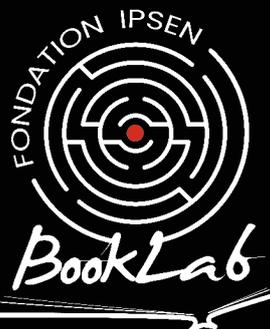


# The Rare Disease Gazette

*Conversations with  
the world's experts  
about rare disease*

Issue #5 | September 2021

## ARTIFICIAL INTELLIGENCE



# Editorial

by **James A. Levine**

MD, PhD, Président, Fondation Ipsen

[www.fondation-ipsen.org](http://www.fondation-ipsen.org)

Rare Disease Detection: Rare But Not Alone



The plight of patients with rare diseases is a critical unmet need of patients in health-care. The statistics are frightening; there are 7000 rare diseases in the world that affect 350,000,000 people. One in eleven Americans has a rare disease. Three-quarters of patients with rare diseases are children and only half of patients receive an accurate diagnosis. The average delay for a patient to receive a diagnosis with a rare disease is 1 1/2 years. It is deeply concerning that one in four patients with a rare disease waits four years for an accurate diagnosis. There is an urgent need to communicate knowledge and expertise in the field of rare disease detection.

The journal *Science*, (American Association for the Advancement of Science) in collaboration with Fondation Ipsen delivers international science webinars for the general public. In 2021 these webinars focused on improving the detection of rare diseases. *The Rare Disease Gazette* is a magazine that broadcasts these discussions.

*James Levine*



## DON'T MISS!

**The Conversation, page 1**

**Experts of the month:** Sean Sanders, PhD, hosts a conversation with world's experts about Intelligent detection and diagnosis of rare diseases

**Journal Club, page 8**

**Article of the month:** Artificial Intelligence (AI) in Rare Diseases: Is the Future Brighter?

**History Corner, page 9**

**History of the month:** The remarkable history of Leprosy.

**What's up?, page 9**

**Highlights of the month:** Movies to move you

# The Conversation

## Experts of the month

**Sandra Brasil, Ph.D.**  
**Julian Isla**  
**Sean Sanders, Ph.D.**  
**Ben Solomon, M.D.**  
**Sylvia Thun, M.D.**

### Sean Sanders (host):

*Hello everyone and thank you for logging into this fifth part of our Science and Life Webinar Series on rare diseases. My name is Sean Sanders, and I am Director and Senior Editor for Custom Publishing at Science. I am happy once again to act as moderator for today's discussion. In this nine-part series that will run through the remainder of 2021, we are unpacking many different aspects of this important topic of rare diseases. If you missed our previous webinars in the series, you can find archived recordings at [webinar.science-mag.org](http://webinar.science-mag.org). In today's event, we are diving into the topic of intelligence testing for rare diseases with an emphasis on technologies like artificial intelligence that can aid this process.*

*I am now delighted to introduce our four fantastic panelists that we have on the line today, and as usual, I am going to give each of them an opportunity to introduce themselves to you, and we will do that right now. So welcome to all of you. Thank you for making the time. Perhaps I can ask Dr. Solomon to start us off.*

### Ben Solomon:

I am Ben Solomon, the Clinical Director of the National Human Genome Research Institute. I'll call it NHGRI, which is part of the NIH here in the United States. I run a small research group that focus specifically on artificial intelligence as applied to genetic conditions. Previously I was the head of a biotech company that did a lot of genomic and genetic sequencing, and of course, as we'll talk about AI and related methods, are very important there. The other thing I want to mention is that I interact and help support a lot of investigators

here who look at rare conditions, genetic conditions -all aspects of them -and they're starting to use more and more of these approaches in their research.

### Sean Sanders (host):

*Next, we have Dr. Brasil. Welcome, Sandra.*

### Sandra Brasil:

I am a Researcher and Patient Advocate, at the Portuguese Association for Congenital Disorders of Glycosylation, which is a rare disease, and we are part of the Glycoimmunology group under University NOVA Lisbon. We have also created an international network, joining together several researchers, medical doctors, and patients, their families, and patient associations from all around the globe, which is called CDG & Allies- PPAIN. Part of our work is to help families and patients and give them resources. One of the things that we do is we go through revisions, and we have been working on the team about artificial intelligence and how it can help push for research and therapies in rare diseases, particularly congenital disorders of glycosylation.

### Sean Sanders (host):

*Fantastic, thank you so much, Sandra. Next, I am going to turn to Dr. Thun.*

### Sylvia Thun:

My name is Sylvia Thun, I am from Germany, Berlin and I work at the Charité Universitätsmedizin where I am a Medical Doctor. I am not only a Medical Doctor, but also an engineer and informatician. For 20 years, I have worked within the field of Digital Medicine and I am the Chair of HL7 Germany. I have also worked together with HL7 International ISO in the global alliance, to have better and more standards in the world for IT in the medical field. In addition, I am working here in Germany together with the genomDE initiative, and in Europe with One Million Genomes and Beyond project.

### Sean Sanders (host):

*Wonderful, thank you, Sylvia. And last but certainly not least, we have Mr. Isla. Julian, please go ahead.*

### Julian Isla:

Thank you. My name is Julian Isla. I am based in Madrid, Spain. I work for Microsoft and my team is working on data and

artificial intelligence. On top of this work, and because I have a son with a rare disease, I started to be engaged as patient representative 12 years ago. Right now, I am the Scientific Officer of the Dravet Syndrome European Federation, which is the European Organization for Dravet syndrome. Dravet syndrome is the condition of my son. Also, I am the Founder and Chairman of Foundation 29. Foundation 29 is a non-profit organisation, trying to use much in learning, to improve or to create better decision support tools for physicians. And I am involved in the regulatory field because I am a patient representative to the European medicine agency, and I am a member of their Orphan Drug Committee.

