The value of adding RNA sequencing to precision biomarker strategy

Summary of a discussion between Joel Dudley and Radu Dobrin during the Precision Oncology Leaders' Summit on March 30, 2021.

Editor's Note: We had a chance to enjoy a virtual chat between two colleagues and collaborators, Dr. Joel Dudley (Tempus Labs) and Dr. Radu Dobrin, during the Precision Oncology Leaders' Summit on March 30, 2021. Their discussion focused not only on the value of adding RNA sequencing to precision biomarker strategy but also underlying reasons of why adding an RNA strategy is important to pre-clinical and clinical research. No written summary can capture the rapport between these experts and the spirit of their discussion, so please navigate to the recorded version on the PMLS website. Note also that this is a redacted version of the discussion – all the content has been retained but rendered in reader-friendly format (see https://www.engagez.net/pmls-precisiononcology2021#lct=conferencecenter--823862-calendar_192930_5ondemand, log-in required).

Radu:
First, may I say it is a pleasure to talk about RNA and biomarker discovery with Joel. We have known each other for quite some time – we met long ago at Keystone Symposia on Genomics, if I'm not mistaken. Could we start by your telling us about Tempus, its genesis and capabilities? It would be interesting to look back and consider the niche Tempus has since occupied in RNA, especially since I am mostly familiar with precision medicine companies in DNA around immuno-oncology.

Joel:
Thank you, Radu. Let me start by talking about what we do in particular since not everyone may be familiar with Tempus and how we have worked with you in the past. Tempus is a precision medicine company offering diagnostics to oncologists and expanding into cardiology and neuropsychiatry. We are unique for several reasons: one is that we are one of only a few precision medicine companies dedicated to multimodal data or multi-scaled data. By that I mean we run

not only DNA diagnostic tests from the tumor in the germline, but we also run full transcriptome on every patient. We also offer digital pathology analysis so we can integrate DNA, RNA, and pathology data along with ingested structured clinical records to generate a smart test for patients. Every test we run is contextualized by patient-specific information.

To me, that is the manifestation of precision medicine – namely, every test is made smart by organizing multimodal data on a patient and then synthesizing the patient's data into a recommendation.

We have an organoid lab as well. We grow and store patient organoids for future research. Our goal is to deliver a smart test back to the clinician with all those dimensions of data synthesized. I would argue that Tempus is probably the only true data-first precision medicine company operating today. When we get a sample in our hands, we ask, "How much data can we generate from this sample that will give us further insight into the biology of this disease?"

That is why we have been generating whole transcriptome RNA from day one. We now have a very large cancer transcriptome database, just five years after launch. It is several times larger than the cancer genome atlas, which was a large government effort.

The importance of collecting something like RNA is that it is a dynamic measure, and as you know, Radu, cancer biology is highly complex, dynamic, and context specific. I always say, DNA is just sheet music. It does not really tell you how the song is being played.

Keep in mind that RNA is just one dimension in a data space; there are other dimensions, such as proteomics. But the beauty of RNA is that it is a sensitive molecular sensor, a read out for what is going on inside the cell or group of cells from bulk RNA. it is almost impossible to get close to the true dynamic biology of a cancer cell without including something like the transcriptome, so maybe, Radu, you could talk about your perspective on RNA.

Radu:
And Joel, as you were saying, I'm a big fan of layered data where we have much better completion of data per patient. We can understand molecularly what happens both before and after treatment. As you noted, most companies just do DNA; perhaps you could expand on how and why Tempus chose to get into RNA?

Joel:
Thanks for that question. Tempus was started by Eric Lefkofsky, who was the founder of Groupon and several other tech companies. His wife unfortunately had breast cancer and that personal experience motivated him to create Tempus. Eric was not in science, but the foundation of Tempus proves the saying that it's sometimes good not to be too burdened by knowledge, especially when you attack a biomedical problem. And I mean that in a very positive and endearing way.

Radu:
Yes, I have heard that. I'm a physicist!

"We can integrate DNA, RNA, and pathology data along with ingested structured clinical records to generate a smart test for patients. Every test we run is contextualized by patient-specific information. To me, that is the manifestation of precision medicine – namely, every test is made smart by organizing multimodal data on a patient and then synthesizing the patient's data into a recommendation."

