ensemble or population average measurements such as gene expression profiles. A basic assumption has been that ensemble averages reflect the dominant biological mechanism of action within individual cells in the population.

So, on the lower left, in figure A, in a distribution that doesn't have any subsets within it, an average measurement maybe reflective of the whole. In example B, if you have a subset of a population within a larger population, measuring the mean is not going to be helpful. In the extreme, in the example C where you have two distinct subsets, in fact a mean value will be meaningless. Of course, this is one of the key elements of high-content screening and high-content analysis is by doing cell-by-cell measurements, you can look at heterogeneities.
Slide 34

Well, there are multiple sources of cellular heterogeneity in the cell population. One simple one is genetics. So, for example in a tumor, you can have cells that have evolved genetically so they actually are distinct. But also there's non-genetic heterogeneity even in clonal populations of cells. One example is extrinsic and for example again in the tumor, the cell may respond differently when it's sitting next to different stromal cells or immune cells within the tumor.

There's also intrinsic heterogeneity and there's two forms of that. One is macroheterogeneity and that again is when you have two distinct subpopulations and a good example of that is in stem cell differentiation. But even within what looks like a homogeneous population, you can have microheterogeneity, population noise is having some small but discrete subsets within the population and temporal noise, which a big part of that is gene expression.

Slide 35

So one of the things we did was to begin looking at heterogeneity. We went back to a large dataset we did a number of years ago measuring on the left 12 different parameters in a breast cancer cell line. In the middle figure, you can do what we call cell maps and actually look on a 2 x 2 basis of two of the parameters, in this case, nuclear condensation and DNA content. As you go from the left to right, as you increase the drug concentration, in this case, it was for a microtubule modulating compound, laulimalide, you can see in that last column that only 74% of the cells actually are the G2M block, which we would have predicted. But a large fraction are still in 2n and in fact a small percentage are exhibiting apoptosis.

Heterogeneity is a serious real biological process just not "simple noise." On the right, we've started looking at these large-scale measurements and this is an example where we looked at 12 feature space and projected the 12 features that we measured on the two principal components. And what we learned by doing this, there were at least five different subpopulations and this was at one dose in this cell model. So one of the questions we're now addressing is how do subpopulations vary with dose response? Do all of those parameters shift at the same way or are they forming yet further subsets?

Slide 36
Well, heterogeneity is important all the way to patient level. I borrowed this image from Cernostics. It’s a cartoon of a tumor where, of course, there are normal cells, there are cancer stem cells. The yellow cells are the cancer cells, some of which have changed genetically. Others are the same and in fact, they can respond differently because of where they’re residing; some cells sitting next to immune cells, other cells sitting next to other stromal cells. We're dissecting this in an evolutionary way. We've started with single cell type models and then now we're building tumor microenvironments.

**Slide 37**

Well, if we’re able to extend the plexing to hyperplexing, that would drive our profiling. It would give us a better handle on subpopulation analysis and enable us to do more sophisticated direct pathway mapping. The key to this will be using and developing advanced informatics tools, being able to analyze cell heterogeneity to look at the correlation between genomics and proteomics data and then do a variety of modeling approaches in direct pathway mapping, and ultimately using machine vision to guide the next experiments.

[0:35:42]

The results of all of this, we hope to be able better define drug mechanism of action, to define heterogeneity and how we can deal with it, to have direct cellular network knowledge, and to then have new targets for discovery and potentially new diagnostic markers.

**Slide 38**

With all of this cellular profiling information, I think, the key will be to integrate that with the large amount of omics information we have from gene expression and proteomics and look carefully at drug effects. And then use all of that data and how it affects the networks to generate and refine detailed models of specific pathways using rule-based techniques. And we think this will allow us to design better assays and to do better drug discovery.

**Slide 39**

With that, I'd just like to acknowledge the key people that worked on these different programs.

Sean Sanders: Great. Thank you so much, Dr. Taylor.