### Sean Sanders (host):

*It is really fantastic to have you all here, you come from very different and varied places and organizations. I think we can have an interesting discussion. I would like to start this webinar by talking broadly about artificial intelligence, what it is and what it means. And then we'll dig in a little bit to the application of artificial intelligence in rare diseases. So, Ben, I think I am going to come to you if I could with this first question. And that is, how do we define AI? And how is it different from machine learning and deep learning, which are other terms that people are probably familiar with related to artificial intelligence?*

**"Artificial Intelligence (AI). Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind."**

<https://www.ibm.com>

### Ben Solomon:

I think it is important to start by defining some of these things. We hear them talked about all the time, in lots of different contexts. The way I like to think about it, is as maybe a series of concentric circles. Artificial intelligence is the outermost circle. And as the name implies, it is the use of algorithms, software to help computers act intelligently. And I am doing that intentionally in quotes, hopefully, intelligently,

hopefully, accurately. There has this idea that maybe it mimics human thought patterns. I think superficially, there are some connections, but there are a lot of differences. And there has this whole fascinating field about of cognitive science in that area.

Within artificial intelligence, as the name implies, machine learning is the process through which machines or computers can use algorithms to learn from a set of data. You let it run and machine learning happens. There are lots of different nuances to this. More recently, there has been a lot of emphasis and a lot of buzz around deep learning, which is, again, a circle within machine learning. And that is the idea that we take layers, and these different layers extract more sophisticated or precise information as you go down in the layer.

A classic example in genetics that happens is we are say we are using deep learning to help diagnose patients who might have a genetic condition. We train a computer to be able to recognize that from images of different people. In the kind of first layers, the computer might be able to recognize, okay, this is a border, or this is part of the picture, and then recognizing, say, facial parts, like ears or eyes, and then the whole face, and then eventually recognize, Hey, this is a picture maybe of Ben Solomon, or this is a picture of someone who does not have the condition of interest, or this is a patient with this condition of interest. There has been a lot of emphasis on that, partly because of the availability of some of these algorithms. Again, I would invite others to comment on this. This was full of inaccuracies and generalities. That is kind of how I think about it.

### **Sylvia Thun:**

I think it was a very good explanation, because artificial intelligence is the broader term. And then you have those other categories that are within the bigger definition.

### **Sandra Brasil:**

I think we need to use large data sets to train these algorithms. So, we should talk about data and about precise data.

Sean Sanders (host):

Absolutely. That is a really important point. I would like to ask you, and maybe, Sylvia, I can come to you with this, where are the areas that artificial intelligence really excels? And on the flip side, where does it struggle?

### **Sylvia Thun:**

First, where do we get the data from? Most of the time, we just use images, like images from our pathology images, or ideology images or even images from faces. We can predict what kind of disease somebody has. The second part are demographic data, and, as you all already said, omics data. And beside that we try to work with diagnosis and symptoms. And we have startups here in Germany or I think around the world who try to predict, with a symptom, what kind of rare disease or disease somebody has.

### **Sean Sanders (host):**

*Great. Any other thoughts particularly on where AI will have some issues?*

### **Ben Solomon:**

One recommendation is the book Deep Medicine by Eric Topol. It talks about this and other topics, and in that book, Eric makes the point that one of the ways that AI is really good at one of the areas is what he would call classifying or labeling. So just like Sylvia said, can I take a bunch of images that a pathologist might look at, or a radiologist, or a cardiologist looking at the heart or the electric patterns, and classifying some of those other things? I think one of the other areas that is interesting, and these really intersect, and some very clever people have figured out ways to make these intersect more than they would naturally, is looking at bunches of words. So, things like medical records and being able to make diagnoses that otherwise would be missed. In my mind, these approaches tend to work really well with large data sets that have been collated just again, as Sylvia said, very carefully to allow us to start to learn from them and allow the algorithms to start extracting information from them.

### **Sean Sanders (host):**

*One of the terms that comes to mind that researchers use often is 'junk in, junk out', and I am going to guess that if the data is not solid that you are putting in, then you are not going to get good results out. So maybe Julian, I can ask you this question: when you are looking at data that you are collecting, and we'll talk specifically about rare diseases, how do you ensure that this data is going to allow the artificial intelligence algorithms to learn what you want them to learn and to be able to give you trustworthy results?*

*"If you try to get good data sets of patient records for people with rare diseases, it is very difficult to get these records. There is no open access data for those records, and the quantity of medical records that we have for people for rare diseases is rarely used."*

### **Julian Isla:**

Right now, this is a big problem. One of the problems faced by the research groups working on diagnostic tools for rare diseases, is the lack of data. If you try to get good data sets of patient records for people with rare diseases, it is very difficult to get these records. There is no open access data for those records, and the quantity of medical records that we have for people for rare diseases is rarely used. The other problem that we are facing right now is system databases. We've created information about rare diseases, there are plenty of mistakes, because the curation of databases has been done by humans for so many years, and humans are prone to make mistakes. We know we are. Here's one example, for my son's condition, it is neurological. Several years ago, I realized one of the symptoms for my son's condition was photosensitive skin. And this is not actually a symptom for his condition. The symptom was listed in Orphanet, which is the European database for rare diseases.