Joel Dudley

Joel:
Exactly! Sometimes you get heavily biased by how everybody's been thinking about a problem and the problem has been approached traditionally. When our tech founder and others with him looked at the problem of cancer precision oncology, they saw these limited DNA panels that people are running. The issue of a limited point of view was obvious to them almost from an engineering perspective. That perspective was a huge benefit to Tempus. They saw that big gaps in data limited our ability to understand alternate solutions. In every other domain, people are collecting huge amounts of data to build predictive models. They saw the mission of Tempus as filling in a lot of these missing data gaps.

And the reason, by the way, that others were not collecting these data were for good reasons. Because, for one example, the transcriptome on its own is yet to be reimbursed by insurance. That is why I maintain that Tempus is the true data-first precision medicine company – the thing that's most important to us is the data news flow for understanding not only cancer, but the cancer per individual patient.

So, we will not just "launch it as a diagnostic test, reimbursement be damned," with the goal of eventually associating data with clinical outcomes to justify a coverage decision. The founders had the fortitude and foresight to pursue multi-modal dimensions of data to fill in data gaps before launching diagnostic tests. They could build a more complete picture, and as we've been stressing here, you don't get a complete picture of the disease without connecting multiple layers of data from clinical to imaging to molecular. You really have to understand not only how those data fit together but how information flows between those data. That is why I had to join Tempus from my very cozy tenured academic position; I had never seen a multi scale data precision medicine company operating at scale.

Radu:
And pharma is definitely interested in this approach – layering data and understanding the transcriptome in the context of the patients. To me, that's what we are talking about today, it's truly the key. It allows us to link from the early development of a target to a critical, coherent development of all the pieces of information. We try to understand the whole picture – from what happens with the newly developed molecule to the patient-centric data generated by companies like Tempus. We're talking about data that comes from human samples, and the vast amounts of data we can collect in 2021 is truly amazing.

With advances of machine learning today, we can analyze everything quite rapidly and to identify biomarkers. We are in a position where everything is for patients. And Tempus, like you, yourself, Joel, embody the strong tradition of asking, "How can we change patients' lives?" Especially in our field and the growing ability to layer data – it's amazing, and the future is going to be even more about layering data. To me, RNA holds the key, because, as you said about the dynamics of RNA, we can look deeply into molecular pathways.

Joel:
And we love dynamic RNA in terms of building network models. But even before we get to the dynamic stuff and more complex presentations of the biology inside the cell, RNA can be far more performative than DNA for things such as fusions. Fusions, detection via RNA can be more fruitful than DNA alone, and by combining the two dimensions you can reveal biology. For example, you can do variant calling from RNA, as you can from DNA; but by contrasting RNA and DNA variant calling, you can look at things such as RNA editing that may be happening. By contrasting these two – even before you get to some of the more complex modeling that Radu and I like (being data science folk) – you can start to see a

rich world of biology revealed by RNA that's simply not possible to see solely through DNA.

Radu:
Maybe we can look back a bit, Joel, because I was not thinking about RNA in this evolution. When we met long ago, when I switched from physics to biology, I was first drawn to microarrays. This was around 2000. I had no proper biology training, and I was amazed to look at so many genes simultaneously and try to understand the connections between and among them. If you look back 20 years, you realize how much harder it was to come up with an RNA biomarker. With the sequencing advancements today, we're in a much better space to develop these novel biomarkers. What's your take on that?

Joel:
That's a great point. I published many papers using RNA microarrays. The data that's generated by them is still potentially useful, as evidenced by the huge number of biomarkers that have been developed from microarray data. But yes, we definitively switched into a new era with next generation sequencing and RNA Seq. With arrays, the challenge is to predefine the probes and build them on the arrays. Those probes address certain questions in terms of capturing dynamic expression, or certain types of sensitivity for expression from low to high levels.

But the beauty of whole transcriptome sequencing (WTS) is that it is arguably unbiased. You're amplifying and sequencing all the RNA that's present. Many straightforward examples of how that's powerful can be cited – and not only because WTS captures a better dynamic range of the RNA. A whole universe of RNAs exists that can be captured with WTS that is not typically represented on arrays – for example, microRNAs, lncRNAs, snoRNAs and picoRNA. Beyond that we can also capture things like viruses, which is super interesting. We see RNA viruses in patient samples now, and you can start to uncover potentially a universe of oncogenic viruses. By using the unbiased transcriptome approach, we survey all the RNA virus material in cancer samples. We would not have seen that material with microarrays unless we knew the sequences beforehand and designed specific probes.