**Slide 40 to Slide 41**

And our final speaker for today is Dr. Nick Thomas. Dr. Thomas is principal scientist in Cell Technologies at GE Healthcare based in Cardiff, Wales. He has a B.S. in biochemistry from the University of Glasgow and a Ph.D. from the University of Wales College of
Medicine. In a 27-year career starting with Amersham International, he has held a number of positions in operations, marketing, and research and development. He is the inventor or co-inventor on over 60 patents covering a wide range of technologies including microfabrication, molecular and cellular sensors, and cellular imaging. Dr. Thomas has worked on the development of cellular analysis instrumentation, software, and reagents for the past decade and has published a number of review papers and book chapters on the subject.

Dr. Thomas, welcome.

Dr. Nick Thomas: Thank you, Sean. It's my great pleasure to be here. So, the two previous speakers have given great overviews of a range of HCA applications so I thought what I'd do for this section is focus down on one particular example, and what I'm going to talk about is the cardiotoxicity assay and multi-parameter assay.

So, although, this is a toxicity assay, I should emphasize on what the kind of approach is and the data visualization techniques I'm going to show you can be applied to a wide range of HCA studies.

Slide 42

So, as I said, this is a toxicology study and we've done it using a recently developed cell model that we've developed in Cardiff. So, these are human cardiomyocytes derived from human embryonic stem cells. And what you're looking at now on the screen, the two pictures taken IN Cell Analyzer are these cardiomyocytes stained for the troponin I component of the myofibrils on the left-hand side and Connexin 43 gap junctions on the right-hand side. So, these are the cells we've used in the assay.

Slide 43

And this is a study, an ongoing study, that we're doing with Genentech in California and what we're doing is taking a library, a series of libraries. And this is the first library I'm going to talk to you about today of compounds in a variety of classes including, in this particular library, quite a large number of tyrosine kinase inhibitors and other compounds. All of these are compounds focused with an oncology focus.

I should also mention, we've included two controls; positive control amiodarone, which is antiarrhythmic, but has known cardiotoxicity and nifedipine as a negative control.
So, this is the assay. It's a live cell assay. It's run in 384-well plates using the Cyteva cardiomyocytes. We exposed the cells to the various doses in a 7-dose response curve and we gave the cells the drugs for 24, 48, and 72 hours. So, we're getting both acute and chronic exposure. At the end of that exposure period, the cells are stained with four different fluors; Hoechst, tetramethyl rhodamine, Fluor 4 and TOTO-3 for membrane integrity.

Then the 384-well plates are imaged on IN Cell Analyzer 2000 and here, we're using a 40x objective and taken advantage of the large 4.2 megapixel sensor on the instrument. So, we get good magnification, large scale coverage, and high resolution. That allows us then to take those images into IN Cell Investigator, which is our image analysis software, and abstract a number of parameters from those images, 19 different parameters in this particular study. Those parameters then are automatically ported using the dynamic link between investigator and Spotfire to do the visualizations that I'm going to show you today.

So one final thing I should say on this study. We received all the compounds in what I showed you on the previous slide from Genentech and those compounds were blinded until we've completed the data analysis that I'm going to show you just a small fraction of today.

So, this is the assay just to give you some idea of what we're looking for in the assay. So, this is a control image or an image from control cells. So, what we're looking at here are cells stained for DNA in blue, mitochondria in red, and calcium homeostasis effectively in green. So, if you concentrate on the red part of the image, we're looking at mitochondria. So, what you can see here is a typical filamentous structure of the mitochondria throughout the cells and that's what you would expect to see in healthy mature cardiomyocytes. They're a very energy demanding cell; they have a lot of mitochondria to make ATP.

When we treat those cells with amiodarone, this is a typical example of what you'd see and I think the most obvious example of the effects of amiodarone is the coalescence and condensation of that mitochondrial structure within the cells. And I've put a couple of
ours on here so you can see. So basically, from a filamentous structure, that is collapsing down and we can take measurements in that channel and we're also taking measurements both in the other three channels in the assay.

**Slide 47**

Those were one of the positive controls, here's one of the compounds in the assay vatalanib and you can see as we go across the slide from left to right, the effect of increasing concentrations of that compound. And what we're seeing here is both changes in the red channel, in the mitochondria, and changes in the green channel since actually the increase in green that you can see in the images is due to the perturbation of the calcium homeostasis of the cells.