I contacted Orphanet asking why do you have photosensitive skin for Dravet syndrome? And they told me, "Well, we don't know." But after some research, they realized that photosensitive skin in the tool they use for tagging is close to photosensitive seizures. And the problem was the coder, the people doing the tagging. They made a mistake, normal, they are human, but the mistake has been in the database for so many years. And this is a problem that is happening because when you are looking into other conditions as well, they have mistakes. There are symptoms that the patient groups, they don't recognize, and we don't have information about frequency, severity, onset. This is still an ongoing problem that we must solve.

### **Sean Sanders (host):**

*This brings me to the next point that I wanted to touch on, and that is, are*

*all data equal, or are some more useful than others? And just going off what you were just saying, Julian, is it possible to use AI to detect these types of errors in the databases?*

**Sylvia Thun:**

Yes, we have different kinds of data. And we should look at the data enterer, who is this? Who is responsible for documentation, and the software vendor? If you are working with terminologies like international terminologies like some entity or the HPO, that is the Human Phenotype Ontology, the software vendor can help to enter and to document the data in a correct manner. If you are not working, and you are allowed to just use free text, it is very hard for AI algorithms to use the data or to train the algorithm with this free text. And even if you are using NLP technology, natural language processing technology, it cannot capture data in a very good and precise way. So first, we should have a very common international language. And besides that, we must have an international syntax like HSM and FHIR, that is fast healthcare interoperability resource that is a language based on XML and JSON to exchange data.

***We must have a common international terminology, not only about the genes, but on the variants, but on the clinical phenotypes as well as for symptoms and diagnosis.***

Many of these diagnosed rare diseases don't even have a name. So, there should be an organization, a worldwide organization like Orphanet or the WHO or some other organization, who is responsible for the governance of giving names to diseases and of putting symptoms to the genes to the disease.

We must have a common international terminology, not only about the genes, but on the variants, but on the clinical phenotypes as well as for symptoms and diagnosis.

**Julian Isla:**

Sylvia raised a very, very important topic, from my point of view, the topic about how we put names to conditions, to diseases. Because, for example, the condition of my son, Dravet syndrome, has the name of

the doctor, the friend's doctor who identified the symptoms and the set of symptoms are the syndrome. But right now, this is not happening anymore. The diseases, they don't have the names of doctors, they have the names of the genes. And this is imparting in a very powerful way. Patient organizations as well, because the patient organizations, they have the name of the genes.

I am not sure if we must look for specific names because from my point of view, a name is just a name. If the name is not able to embed information, it is not useful. Let me get you an example, in astrology or astronomy, the astronomers are naming the stars with coordinates because the coordinate and the naming is giving information. If we name conditions like Dravet, Rett, Lennox-Gastaut down, you know, it is useless for our computer. The computer is not able to process this. I am wondering if in the new way we are naming medicines, the medicines would be a new methodology to process this information and that information into the name. That is the way to do it.

**Ben Solomon:**

I want to add a couple points in response, or at least echo what Julian and Silvia said. First, so there has a neat paper they have published about a year ago in the American Journal of Human Genetics that talked about clinical diagnosis, molecular diagnosis, and clinical molecular diagnosis, which as the name implies to it both, and I think I agree, there has a lot of work to be able to allow us to better define and think about what these conditions are and what's biologically causing them, as well as to allow smart computer systems to parse the data a lot better.

We are decades behind where we should be, or we are still using methods that were done decades ago to try to do things that are much more exciting, but were hampered by these older methods. The other point I want to make, again, just to echo what Silvia said, is that computers can sometimes recognize these problems. What I think is also exciting, the problems with data sets rather, and so on and so forth, what there has some interesting research about, is that computer algorithms can actually correct some of these problems, so for example, take a lousy image of an X-ray or some kind of neuro imaging, like an MRI or a pathology slide that is not really too good, can computers fill in some of the gaps and correct for that.

Computers will accept the biases and problems and everything in the original data set, so we can allow the computer to correct things, but we don't want them

to correct things based on lousy implications or suppositions and so on and so forth. Unfortunately, a lot of the data in the system so far are based on patients and control people, people without conditions from European, from White European ancestries. And extrapolating from those populations to other populations may not work, or may work much worse, or may introduce all sorts of problems. So, there has efforts to be able to correct that among a lot of the other issues there..

**Sandra Brasil:**

I just wanted to add because I think that is the nomenclature is definitely very important, and also I think that it needs to be, we need to have consensus even between the doctors that do the denudations and register the data because for example, again, from my experience working with congenital disorders glycosylation like CDG between Europe and, for example, the United States, there is not a consensus, a definitive consensus between the nomenclature within this particular disorder. I think this is also very important and very impactful because some doctors may not consider one specific disorder as CDG and others may do. You will have those differences in the records, and then that could have an impact in using these methodologies and artificial intelligence because the annotation will be different -then you are losing information in that sense. I think we need to have a way to establish very well what is the condition and the definitive nomenclature so we can all run on the same basis and give, use the same names, so we can use the same data.

**Sylvia Thun:**

There has a very nice article in Nature Digital Medicine from Cirillo about sex and gender differences in AI ( <https://doi.org/10.1038/s41746-020-0288-5>). We need to closely look at the data, for instance, our AKI algorithm about kidney injuries is based on data from 96% men. This algorithm can cause harm to women if we are not calculating this bias out of the data.