Radu:
Yes, exactly, and I remember long ago, we did try to design special probes to capture that information. Everybody tried that! The technology evolves so rapidly. These days, because of sequencing, we're able to uncover things that we did not even know. We were hoping back then that it might have been possible to uncover such things, but now the data are accessible! If you think about biomarkers, there's still a distance to go for the more clinical aspects – to be able to move from research-purpose analysis on to the samples we're collecting and the data we're generating for the clinic. And now in oncology, we are moving from extracting RNA in bulk samples towards single cell RNA sequencing. That's an exciting recent development.

Joel:
It's funny and still amazes me to this day – sometimes you have to step back to appreciate how far we've come. Every time I use our 10X sequencing machine I'm just blown away that sequencing has become a routine measurement that we can capture with such reliability and such fidelity. It was unthinkable not too long ago. But there is a middle ground between the bulk and cell-level analysis, and many people do not appreciate that informatics serves as that in-between. Let's say we perform many RNA Seq runs on a bulk tumor sample; by combining bulk RNA and single-cell sequencing data, it's possible to deconvolute bulk samples computationally. So, without doing single-cell sequencing on a sample, you can deconvolute an immune cell that may be infiltrated into a tumor and get a computational estimate of how many CD8 or CD4, XYZ set of cells are infiltrating into the tumor. For immuno-oncology, that deconvolution can be very useful. But many people do not realize that you can calibrate the bulk RNA Seq on some single cell data, and then with the bulk RNA Seq alone, get single cell, or at least cell fraction, estimates.

> *"To me, that's what we are talking about today, it's truly the key. It allows us to link from the early development of a target to a critical, coherent development of all the pieces of information."*
>
> Radu Dobrin

Radu:
Another amazing thing about single-cell RNA Seq is the ability to understand the shift in cellular populations, especially on treatment samples. When you're looking at the tumor microenvironment and the complexity of the tumor, it's quite interesting to look from the point of view of single cell RNA Seq to assess what happens in the tumor in treated patients. In fact, one can see cell populations shift from immune cells to tumor cells.

It's one of the greatest advancements that allows us to dig deeper. I know, Joel, you looked in your career at this, but we relied too heavily on de-convoluting data without single cell RNA seq. That is quite challenging, isn't it!

Joel:
Yes, and this thread of it – whether it's cellular de-convolution from bulk RNA Seq or doing single cell sequencing on patient samples – to get extremely high-resolution data on what novel cell populations could be in a tumor. Many of the biomarkers we use clinically today are fairly rudimentary. Take tumor mutation burden (TMB) for example, which is a clinical biomarker that's used for pembrolizumab and other immuno-oncology drugs. It's really based on some per unit aggregation of tumors and a biomarker for neoantigen potential and such – that's pretty rudimentary when you think about it. With data like bulk tumor RNA and single cell sequencing, we'll eventually move towards a a TMB 2.0, which requires the discovery of new antigens derived from immune cells.

If you have neoantigens without immune cells, then it doesn't matter, right? A paper came out from Andy Anderson not so long ago, where he and his team analyzed several large data sets, along with the cancer genome and other atlases; they suggested, and this makes total sense, that TMB is a poor predictor without understanding the immune infiltrated cells into the tumor. Again, you can't see the bigger picture without RNA data or whole side imaging data. Anyway, we are going to move to a world where these fairly simplistic clinical markers like TMB will be enhanced with this extra information that can be extracted from RNA and imaging.

Radu:
We have looked at papers, lots of papers, regarding inflammation of a tumor. We're talking about hot tumors, cold tumors, and that's been possible by adding another axis – RNA. New insights were not coming from DNA, so understanding tumor inflammation, and especially correlated with high tumor mutation burden, could give a higher probability of success. As you noted, DNA's the sheet music and RNA's the song. The inflammation side will open a new avenue through combining RNA with DNA. However, I just want to talk about signatures. Some of the first work I did was at Princeton with Ihor Lemischka regarding signatures, and it's interesting looking back on such studies that included RNA.