**Slide 48**

So, on the slide you're looking at now, this is the -- if you like the typical way that people would look at data from classical biochemistry experiments, you do a dose response, you plot those response curves. What I'm showing you here is only a fraction of the data because I'm showing you only 4 of the 19 parameters and I'm only showing you the 72-hour data point. Even so, it's very difficult to get your head around what's actually going on in this data looking at in conventional terms, which is why the tendency with high-content analysis is to use the direct linking from the analysis software into some form of visualization. This is illustrated on the next slide.

**Slide 49**

So, this is what we might call data profiling here using parallel axes plots to display the data from the 19 parameters. And you'll see along the bottom of the slide, I've color coded the areas of the grass according to whether the parameters are related to mitochondrial integrity, calcium homeostasis, or viability. So, what we're doing here just for anybody that's not familiar with this type of plot is we're taking 19 different parameters. Each of those parameters is plotted on one of the vertical axis on the plots and then the profile is formed by effectively joining the dots.

So, if we just start with looking at nifedipine that sit in the top left, you can see there on that plot, there are two profiles, one in pale blue, one in red. They're essentially super imposable indicating that there's no change in that profile from a low dose to a high dose. Similarly, below that one, tasocitinib, one of the test compounds,
again virtually super imposable profiles indicating no effect of tasocitinib across the dose range tested.

But in marked contrast down the bottom right is amiodarone, the positive control in the assay. So, we know that interferes with mitochondrial function. And you can see that very sharp peak, which I've shown with a purple arrow there in one of the mitochondrial integrity parameters, which happens to be form factor. Form factor describes the shape of the mitochondria. So, with nifedipine and tasocitinib that value was very low because you have a diverse spread filamentous structure. With amiodarone and similarly with sunitinib directly above, you get a massive increase in that form factor, which is basically because you're going from a long filamentous structure to a good dense mitochondrial architecture. So, that was just the -- I've selected four of the compounds.

Slide 50

Here is all of the data from the 72-hour time points. So basically, we're looking at profiles for all of the compounds in the study. I've marked nifedipine in green and amiodarone in the bottom right in red. So, you can see a variety of effects across the different compounds.

Slide 51

So, what I'm doing here, I've highlighted nifedipine, mubritinib, and tasocitinib and you can see those three compounds have pretty well conserved profiles across the dose range from 0.05 micromolar to 1 micromolar. In contrast, the compounds highlighted in red have similar profiles to amiodarone in that they changed markedly with those.

Slide 52

So what does that tell us? It tells us that those compounds are changing the phenotype of the cell in a toxic way, but what we can also look at is how the different compounds affect the profiles. And what I'm showing here is a range of compounds with their range of mechanism of actions. From pazopanib on the left-hand side, which has very little effect, maybe a slight change is significant in calcium integrity, through sunitinib, imatinib, and lapatinib, which they have very marked effects both on calcium integrity of the cells, calcium handling, and the mitochondrial function of the cells.

Slide 53
So, that's an observational way of looking at the data. What I want to show you quickly now is an unbiased statistical base to look at the data. So, this is using self-organizing maps within Spotfire software and what the Spotfire software is doing is clustering those profiles together on the basis of similarity. So, these are the profiles at the lower dose, 1.24 micromolar, and you can see most of the profiles from the compounds cluster towards the bottom of the plot in clusters 1 and 2.

**Slide 54**

If we look at 33 micromolar, you'll see now in cluster 1 in the bottom left, there's only two compounds left, tasocitinib and the negative control nifedipine where the other compounds have moved right or up and a lot of compounds are now in cluster 6 at the top left.

**Slide 55**

So if we take those clusters and look at the data in heat maps, you can see distinct patterns going from tasocitinib, which is in cluster 1, through to the compounds, which I've ranged along the bottom of the slide, sunitinib, vandetanib, and amiodarone, which you can see have very similar looking heat maps.

**Slide 56**

So, that allows us then to take the compound list that we started with and put them into a rank order based on the cluster that they end up into. From cluster 1, which contains the negative control nifedipine and tasocitinib, through to cluster 6, which includes the positive control, and interestingly a number of compounds including imatinib, lapatinib and others, which have reported cardiotoxicity in the literature. And there I'll finish.