**Sean Sanders (host):**

*I guess just to reiterate the point again is if the algorithms are biased in the results that you get out are going to be bias. We really need to address how these algorithms might be biased and make sure that we are providing them with data that is going to remove that. I did want to come back to something.*

*So Silvia, you talked about having a standard syntax, having a common international terminology, having consensus, it seems that this is not really happening right now, maybe to some extent. What can be done about that, and also what can patients or patient families do to advocate for this? Now that patients might understand that this is an issue, that there is no common syntax, is there anything that they can do, patients and the patient's family to advocate for this type of thing?*

**"I think that patients should also be included in the research, so we should have patients-inclusive research, which is something that is starting to happen but it is not so common."**

**Sandra Brasil:**

Yes. I think that patients should also be included in the research, so we should have patients-inclusive research, which is something that is starting to happen but it is not so common. And, I think that we need to educate both sides, so researchers because from my experience as a researcher for the first part of my scientific path, we didn't have much contact with patients, so it was all very personal. And we were in the lab, we did not have any feedback from them. Now, working with the patients association is completely different. We had many projects that were led by the patients, and we discovered things that were new and were not known by the researchers. The patients have a lot of information, that sometimes they are not even aware of, and the researchers are not also aware of.

**"We need to create these synergies between patients, patient organizations, researchers, and put everyone in contact, and talking, so they can share their symptoms and what is important for them."**

So I think that we need to create these synergies between patients, patient organizations, researchers, and put everyone in contact, and talking, so they can share their symptoms and what is important for them. Researchers can learn with their experience, and then incorporate that knowledge in the research. I think that is definitely the first step that needs to be taken in that sense, to make sure that patients are included, so we could have researchers reaching out to the organizations or vice versa, and then forming little patient committees in which they have inputs about the research - then you can have their opinions. And also, I think that it will be very important to include medical doctors because they have a lot of information. I think that is very important to create synergies between all stakeholders, because all of the information is there, and the problem is that I think it is dispersed and we are not communicating properly, so I think we need to address that.

**Sean Sanders (host):**

*Julian, maybe I could come to you and, actually, the story that you told about your son's condition, and you advocated it for him and went to the organization, the databases, and told them there was an error. I would love to get your thoughts.*

**Julian Isla:**

One example of one of the experiments that we did in Foundation 29 (<https://foundation29.org/en/>) because we have been working on a tool to facilitate the process for diagnosis. And the tool has two versions, one is for patients, one is for physicians. The idea of the tool is that the tool could be a single platform to facilitate collaboration. In our idea, we thought, okay, we have patients with needs, we have physicians willing to help, why not create a platform to allow all of them to collaborate with each other? The reality is that general physicians, they don't have time to invest in additional platforms. In Europe, the average time, when you have a consultation for a physician the average time is around half an hour, in half an hour they must check the patient, they have to fill the medical record, they have to fill prescription, they have to book the next appointment, and the feedback that we have from them is, we don't have time.

We created the tool with the idea of, okay, maybe the physician can work after consultation, after hospital hours, but the reality is, they don't have time because

they are overloaded. And the problem is, connecting this issue with the point from Sandra is that, I guess we have to work together to make sure the patients with diagnosis needs are connected with international networks because general specialists they don't have access to these networks, they don't have time for research, they don't have time to put a patient in a nice tool if they have to invest six hours for each patient. I guess this is a medical need right now, how we can mentor a patient with no diagnosis in Brazil, phenotype and genotype is connected with the international research networks because I am pretty sure the specialist in Brazil is not connected with the other colleagues, most of them for sure they're going to be a specialist, and that if we want to do a tool, a platform for everybody, everybody should be able to use it, not just a specialist.

**Ben Solomon:**

I want to say the COVID pandemic has been terrible, but some good things have come out of it, lessons learned, telehealth is often cited. Another one is the ability for scientists, academicians, patients, lots of folks to work together collaboratively when things really matter. And I think there are lots of lessons that we can take from that experience and apply to many areas of biomedicine, including the topic of AI in rare disease research. I think there is a lot that can be done if we can, just like Julian said, get together, and work as a much larger unit rather than all these little silos, it is a cliché, but I think it is really important.

**Sean Sanders (host):**

*I would like to spend the second half of this webinar focusing more directly on rare diseases, and the application of artificial intelligence in that field, so maybe we can ask and Sylvia I'll come to you with this question. Why do we need AI for rare diseases? What is unique about it? And why is the field of rare diseases, or the group of rare disease, particularly suited to AI analysis?*

**Sylvia Thun:**

The challenge is that we must deal with a lot of data, and if you think about lab data or genome data and images and grey scales and so on, we put the data together and then we create a diagnosis, this is how we work. Doctors work, and right now we have so much more data and we have to memorize huge quantities of digital information and we cannot

do it anymore, and beside that, now we are just getting more and more and more on so many publications, right now in this second, are made here on PubMed, and we have to read the new publications and so on, so we have to use machines to help us, and no. We still need doctors sure, but machines or AI algorithm should just help us to find the right therapy and the right diagnosis. So beside that, we can accelerate diagnosis. Here in Germany, we have like three to five years until one gets a diagnosis, if he or she has a rare disease, and that is too much, and there are so many people who have a rare disease - because rare is not rare.

***“You have more than 7000 different rare diseases that have been described so far, when you consider everything all of them put together, you have a very large part of the population that is being affected.”***

#### **Sandra Brasil:**

I would just like to add to what Sylvia said. Well, on the rare, it is not being rare, because when you consider a rare disease by per se, it is rare and affects a few people, but considering that you have more than 7000 different rare diseases that have been described so far, when you consider everything all of them put together, you have a very large part of the population that is being affected. Really, the time, I think that not having a diagnosis, it is a very big problem, but a miss diagnosis, it is also a very big problem because that patient might have been taking a wrong medication that is not helping or might even harm, I mean not harm the patient, but will not benefit the patient, so I think that we have to address both, not diagnosing the patient, but also misdiagnosing the patient is also very important.