As soon as people saw that inflammation matters in immuno-oncology, there was a flurry of activities and papers about the diversity of signatures, which is a measure of the power of RNA. Scientifically, we need to go towards a more, I'd say, focused point because some of the different signatures do correlate with each other. The focus should be to understand the signatures much better.

Joel:
Yes. What's unusual about Tempus is we do the opposite of what other people do. We go to market with the broadest possible assay so that when we have novel insights in the future, we can just layer those insights onto existing tests in Algos, our business for our digital diagnostics market. The deeper data and insight will be more specific for that result.

As you mentioned, there are many uses of RNA signatures. I'll give you a concrete example: prognostic RNA signatures – from therapy response to radiation therapy response for disease progression. You could curate the literature and find literally thousands of these signatures to varying degrees of quality and clinical evidence.

But because we run whole transcriptome, one could discover a novel predictive signature from RNA – let's say responders vs non-responders. Researchers are going to say, "Now here's a case control study where we have people who responded and didn't respond. Let's use whole transcriptome as our discovery tool." And they might find through statistical analysis, "Oh, here's this 15 gene signature that predicts who's going to respond and not respond." A typical diagnostics development company would then go to market with that 15 gene signature through a diagnostics commercialization partner; from then on, the assay would be that 15 gene signature and if they wanted to update it, go through the whole process again.

Tempus is unusual in that we run whole transcriptome by default. So, if we have that 15 gene signature, we could then layer that signature on digitally. Basically, because we measure the whole transcriptome, we're trying to get to where existing signatures become digital tests in Tempus' platform that get layered onto our whole transcriptome data. As I said, this tends to be the reverse of how people usually do it. For a host of reasons, they develop bespoke specialized tests. Tempus' mission is to transport a whole world of these signatures into the RNA Seq transcriptome universe.

Radu:
From a technological perspective, Joel, what is your take? I know there's a lot of discussion in the biomarker and bioinformatics field about signatures per patient. Do you think we're going to be able to customize to patient signatures? What's your thought?

Joel:
Yes, that's the world we're moving towards. That's the dream, of course, of mine and others. We're starting to see this in the minimal residual disease (MRD) space where we make very simplistic DNA signature markers for patients and track their progression post-treatment – or at least the emergence of certain mutations as they may respond or not respond to treatment.

"I'll never forget the many drugs we studied that "broke their profiles" and were off patent. The challenge is the lack of good internal RNA data from pharma on their new drugs. For the most part, not all pharma generate RNA signatures on their drugs, and pharma that do generate data tend not to publish it."

Joel Dudley

But in RNA land, that's the dream. We at Tempus plan to realize that dream by our ability to collect data longitudinally. Most of the cancer data we have is our snapshot for some specific time. If we were to look at the number of datasets where we have multiple molecular measures – dynamic molecular measures, such as RNA or methylation longitudinally over the course of patient disease – that's a relatively small number of datasets. We have some at Tempus – we generate those – but we need to generate more for longitudinal control cohorts and longitudinal profiling of disease with these dynamic measures. I know people are doing this with DNA and cell-free DNA, but more powerful, dynamic measures such as RNA and methylation will be key going forward to developing that data universe.

Radu:
That's an interesting discussion to have regarding the longitudinal data. We all have access to a baseline of archive samples. But RNA, as you said, is a dynamic measure. What you need is the time course of the patient, the disease, and the response to therapy. If you think about how a clinical trial is done now, we have snapshots of patient. It's not truly as dynamic as, say, stem cells. Of course,

Joel and Radu addressed a few questions raised during their chat. We provide the questions and answers:

Q. How do you infuse patient experiences input into your development?
Joel: I can speak to it from the Tempus side, which would be more on clinical cancer through the oncologist and the clinician. However, we have a neuropsychiatry vertical that has spun up where we're doing pharmacogenomics testing and other types of testing. In the neuropsychiatric instance, we have a patient-reported outcomes mobile application deployed since EHRs are not great at capturing the salient features of those phenotypes. We have expanded that mobile application into our COVID testing going to patients. We see this moving forward into cancer as well, where you're directly interfacing with the patient and getting patient-reported outcomes.

Radu: For pharma, I don't think we've been looking at patient reported outcomes for the last couple of years. With the advent of combining EMRs with more complex data and the push of understanding what happens in a real world setting with a patient – that's driving us, and we're looking quite deeply into it. I cannot tell you details about it, but it's something that is going to come fast in the future.