**Slide 57**

Sean Sanders: Great. Thank you so much, Dr. Thomas. We're going to move right on to our Q&A after I thank all our speakers for their excellent presentations.

We're going to look at some of the questions submitted by our online viewers. Just a quick reminder to those of you watching us live that you can submit your questions by typing them into the text box and clicking the submit button. If you don't see that box, just click the red Q&A icon and it should appear.

We're going to start off quickly with a polling question for our audience and this asks, which of these choices best describes your
current work environment? Is it research, screening, drug discovery, diagnostics, or other? And while we have that question up and the audience is taking a look at it, I'm going to put a question to you, our panel. What applications do you see as the sweet spots for HCA technologies and what applications are challenging currently for HCA? So maybe we'll start with Dr. Djaballah.

Dr. Hakim Djaballah: Good question. I think our sweet spots for HCA currently is anything to do, at least within oncology that I can speak to with some degree of confidence because that's what we do at Sloan-Kettering is most of our assays have moved to cell-based assays and we try to take images of as many things as possible. Even in the context of we have a target-based assay whether it's been engineered in the cell as I showed earlier and used as a transformed phenotype for an oncogene as such, etc. And in my lab, we see more and more of that portfolio with more of an HCA, HCS-rich than target based and I think that's becoming the norm in other places as well.

For the second part of the question, it's a little bit challenging because I think we want to be able to run live assays over a longer period of time. We want to be able to have the ability to look at the same cells within the same positioning over a long period of time. And that I don't think it's there yet though we try to do it, but we are limited to the number of hours or number of days you could leave those cells there before you start getting into issues of crowding into things that are not really a problem for people doing, you know, classical, typical one-on-one cytology on slides.

You know, the 384-well plate has a limited surface area, it's 2 mm² and when these things grow especially cancer cells, they pile on top of each other. I think being able to have the resolution sensitivity from the get-go to measure fewer cells than when you have a huge problem of crowdedness over a period of time. I think that remains, in my opinion, a huge challenge today.

Right. Dr. Taylor?

Dr. D. Lansing Taylor: I think high-content analysis has evolved quite dramatically over the last 13 or 14 years and certainly, the field started out with a focus on fixed end-point assays looking at just one or two parameters in a single target-oriented assay. Today, with the developments of all of the related technologies, we can do live cells, research experimental animals generating literally massive amounts of data. So, I think it's
the direct measurements of functions whether it's in cells or experimental animals, which is the sweet spot.

In terms of the challenge, because we're generating this massive amount of data and we have massive amounts of data from gene expression profiling and proteomics, how do we integrate using computational methods and systems biology tools all of that data to really understand pathways and mechanisms of action within cells.

Sean Sanders: Right. And, you, Dr. Thomas?

Dr. Nick Thomas: I think there are a number of well understood sweet spots and perhaps one that most people would think of are kind of like the one I illustrated in early stage toxicology, large-scale RNAi screening. But I think the important issue to consider is that that sweet spot is continuing to expand these days. It's driven by the experience of users. They learn new techniques; they're looking how high-content analysis can be applied. So, you know, if we go back 10 or a dozen years or so, high-content analysis looked at applications, which were basically the same type of applications that were done with research scopes.

Sean Sanders: Uh-hum.

Dr. Nick Thomas: Nowadays, those applications are still done, they're becoming much more sophisticated in a multi-parameter way and multispectral ways Lance alluded to. But I think the demands of the community have driven the development of the instrumentation. They've driven the development of cell models and software. Those still remain challenges and that's kind of the thing that keeps me interested in my day job, how do we enable applications where high-content is the ideal and the logical and the sensible way to approach a particular scientific question.

Sean Sanders: Excellent. So, I'm going to have those poll results put up for our audience and as expected, the vast majority are in research environments, only about 3% in screening, and about almost 9% in drug discovery. So based on this, the next question I have for you is confocal imaging is currently very widely used in research microscopy, what do you see as its value in HCA? So, why don't we start with Dr. Taylor?