Considering that you have so many rare diseases, I think that artificial intelligence is going to be very important because it is impossible for doctors to memorize and know all of those different rare diseases, so I think if you have a tool that can help in the diagnosis by the recognition of for example, facial characteristics that are specific to a rare disease, I think that it will be of tremendous help because it is impossible for the clinicians to know the symptoms and presentations of all of these disorders.

#### **Ben Solomon:**

If I could just add to that, I think one of the things that these methods can do is really level the playing field, so I won't speak for other countries, maybe you guys have different experiences, but here, if a patient is... I don't want to say lucky enough, but if they are lucky enough to be near a large medical center, an academic medical center with lots of experts, they are more they are more likely to get a diagnosis quicker, they have more access to some of these things. Perhaps one goal of this is that AI will be able to level the playing field, so it doesn't matter where you are in the United States or in other countries, these algorithms will provide access quicker and hopefully cut down the time to diagnosis. Just as importantly, the time to making sure that one is following best management guidelines and connecting with the best experts, and just as importantly, other family groups who often know more than the experts in terms of the best way to manage patients.

***“In Europe, we have 117 medical drugs with orphan drug designation for around 100 conditions. Just 100 conditions, rare conditions, have a treatment with a valid indication that means that most of the rare diseases, they don't have a treatment for the condition.”***

#### **Julian Isla:**

Just to add to that, diagnosis is just part of the medical need when you have a diagnosis, broadly you will need treatment, and unfortunately, if you take the number of our orphan drug designations that we have in Europe, we have 117 medical drugs with orphan drug designation for around 100 conditions. Just 100 conditions, rare conditions, have a treatment with a valid indication that means that most of the rare diseases, they don't have a treatment for the condition. But one important, and I would like to ask to my colleagues what they think about the pairing of precision medicine and rare diseases. Doctor Topol - I really like the books from Doctor Topol - he says that in the future, precision medicine means that everybody will be rare. Because we will be able to stratify our population into the individual level, and when that happens, everybody

will be different, and the border is when rare conditions and prevalent conditions will disappear.

This is a phenomena that we are seeing right now in the other way with European regulation, because we can see how pediatric cancer is using Orphanet regulation to position therapies for cancer. And this is not why we created the Orphanet regulation. Orphanet regulation is created to protect the development of new drugs for cancer. There is a reset. They are in a much more different situation than so many patients with no more options. The same way that their cancer is rare. I am wondering if their rare conditions will be not so rare because if we know more information about people, everybody will be rare, and this distinction won't be valid. I am not sure what my colleagues think about that.

#### **Sandra Brasil:**

I think it is a really interesting point, and also, I think adding on to that, I read an interesting article the other day: for example, if you develop gene therapy that can reach the liver with good results, then that therapy should not only be used for that specific rare disease, it should be open to all rare diseases that have the same issue. But the problem with that is that it implies, once again, communication and collaboration and the willingness to share the results, and products of research. I think that, once again, we need to boost collaboration between groups. In science we hear a lot, “publish or perish,” but in this case you need publications. We need to think about the greater good, which is patients and families, and the impact that not having a treatment has on these populations. I think that we need to put our strengths and our resources all in the same pot and go forward with research.

#### **Ben Solomon:**

I think precision medicine, however you define it, is really exciting and all the new technology and methods I think are greatly going to improve health, both for those affected by rare diseases and others, their family members and others who don't have conditions. What I do worry about, just to be a little bit of a devil's advocate, is this idea of the narcissome. You notice as we talk about the genome, looking at DNA, the transcriptome, looking at RNA, and all these other -omes that we talk about, the microbiome. There has also been this movement saying, “Okay, we are put a lot of these emphases and a lot of these technologies to focus on healthy people and really get to know every single bit of data around them for health reasons.”

That is not a bad thing at all. I don't mean that it is one or the other, but I would not want the emphasis on that to take away from the emphasis on helping patients with severe congenital or any kind of rare or genetic condition where, just like Julian said, there just are not enough therapies and there are not enough ways to manage these patients to help them have better quality of life.

**Sean Sanders (host):**

*I think that is a fascinating point, Julian, thank you for bringing that up, and something that I think is important to think about in the future. I did want to come back to something else that you said about therapies, because we haven't really talked about that. And Sylvia, I would like to put this question to you. In what ways could AI be applied to matching therapies to diseases and especially often drugs or drugs that offer treatment of other diseases that might be applicable to certain rare diseases. Are people thinking about this? Is there any work in that space?*

**Sylvia Thun:**

We are thinking about our pandemic situation and not about the rare disease situation. We had the chance, because we had so much data throughout the whole world, and the FDA asked us here in Germany, our Big Pharma and European Medicine Agency, are there any therapies? Why are so good in the beginning of the pandemic in Germany, is there a chance that your people get other drugs or something so that you don't get COVID? We were wrong, we had the same, as hard as in the US, in Germany as well, but later. And so we can just have a look at the data, which kind of therapy can work for different diseases and even rare diseases. That is one point. The other point is more the research area, and I am not a researcher in the drug development.