Q2. How can longitudinal data be economically feasible for regular clinical patients?
Radu: Actually ... I'm not sure I have an answer! As a scientist, deep down in my heart I do want to have longitudinal data. I see the biggest power in having longitudinal data. it could change

doing experiments in a dish is much easier – you can profile every hour if you'd like.

What do you think, Joel? How can we move towards a better collection of samples so we can see the trends in a patient's RNA and power the data for precision medicine?

Joel:
The challenge here is always finding ways to do it in a clinical trial. In a clinical trial you can use part of the trial for collecting samples, but Tempus is focused on how to insert data collection into the normal clinical workflow at scale. Let's try not to introduce new behaviors for new tests that doctors are not accustomed to ordering. We want to develop tests that they're accustomed to ordering in their workflow and to make those tests smart by collecting more information than they have in the past. When you look at opportunities to do this, it turns out that minimum residual disease testing is a great opportunity. But, MRD tests currently on the market are very simple – two or three gene PCR based tests, or methyl converted PCR.

Cell-free DNA is being used for these tests but again, the biological context can be limited. You're detecting a signal, but it'd be much more interesting to get circulating tumor cells and do single cell profiling because now you're capturing biological context. Not that cell-free DNA isn't a useful biomarker! As the minimum residual disease space evolves from these cell-free DNA PCR based tests to include other types of tests – whole methylation profile, whole transcriptome profiling – that's when we'll reach the next level of diagnostics in that space.

Radu:
Yes, so you're talking about technology. What is the Tempus' point of view from the center – from generating data, to working with patients, to working with pharma, and your personal experience from having worked with lots of big tech companies in US and outside? What do you see as the evolution of the technological landscape? Because one of the discussions we've been having for the last few years is reducing costs. One of the key components of doing RNA Seq is cost. That's why many develop reduced signature sets. Where do you see it going?

Joel:
Yes, that's a great question. One thing that's interesting, by virtue of our scale as a company, we believe that we can drive the cost of RNA sequencing down. Prior to Tempus, there weren't any companies clinically offering RNA sequencing at scale. We do have economy of scale that come from our size and our scale of sequencing. The first straightforward step is to offer a whole transcriptome sequencing at scale to lower cost. Of course, there have to be technological advancements that drive this down, and that is something we're always focused on at Tempus.

We are aware that alternative sequence platforms are evolving. I just saw a preprint from a friend of mine who's working with Oxford Nanopore on a low-pass whole genome methylation profiling method. Oxford Nanopore may be able to provide a low pass whole genome methylation profile that looks as good as Methyl Seq on an Illumina platform.

Those assays are in progress, and it may take some work to deploy them in a clinical scenario, but there are other alternative sequencing platforms that continue to get interesting. Computational improvements will continue, so as we can start to leverage machine learning and deep learning to move, for example, from low fidelity to high fidelity imaging, specifically from fuzzy images to sharp images. One of the main tasks of deep learning image processing is to train on large datasets of

with the way we do medicine. If your doctor draws blood, you could collect RNA profiles along with blood pressure. At some point, it could be economical to look for relationships between, say, RNA, DNA, cholesterol, and blood pressure. That's something that might help us. Right now, if you think about just a single pharmaceutical company spending so many resources to generate longitudinal data in a clinical trial, we're talking a dramatic increase, and I don't know if that's easily translatable into a high-density dataset.

Joel: We just look at opportunities where longitudinal profiling already exists. To some degree, it is part of the clinical workflow.

Q. Given the personalized nature of RNA Seq from patient to patient, how do you determine the accuracy and precision required for clear data and decisions?
Joel: What makes that a great question is that it is all about technical validation and reproducibility. What is so challenging is that RNA is dynamic, even minute to minute in some patients. We need to establish those patients' history using informatics to normalize and harmonize that data correctly, then we can be accurate.

Q. Can you talk about the benefits of combining DNA and RNA sequencing for clinical trial recruitment and companion diagnostic development?
Radu: Both of us can speak about this. For immuno-oncology, combining inflammation with tumor mutation burden tells you a lot about which population segments in a clinical trial would respond to your drug. In the whole IO space right now, just about everybody is using RNA with DNA and looking at even more complex ways of combining the data.