Dr. D. Lansing Taylor: Well, I think it's important as the biology has gotten more sophisticated for investigation. Building 3D models for example,
tumor microenvironments are looking at stem cell differentiation where there's multiple layers of cells or looking at an experimental research animal like C. elegans or zebrafish, the ability to optically look at a single plane at a time to make measurements in these complex biologies is central going forward.

Sean Sanders: Uh-hum. Excellent. Dr. Thomas?

Dr. Nick Thomas: I think confocality is an important technique. I think the first major thing to remember is that you can't change the laws of physics. So, there are certain limitations to any confocality technique. Some confocality techniques particularly those used in high-end research microscopes work fundamentally by throwing away light. So, that's good if you want to have the best quality image, that's not so good if you want to run at very high throughput. So right now, we're taking an approach where we're trying to be flexible in our systems and having systems that cannot please all of the people all of the time, that's generally impossible.

Sean Sanders: Uh-hum.

Dr. Nick Thomas: Equally impossible is breaking the laws of physics. But basically, it's a question both I think of development of instrumentation but also people learning the appropriate time to use a particular technique.

[0:55:00]

Sean Sanders: Excellent. So, I'm going to put up another quick polling question for our audience and this asks what are the most pressing limitations that you face in your current microscopy efforts? And these include hardware or instrumentation, software, model systems, reagents, or other. And while people are taking a look at that one and answering it -- by the way, to answer the questions, just click the relevant box on your slide viewer and click submit and that will get through to us.

So, I'll put this question to Dr. Djaballah. Do you believe that HCA users will adopt 3D culture systems as cell models, for example hepatocytes for stem cell development?

Dr. Hakim Djaballah: It's a very good question. I think this whole business of 3D cell modeling is kind of getting out of hand. I get fascinated by conferences that have a topic on physiologically relevant cells for screening and in my opinion that's all mumbo jumbo. I think as soon as you take a cell from on organism, it's never going to be
physiologically relevant. We are really taking care of surrogate assays.

Having said that, a human being is not a mouse and a mouse is not a zebrafish, and I think we have to be realistic about that. Many years ago when I was at school, one of the exam questions was what was the difference between a mouse and an elephant in term of DNA content, and to my surprise all of us got it wrong. We thought that the elephant had more DNA because they were bigger in size.

You know, having said that, I think the tools that are available today to address certain questions in terms of drug discovery are not sufficient and that's very apparent because we really are not very good at drug discovery and putting new drugs on the clinic, in the market for people to use and to treat disease. We have been fascinated by sidetracks of physiologically relevant cell-based assays.

You know, you take a normal cell, you put it in a Petri dish or in a 384-well plate, the gap junctions are active, there's contact growth inhibition, they form beautiful monolayer and they stop.

Sean Sanders: Uh-hum.

Dr. D. Lansing Taylor: You take a cancer cell, as long as it has food, it just keeps going, it would never stop. Now, when we get into the 3D systems where now we have something in the shape of a cocoon, back to what Nick was alluding to, the laws of physics are different in a microtiter plate; oxygen dispersion, CO₂ dispersion, food dispersion, and the necrotic events that are happening inside the cocoon that's being formed. So, at the end of the day, what do you measure.

Sean Sanders: Uh-hum.

Dr. D. Lansing Taylor: People have come with ways of this nano fabrication where you put the scaffold in there for these cells to grow on. And I think though in theory, they're great. If you are in engineering department, that's fantastic, but for drug discovery screening, I don't think they have much of that relevance anyways because most of our microscopes don't have that breadth. Even in confocal modality, you have a much wider Z-stack to be able to start stacking into that microfabrication to see what the cells are actually doing.

So, I am pretty satisfied with cells in 2D. I think we've been successful with those. We've made a tremendous advance. But when
it comes to 3D, my bias has been to oncogenic transformation, how
do you take a normal cell and you introduce an oncogene and
suddenly it starts doing things differently and try to understand that.
But back to, you know, concluding the question by the point of
physiologically relevant assays, I think that's just a myth and I don't
believe in it for one second with the reasons I've just said earlier.

Sean Sanders: Uh-hum. Any other comments from the panel?

Dr. D. Lansing Taylor: Sure.

Sean Sanders: Dr. Taylor?

Dr. D. Lansing Taylor: I would pleasantly totally disagree with my colleague.