**Sandra Brasil:**

Well, just for everyone to have an idea, developing a new drug from the start, from the beginning, it is a very time-consuming and expensive process. The latest estimations tell us that developing a new pharmaceutical compound, a new drug, takes over \$2 billion, so it is very, very expensive. And one area in which artificial intelligence can help with is finding new therapies for rare diseases by looking for therapies that have already been

approved. Or, that are, for example, failed clinical trials and phase two, because for the ones that are not familiar, clinical trials have several phases. In phase one, you test safety in the general population, and in phase two you test efficacy, so if the drug will work in the patient population once you know that it is safe. Sometimes it fails in phase two because it works in the lab, but when you get to people, it does not work. The idea is to use artificial intelligence tools to go through, because now you have several databases with a lot of information on approved drugs, and drugs that have been put on hold.

You have all that data there and you can check if those drugs will be able to work in your disease. But to do that, you need to have a lot of information about your own disorder, about that rare disease. In order to make this approach work, you need to study very well, the rare disease, and have molecular data, data from metabolomics and other omics, so you can incorporate all of that in one unique disease model and see if the information you are retrieving from the drug is able to help the patients. I think that now the biggest problem is having that amount of information from the rare disease to make sure that you have a good disease model.

Because if the disease model is not good, you will not have a good result, or even if you spend the time and money, when you reach the patients, you will not have a good result. I just wanted to point out, because we did a literature review that was published in 2019 about artificial intelligence, the tools for rare diseases. And I just wanted to point out that only one study, only one paper was available, that mentioned artificial intelligence for this drug reposition, because it is the name that we give to the new use for known drugs. I think that this is an area that needs to be approached for rare diseases, and I think that it is a very interesting field that needs to be explored.

**Ben Solomon:**

If I could add just a couple thoughts to that. One is, I encourage everybody to read Sandra's group's paper (Brasil, Sandra et al. "Artificial Intelligence (AI) in Rare Diseases: Is the Future Brighter?" Genes vol. 10,12 978. 27 Nov. 2019, doi:10.3390/genes10120978), which I thought was wonderful and just fascinating review of the literature. It will be interesting just in the years since that was published, I am sure there has so much more that is come out. But one point, going back to the sharing, is that it is important. And I understand this is hard to do in certain contexts, but to make algorithms

and the underlying data sets available that are published in the papers that Sandra is talking about, so that others can look and say, look, I need to make sure that this works. I need to make sure this works on another data set, or I want to build on this. If those algorithms are not available, then, and I understand the reasons for them not to be available, but it hits a brick wall. Sharing in that context is important as well. The other point I want to make is that artificial intelligence, one of the other things it can do, taking a step back from what Sandra was mentioning, is that sometimes it can make these biological connections that we wouldn't realize.

I am not an expert in CDG, in congenital disorders of glycosylation, but there has so many of them, and on the surface, they wouldn't necessarily appear to be biologically connected. As new CDG conditions are discovered or ones that might be CDGs, these methods, these artificial intelligence methods can be used to say, "Okay, maybe these are biologically connected and therefore, maybe this can give us a route to therapeutics or research into ways to directly treat the conditions."

<https://rarediseases.org/rare-diseases/congenital-disorders-of-glycosylation/>

**"Congenital disorders of glycosylation (CDG) is an umbrella term for a rapidly expanding group of over 130 rare genetic, metabolic disorders due to defects in a complex chemical process known as glycosylation. Glycosylation is the process by which sugar 'trees' (glycans) are created, altered and attached to 1000's of proteins or fats (lipids). When these sugar molecules are attached to proteins, they form glycoproteins; when they are attached to lipids, they form glycolipids."**

**Sandra Brasil:**

Yes, definitely. For example, some CDGs can be treated by giving some sugars that doctors are testing. Some CDGs are in the same metabolomic path, so they are connected. If one type of sugar works on one, then they are thinking, "Well, we are testing and now we are experimenting to see if it works on the other patients." I think that is definitely the idea and artificial intelligence will be a tool that can help doing that with other patients in the other CDG types that we have. That will be very interesting.

Going even back at the beginning, we thought that rare diseases were simple conditions. You had one or two mutations that were affecting a gene, and only one gene was responsible for the disease. And now we know that it is not like that, we know that different genes in different pathways, can have impacts on the disease and on the presentations and symptoms that the patients have.

I think that that is very important to consider when we use these artificial intelligence algorithms, we need to have a very clear idea of the general picture. We cannot think about rare diseases as simple diseases. They are very complex, and we need to take that into account and definitely be able to make these connections, because sometimes we think that by being on different metabolic pathways, they are not connected, and by using artificial intelligence too, we might find connections, and that might be the solution for therapy, for example.

***“We cannot think about rare diseases as simple diseases. They are very complex, and we need to take that into account and definitely be able to make these connections, because sometimes we think that by being on different metabolic pathways, they are not connected, and by using artificial intelligence too, we might find connections, and that might be the solution for therapy.”***

**Sean Sanders (host):**

*I would like to come back to data sharing, and I would like to look at it from a couple of different perspectives. The one is, and Julian, I am going to come to you with this question because you work for Microsoft. And the question is, where is the AI research in rare diseases currently taking place? Is it predominantly private companies and laboratories? Is it in academia and hospital systems? And how does that impact the sharing of those algorithms and that data? If you could talk to that.*

**Julian Isla:**

From a Microsoft point of view, our research team is working, not really in rare diseases, but in technology that can help the development of new solutions for people with rare diseases. For example, NLP, natural language processing. This, as Sylvia already mentioned, this is very, very important to move from unstructured medical records into structured information. Microsoft has a team of, a very important team, working on NLP, and right now we have a product, we have a solution that everybody can use. It is called Text Analytics For Healthcare, and this is a service in our cloud that everybody can use. But it is not just Microsoft, you know? My colleagues from Google, they have a similar service, or my colleagues from Amazon, they have medical comprehensive as well. The IT industry is focusing on getting a structure from unstructured medical information, this is the priority for all of us.