IO opened a gate to understand patients better, but now, people realize that a single biomarker is not enough to identify the patient's response to the drug. That's why we are looking at different ways and different modalities of combining data and that's why we're talking about so many different approaches to classify the patients.
Joel: My quick follow-on is also those combined models. If I had my slides I could show some examples, but those combined models tend to perform better from an area under the curve (AUC) standpoint.

Q. How are pharma leveraging RNA data across the drug development process?
Radu: We leverage data through forward and reverse translation. Basically, we're using the patient data to understand new targets and then we take the patient data from the real-world setting to understand who might respond to the drug. We leverage RNA Seq data across the entire pipeline. In 2021, there's no space where we don't leverage it. We collect clinical trial data and will also work with partners to understand what happens with the patients.

So, I'm not sure there's a therapeutic space or a pharmaceutical company that does not work with RNA these days, especially to understand the pathways and MOAs. We're not thinking just from a biomarker perspective but also from a new target perspective to power our next generation of molecules. Both components are actively pursued in pharma today.

low-resolution and high-resolution images. In that case, you can think about doing the same thing in the RNA sequencing space. For a given RNA Seq profile, you can artificially create a simulated low coverage profile while at the same time you have the 50 million read version of that patient for training purposes.

What we'll start to see is that deep learning may afford the ability to start sequencing at shallower and shallower depths, without losing fidelity of the signal. We believe we'll see that the vast RNA datasets, like the ones we're building, will enable the creation of deep learning-based AI algorithms that will restore the fidelity of a lower pass signal

Radu:
Can you say if Tempus is involved with that, or if you're looking to collaborate?

Joel:
Yes, both. We're collaborating with some folks who are developing these capabilities and we're also looking at developing new methods internally. I can't say too much more about that, but the key thing on that is you can't develop those without first collecting those massive, massive datasets, right? Deep learning, as you know, is extremely data hungry, and you have to have massive, massive data sets to train those algorithms. We're finally getting to a point with the scale of our data that we can start thinking about those types of applications.

Radu:
You know, we've been talking about RNA and, as researchers, we do like the research angle of combining data and understanding molecular pathways, coming up with new drugs. I want to ask you more about your take on drug repurposing and RNA. Everybody knows about your papers, about how to use RNA and the ability to look at drugs and patients using RNA, not DNA. If you didn't have RNA, I don't think there was any way to do that, so what do you think? Where's the evolution, and for us in pharma, what lessons can we learn from the last couple of years of drug repurposing and measuring RNAs?

Joel:
Yes, so the work I've done in this space was enabled by some large resources that were created by the Broad Institute many years ago now, called the connectivity map dataset. As one of the first large efforts to perform RNA sequencing, it was not yet the full transcriptome, but RNA sequencing done large. And now the Lynx Project is doing similar things. What was eye opening about the connectivity map project and how we used it was, first, that when you measure the transcriptome of the drugs and the mechanisms of action (MOA) as they percolate through the cell, one finds that the transcriptome and the MOAs are much more complex than anyone realizes. I'll never forget the many drugs we studied that "broke their profiles" and were off patent. The challenge is the lack of good internal RNA data from pharma on their new drugs. For the most part, not all pharma generate RNA signatures on their drugs, and pharma that do generate data tend not to publish it.

On the other hand, I was working on a project with a large pharmaceutical company, and we were getting RNA sequencing data on a bunch of different cell lines for a drug candidate. I asked the chemist who made that particular compound, "What do you think we're going to find in the RNA data?" He goes, and I'm paraphrasing here, "Well, let me tell you. It was a small molecule, a very specific, exquisite inhibitor of this GPCR. What we're going to find is a very clean signature where this gene gets down regulated and this pathway is up regulated." That's a very idealistic representation of the mechanism of action.

When we performed the RNA Seq on various cell lines, I said, "Did you know that this drug differentially expresses 800 different genes, and it causes the differential splicing of 200 other genes?" The chemist was blown away! He assumed that because he designed it for this one target, that the mechanism of action would be very clean and simple, and of course we know that's not the case for these small molecules. But that is not a bug. That's a feature, right! Even if we don't fully understand that drug's mechanism of action in various cell types, we can pattern map the RNA profiles to diseases for RNA profiles of diseases we've already captured.