Sean Sanders: [Laughs]

Dr. D. Lansing Taylor: I think the history of cell analysis has shown us that the environment
matters. People used to plate cells directly on glass or plastic and
then matrix materials were used and there was a change in the
responsiveness given that simple 2D environmental change. It's very
clear now that three-dimensional relationships and connections have
a big effect.

Now, 3D models are much more complicated and doing a high-
throughput screen it's pretty tough to do that. But in terms of really
understanding the biology, I think it's critical to have these more
complex models at least as secondary screens.

The last comment I'll make is just a comment that Nick made about
breaking the laws of physics, which we usually can't do, but in optics
we thought the limit of resolution was defined by the Abbe
principles. And with super resolution using structured illumination,
we had to change the laws of physics.

Dr. Nick Thomas: I think that's bending the laws, is it not, literally?

Sean Sanders: [Laughs]

Dr. D. Lansing Taylor: It's getting around the laws.

Dr. Nick Thomas: Yes, not breaking. Law avoidance or...? I wouldn’t go there.

[1:00:06]
[Laughter]

Sean Sanders: So, I'm going to -- we're coming to the end of our hour, but I did want to push out these poll results to our audience just for the last question and take a look at what we see. So, it seems like most people are having trouble with their hardware and instrumentation or at least these are the limitations. Quite a number, almost a third software and about 25% face limitations in model systems.

So let me ask you, Dr. Thomas. Do you see that this fits in with our experience? It seems like instrumentation and software are the big hurdles.

Dr. Nick Thomas: I think those are the ones that most people focus on, definitely.

Sean Sanders: Uh-hum.

Dr. Nick Thomas: I mean the instrumentation is the thing usually at the top of people's mind and then that links that software. I think that's where most of the development focus has been up to now, a little on the reagent side, but reagents tend to have been borrowed from other areas, flow cytometry. And interestingly, I think given the discussion we've just had, the cell model is usually the one that's given the least attention. So, people get very fixated on, you know, the number of objectives, the number of filters, the number of channels on the instrument, how good the software is, and then they focus all that power in the HeLa cell.

Sean Sanders: Right.

Dr. Nick Thomas: So, what we're trying to look at is to take a more holistic approach. I think I'm in the middle between these two gentlemen on where 3D models come in. I think they will be used in the future, but I think you have to be careful that you use something when you should do it not just because you can do it.

Sean Sanders: Uh-hum. Good point. Any other comments?

Dr. Hakim Djaballah: You know, I think in terms of instrumentation and hardware, software, and tools, I think people have to remember that this field is continuously growing and there is no solution that's going to fit all. You know, there are people the same with what Nick is saying that they want six colors to do an exercise that very well that the realization that in the cell population may be one or two cells out of
a hundred will actually respond. And if you can't see them, you blame it on the instrumentation. You know, my favorite example is non-homologous DNA ligation. You know, that event is poor in a cell. You know, at best it is 4%. If you try to do it in vitro, it's going to get really bad to find those cells.

In terms of reagents, I think there is a scope that, there are limitations not of the ability of the reagents but the ability to use them. We have to remember whether you have a matrix or cells grown in glass or in plastic becomes irrelevant. When you start doing immunocytochemistry on those plates, there are certain techniques you can't use. As an example, you can't use ethanol fixation. We can't do that in HTS. So, I have to be careful here and separate from the aspects of basic research. People can go to molecular cytology core facility same way we have at Sloan-Kettering and even folks in my group use that to do things on a one-to-one basis and you can go to 100x resolution. You can do six colors and you could get beautiful pictures. Those we'd never be able to do in HTS and they're not really needed.

And by the same token, I think, the model that you're going to study whether it's a cell line or a 2D or a 3D or 4D, it becomes very, very relevant to what's the question you are trying to ask. If the response of what you're trying to perturb is not greater than 10%, or 20%, or 30%, you'd be lost in the noise. And then it's going to be very hard to distinguish whether the tools that you have in terms of the probes that you're using versus the instrumentation versus the software can do segmentation this way.