This is the first step, but I guess I want to highlight that the problem really is the lack of data, you know? AI is fantastic. We can do a lot of things with AI, but the main problem I see is the lack of data, because this is a problem that the medical field is having for other conditions and for rare diseases, it is much more important. Regarding sharing the data, Microsoft is also working in ways of allowing institutions, physicians to share the data with privacy and security. This is very important, and technology like anamorphic encryption or differential privacy, they are already available, and these allow researchers to use the data with privacy and security. But from the patient point of view, I guess, the system is broken. If we really want to move forward, we must empower patients with the data they own.

**The term “data gravity” refers to the desire to have applications and data attract more applications and data on a network.**

**Why Data Gravity Will Grow Stronger - Forbes**  
<https://www.forbes.com › forbestech-council › 2019/01/14>

Right now, the data gravity and gravity is like a physics phenomena, where the data is going, the data goes to the hospital. This is the gravity, the gravity goes to the hospital. But I guess, we have to change this gravity and allow the data go to the patients, because the data is owned by the patients and we provide to the patients, the mechanisms to store the data and

share the data with the people they want, we will solve a lot of problems. Especially the privacy and sharing problems, because they will be the owners and they will make decisions by the data. And I guess from my point of view a change in the way we are managing the data and changing this gravity is going to be needed if we really want to make a bigger step forward.

**Sylvia Thun:**

I think this is the reason that we need something like an international patient summary on EHR that is interoperable in the whole world, and we are right now working on this international patient summary and on a new addendum for a minimal dataset for our diseases, and this is based on a French work from 10 years ago or something, but we must talk about the semantic interoperability. Because like us Europeans, and especially Germans, we don't want to use centralized databases, but they want to share data, so we have to have a cloud infrastructure and with federated learning methods, so that you can bring the algorithm to the data, and not the other way around. This is one step to build the infrastructure, and the next step is to collaborate on the dataset and the international work here. Hopefully, this will be in the next few years.

**Julian Isla:**

In Europe, I am a little bit disappointed with the speed of innovation in Europe, because I guess we talk a lot, we have to make consensus, we have to put everybody on agreement, but we are not running really fast. Do you think innovation will come from institutions, or will innovation come with a new company, a new entity, a new, a new something new that will be able to transform the way we are doing medicine and we are dealing with rare diseases, you know? I don't know, I don't have an answer. It is again a question for the panel.

**Sylvia Thun:**

Perhaps a good example is our vaccination certificate in Europe, so you can go through Europe and have one certificate, perhaps this is the very beginning of things that are moving faster.

**Julian Isla:**

We have the certificate, and on top of the certificate, there are different rules in different countries. If you go with it to, if I go with my certificate to Germany, the conditions are different than if I go to Italy, you know what I mean?

**Sean Sanders (host):**

*And even more of a challenge internationally.*

**Julian Isla:**

Internationally, it is even more challenging. The technology is not the problem, we have the technology, we have a certificate, we have a very solid infrastructure, but we don't know how to use it with efficacy. This is the main problem, from my point of view.

**Sean Sanders (host):**

*Well, we are running out of time, and I would like to just get to one more question. I did want to say though, just on the data sharing, I would like to mention to the audience something that Sylvia pointed out to me and that I was reading up on and which is the term FAIR for findable, accessible, interoperable and reusable. And I think it is an important guideline for everybody to read and look at and consider, because it seems like that is going to be the foundation for data sharing in the future. But the final question I wanted to put to the panel, just looking to the future, what might this look like if AI can be applied successfully to rare diseases and what might be the main challenges?*

**Julian Isla:**

Maybe I am a little bit repetitive, but the big challenge is going to be changing the way we are collecting data, and how we are putting this data, and that these data sets are available to researchers and the medical community. This is still an issue that and we need to work together as citizens to be able to create with data sets, because the data set are the foundation of good AI models.

**Sylvia Thun:**

I think the main challenge is to educate our young people and our doctors to use AI in a very good manner.

*“I think the main challenge is to educate our young people and our doctors to use AI in a very good manner.”*

**Sandra Brasil:**

Yes, I agree with Sylvia. I think that education will be a challenge. And we need to focus on that, and for the patients to educate the patients and the patient communities that for them to know what artificial intelligence is, how it can be used, what are the advantages. And, with the data, as Julian mentioned, which is the important data, how we need to collect it. Because when patients are educated, they are empowered, and they can move forward. And they can think about different solutions, and they can partner up and they can make their own decisions and help push forward the research. I think that we must shift the balance - not the power balance - but the information that there is from researchers and clinicians and give them to the patients also.

*“When patients are educated, they are empowered, and they can move forward. And they can think about different solutions, and they can partner up and they can make their own decisions and help push forward the research.”*

**Ben Solomon:**

I am optimistic. I think if I look at the younger generations, my kids and other folks, they're much savvier and much more comfortable with these technologies and with some of these big issues, like data sharing. They look at us all tied in knots about our concerns about data sharing, and they say, “No, no, no, this can be a really good thing.” So hopefully, there'll be a generational shift. I expect there will be, and so some of these things will solve themselves, including with some technological solutions.