I almost think the analogy is like you throw a rock in the pond, you take a picture of the puddles, or the ripples that come from that. And it doesn't matter because you can find another pattern of ripples that matches that pattern of ripples that you know and overlay those. We can kind of do that with drugs and disease. That's the beauty of RNA – it can capture all of that complexity. Even if you don't fully understand the explicit biology, the pathway diagrams (and everything underlying those diagrams) could still be used to match drugs to disease.

Radu:
That's exactly why we in pharma were so excited about RNA to start with years back – being able to link discovery to development. I have led teams heavily involved in that research, and the beauty of working in pharma is making those links and connections. But other than start with a molecule and make those assumptions that everything is clean, everything is perfect, we have the ability to generate data at very early stages. We can understand what happens from simple individual experiments to more complicated animal models and then into humans. And that was something that has, for pharma at least, opened the avenue to understand underlying biology much deeper. We can look at the risk we're taking with some of the molecules much earlier in the pipeline versus waiting very late into our Phase 2 clinical trial to understand what happens with the molecule (so-called "fail early"). We're not driving blindly later.

I don't mean to discount all the protein and other biomarkers that people discover. But RNA profiles give us a better understanding now about the molecules and how to position them in different indications and truly drive towards precision medicine. We can identify subsets of patients that the molecule is working on. You see in the biotech space and pharma that drivers of segmenting the population segments into smaller and smaller patient subsets was mainly due to RNA. We can eventually see the landscape to segment anybody, correct?

Joel:
Yes, absolutely! And I have a question for you, Radu, which is why pharma doesn't perform whole transcriptome profiling on drug libraries, given there's so many papers that show the value of these profiles. Why not profile even the mature assets in the pipeline to understand more broadly the cellular mechanisms and match signature to patient samples or patient subgroups based on clinical RNA signatures? My observation is that pharma is slow to adopt the use of full transcriptomes, both for understanding their drugs and patients in trials. Is that right?

Radu:
That's a great question. That's why a company like Tempus came to fruition – there is a gap! I'm not going to have the perfect answer for all of pharma. it's quite complex to use RNA in a clinical space. I remember when I was a junior scientist I thought, "Oh, I can ask clinicians, just take 100 biopsies from the patients in the next couple of months because I need a very tight time course." And once you start talking to patients and clinicians, you realize that you have to find a middle ground between the number of samples, the types of samples, and sometimes the priority on clinical trials. You have to deal with a very complicated landscape to push the drugs forward.

And all this time, we keep in mind that we want to help the patients. We have to take really tough decisions about, "Are we going to profile the entire cohort or not?" People do run full transcriptome

profiling for most of the clinical trials across pharma. We're talking 2021, and we're going to see more of that because it does allow you to see molecularly what happens with a patient.

Pharma teams ask more and more about what happened in a trial, what happened with this molecule, and it's not a simple analysis – It's another level of complexity. People want to understand if it's positive, why is it positive? If it's negative, why is it negative? How do we differentiate these drugs? You know, we're all competing on maybe a couple of pathways in immuno-oncology, why one molecule is better than the other one. We're all thinking about patients. That's a convoluted answer to your question but I cannot say we can easily do it across every single trial. We always have to keep balance, but we're moving that way.

Joel:
Yes, and when you talk to clinicians, they often like DNA because it's a very binary, if you will, and deterministic in a way. Either the variance is there, or it's not. And then there's the decision to be made if it's there or not. Other clinical tests that clinicians run have standard normal ranges established. If an immunoglobulin is high, then I do this thing, if it's low, I do this thing. They're able dichotomize a lab test result. The challenge with RNA is that it's so sensitive and so dynamic that there's no such thing as a normal range. It looks more like an analog signal than these clean digital signals that they think they're seeing when they look at DNA for clear decision points. That's the challenge with RNA sequencing, not only clinically but also in clinical development. There's no such thing as a standard normal range for an RNA transcript, or a gene signature. And it makes it complicated for people to hold that probabilistic signature in their head and determine what to do with that information.

Radu:
That's a very, very good point. One of my early career experiences was the discussions I had with clinicians when they asked me, "So how do you want us to involve people in the trial? What is high, what is low? What is the cutoff?" I was looking at correlations and p- values, so it was hard to say. I would reply, "What do you mean? Do you want me to draw a line and say that this is good, and this is bad?" As you know, it's quite difficult to draw that line. The way we do medicine by changing the future, that's all I'm looking for, especially with the combination of data. When I go to the doctor I want to know, "Am I sick or am I not sick, do I get the drug or not?" It would be weird if the doctor tells me, "There's a probability and a correlation and you choose." This is not what you want to hear from your doctor, right? It's quite a challenge for us to come up with cutoffs.