Sean Sanders: Uh-hum. Right. Well, I think we're going to have to end there because we are unfortunately out of time for the Q&A session. Sincere thanks to our viewers for submitting your questions to the panel and apologies if we didn't get to all of them.

Before we close out this webinar, please do stay tuned for a couple of minutes. We're going to have a short presentation that will point you to some of the exciting resources related to today's webinar. So, Dr. Thomas, over to you.

Slide 58
Dr. Nick Thomas: Thank you, Sean. So I just thought I'd sum up at the end of this very interesting debate we've had here today and look at where we've come from and where we're going, I think, in high-content analysis from a personal perspective.
So I think, you know, what's the value of high-content analysis? I tried to summarize it here on this slide in I think two main areas: Getting deeper insight into biology whether that's in academia or in pharma, whether that's to find fundamental scientific questions or for drug discovery. But getting deeper insight by probing deeper into biology but doing it in context and in situ. And I think that's the real key for high-content analysis. Don't blow open your biology and try and dissect it molecule by molecule. Go into the cell and study it.

The other obvious one that everybody focuses is improved productivity, getting more samples through. And that has its place either taking imaging assays done on research scopes through into high-content analysis. Or as an alternative way to doing, tackling traditional biochemistry, as an alternative way of doing protein identification instead of doing Western blotting.

If you look at where we've come from, the field looking back over the past 10 years or so, this is basically my history in high-content analysis from our early days through to the present day. And one of the major things I think over the last few years is the increase in camera size and it's interesting that this is happening in the same way as it happens in hobby and professional photography. People want to cover more area without throwing away resolution. So, now, as Hakim has illustrated before, now we can do whole well imaging, now we can look at the patterns of cells, we can look at stem cell colonies. So, we can gather all that information that we can get from flow cytometry, but we get the spatial information as well.

So, this is kind of, you know, alluding to what we said before. Everybody focuses on the box, that's the term, focus on the box, look at the specifications of the image and on the software. But we're taking, as I said, a rather holistic approach and this is why we're starting to develop stem cell derived cell models. Because I've said before, we spent 10 years developing very sophisticated hardware and software and then apply those to a less sophisticated cellular system. So, what we're trying to do now is bring up the level of the cell models and bring up the level of reagents so everything is on an equal, sophisticated plane.
And that’s being driven by the users out there in the field, which are becoming, you know, large in numbers every year. People are moving, they’re translating their research assays. From confocal microscopes, they want to move out into HCA. So, we’re providing them with the means to do confocality and high-throughput. They want to do more demanding and complex 3D systems with sectioning again using confocality.

More in pharma and biotech industry, there’s the use of these instruments in GxP, GLP laboratories so the software has to be adapted to account for the rigorous requirements of those environments. And more and more people are using it. So, there’s a greater range of uses and they have different skill sets. They may have a microscopy background, they may not have so the software and the hardware and the interfaces have to take account of those things to allow cell biologists to actually use the instrument without a specialist to do it from.

Slide 63

Basically, just to sum up, this is where we are today in terms of our imaging platforms. The IN Cell Analyzer 2000 on the left-hand side, very flexible, modular lamp base system and the newly launched IN Cell Analyzer 6000, the high-end laser based confocal imaging. Now, the really important part about this instrument is again, we can't break the laws of physics, but we are attempting to bend them here by allowing you to have on the fly, adaptable confocality in different channels. So, this is your top end machine.

So, thank you, Sean.

Sean Sanders: Great. Thank you very much, Dr. Thomas.

Slide 64

And many thanks once again to all of our speakers: Dr. Hakim Djaballah from the Memorial Sloan-Kettering Cancer Center, Dr. D. Lansing Taylor from the University of Pittsburgh, and Dr. Nick Thomas from GE Healthcare.

Please go to the URL now at the bottom of your slide viewer to learn more about technologies related to today’s discussion and look out for more webinars from Science available at www.sciencemag.org/webinar. This particular webinar will be made available to view again as an on-demand presentation within approximately 48 hours.
We'd love to hear what you thought of the webinar, send us an email at the address now up in your slide viewer; webinar@aaas.org.

Again, thank you to our panel and to GE Healthcare for their kind sponsorship of today’s educational seminar.

Goodbye.

[1:09:05]  End of Audio