Joel:
To be fair, the question is should we standardize a normal range or cutoffs as high, medium, or low in the RNA world, whether it's an individual transcript or a pathway? We simply didn't have the reference datasets to think about that. So again, putting on my Tempus hat, that's the beauty of creating a large cancer transcriptome database. Not only do we have that transcriptome data at scale, but we also have the clinical outcome and the clinical data associated with that patient. This is probably the first dataset where we can start to think about creating these clinical cutoffs around concepts like the patient transcriptome, because we have not only the clinical data done large, but also a representation of patients for different ethnicities, different cancer types, different drugs. We can start to consider how we could help clinicians think about using RNA in a more precise way.

Radu:
That way of thinking has a lot to do with technology and why I was talking about microarrays versus RNA Seq. I remember the first cohorts that I was working on – they had so many biases. Granted, the technology was not ideal for translation into the clinical space, and as it was early days, very difficult to get clean data in the sense of easily analyzable. We had to spend, I'd say, 80% of the time just cleaning up artifacts from the technology. With technologies like RNA Seq and Oxford, we can get to a place we could not reach before. It's going to enable us to bring transcriptomics into the clinic and that is the most important thing to me.

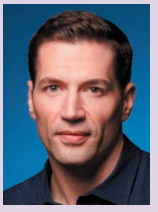
Joel:
Let's talk about future things that will be important in the RNA diagnostic world. I'm going to be very shameless here and talk about something we're doing with our RNA V2 product, which I'm really excited about. I hope other companies advocate their transcriptome capabilities because we need more of this type of data. So, for our Version 2 RNA Seq assay for whole transcriptome assays, we've baked in a bunch of TCR and BCR, T-cell receptor and B-cell receptor probes similar to what you would find from an adaptive biotech. Again, we know that data is going to be useful.

As RNA Seq and its uses evolve, Radu, do you see anything else that we should think about measuring or baking into RNA profiling? or how we use RNA?

Radu:
What's interesting to me would be to see single cell RNA Seq evolve a bit and start going from atlases as it happened with microarrays. There were a whole bunch of atlases that we started with and then we started doing perturbation, so I'd like to see us get to perturbations, too. The technology needs to evolve to allow us to use it more easily in clinical trials. Currently, it's not easy to use it in a clinical setting, especially for single cell RNA Seq. To have the ability to profile tumors and tumor cells, since tumors are so heterogeneous. The mutation landscape is so complicated, we need to have better technology to really integrate the data.

And I'll tell you my shameless pitch, I would like for us to think more holistically about signatures. As a field, we in bioinformatic spend lots of time coming up with lots and lots and lots of new signatures that most of them prove to be correlated to each other. I'd like us to move to a place where we understand signatures. it's been 20 years of research around signatures and molecular profiling. We need to nail down what is different and what is common among signatures? Versus, "Let's write another paper about the new one."

Joel:
That's great. yes, totally agree. Well, this has been fun. we're at the point where we take questions. We should stop flapping our jaws and listen now. JPM



Joel Dudley

Dr. Dudley is the chief scientific officer at Tempus Labs. Prior to joining Tempus Labs, Dr. Dudley was the associate professor of genetics and genomic sciences and founding director of the institute for Next-Generation Healthcare at the Icahn School of Medicine at Mount Sinai. He also served as the executive vice president for precision health for the Mount Sinai Health System. Dr. Dudley received his MS and PhD in biomedical informatics from Stanford University School of Medicine.

Radu Dobrin

Dr. Dobrin served as the head of bioinformatics, immunology, and immuno-oncology for the thematic research centers at Bristol Myers Squibb (BMS). He holds an adjunct professor position at the Icahn School of Medicine at Mt. Sinai. He was responsible for advancing BMS's pipeline through the application of cutting edge bioinformatic approaches. The group engaged all bioinformatic aspects of immunology and immuno-oncology at BMS, including early discovery, translational research, and clinical development. Dr. Dobrin has a dual PhD in physics and an MS in electrical and computer engineering from Michigan State University.