I am a pediatric geneticist by training and so what I hope in the future, I don't know if it is going to be five, 10, 20 years, it'll be a process, but there'll be a lot more artificial intelligence and related technologies that are supporting what we are doing. So instead of me, when I see a patient, going to the computer and Googling or using a fancier version of Google, some database, to look up their symptoms and try to figure out treatments or try to figure out what they have and decide if I should send a genome on them or enroll them in a research study, there'll be a lot of stuff under the surface that will help underlie

and underpin what I do, and make it faster and easier, so that I, as a doctor, am not spending all my time when I am supposed to be talking to a patient and a family, typing on the computer and writing my medical note and looking things up. But I can talk to patients and families. Hopefully, a lot of these technologies will allow us to get back more and more to human contact, and I think that is the reason all of us went into, in various ways, into rare disease research and clinical medicine.

## Journal Club

### Article of the month

#### Artificial Intelligence (AI) in Rare Diseases: Is the Future Brighter?

Brasil S, Pascoal C, Francisco R, Dos Reis Ferreira V, Videira PA, Valadão AG. Artificial Intelligence (AI) in Rare Diseases: Is the Future Brighter? Genes (Basel). 2019 Nov 27;10(12):978.

Free at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6947640/>

The authors discuss how artificial intelligence can help in rare disease diagnosis and treatment

1. Of 7,000 rare diseases only 5% have a treatment.
2. The amount of data in medicine is hugely increasing.
3. We must learn how to understand, analyze, collect and analyze all this information.
4. Artificial intelligence can help.
5. Artificial intelligence is already being used in medical research, diagnosis, and clinical trials and for discovering new treatments in rare diseases.
6. In future, artificial intelligence could boost diagnosis and treatment for rare diseases.

A detailed analysis appears at: <https://www.mdpi.com/2073-4425/10/12/978/s1>

## History Corner

### The remarkable history of Leprosy by Florian Delval

Everyone at some point in time has heard of leprosy. It is indeed mentioned on numerous occasions in biblical or historical contexts, or in great cinema classics such as *The Ten Commandments*, *Ben Hur* or *Papillon*. But do we really know what causes this disease or the course it takes? Stemming as a source of stigma for those affected throughout its existence, leprosy has now become a distant memory in the Western world.

Yet, according to the World Health Organization (WHO), there are between 150,000 and 250,000 novel infections each year. The WHO also states that there are currently between 2 and 3 million lepers worldwide. In 2015, three countries accounted for 80% of new annual cases worldwide: India (60%), Brazil (13%), Indonesia (8%). Leprosy remains endemic in many other countries such as Angola, Madagascar, Central African Republic or Tanzania.

It is a chronic infectious disease caused by the bacillus *Mycobacterium leprae* – a bacterial disease that mainly affects the skin, peripheral nerves, the upper respiratory tract lining, and eyes. The bacillus is likely transmitted via droplets, from the nose and mouth, during close and frequent contact with untreated cases. Incubation is often described as a long and silent process, with a duration of approximately 5 years. Sometimes, symptoms of leprosy can appear 20 years following exposure. Symptoms start with the appearance of depigmented or pinkish skin patches and may be preceded by tingling sensations. Skin anaesthesia is also present: the patient no longer feels pain or heat. Leprosy is not very contagious, but can be mutilating, sometimes even paralyzing and disabling.

It is also characterized by a deterioration in physical appearance of the patient, and it is this deterioration in physical appearance and its unique characteristics that have made leprosy such a stigmatizing disease throughout history. It was therefore a source of exclusion due to its strict isolation of infected individuals. It appears difficult to estimate the number of people who have been affected by leprosy throughout history. We know that in the heart of the Middle Ages, there were nearly 20,000 places in Europe where the sick was isolated and treated: leper houses.

Throughout the duration of these three episodes, we will focus on a fascinating study carried out by the Institut Pasteur in Paris. As a result of their research, an article was published in Science magazine, entitled “On the origin of leprosy.” Researchers involved in this study asked themselves the following question: “is leprosy a disease of a single strain, which gradually spread around the world through migration, or did it appear in different places, randomly and independently, across the ages?” This series also discusses identification of the bacillus *Mycobacterium Leprae* by Norwegian doctor Gerhard Armauer Hansen, and the subsequent race for treatments, before ending this discussion with an overview of stigma and care which has accompanied patients over the ages.

## What's up?

### Highlights of the month

#### Movies to move you by Morgan Packer

Artificial intelligence has changed the landscape for not only modern technology and our world, but also for rare disease diagnosis and understanding. Here are a few of my film recommendations that focus on either artificial intelligence or rare diseases - worth having a watch:

1. A classic Steven Spielberg film, ***A.I. Artificial Intelligence***, narrates the story of David, an 11-year-old boy (starring Haley Joel Osment) who is the first robotic child programmed to love other humans.
2. Spike Jonze film ***Her*** starring Joaquin Phoenix (Theodore Twombly) and Scarlett Johansson (Samantha) focuses on the use of AI to create a virtual vocal assistant. Samantha, this virtual AI-created voice, accompanies Theodore Twombly throughout his daily life, eventually evolving into a romantic relationship. This film highlights the complexities of AI, relationships between humans and AI, and ethics behind this evolving technology.

3. One special mention is ***Disorder – The Rare Disease Film Festival***, taking place every other year in various worldwide locations, including San Francisco, Italy, Germany, Japan, Miami, Philadelphia, Sioux Falls and Boston. While the next edition of Disorder – The Rare Disease Film Festival, is yet to be scheduled due to Covid-19, many of the festivals' previous films are available for free online here: <https://www.rarediseasefilmfestival.com>

### Manga with the French Paralympic Sports Committee

All our publications  
are available at



### Boccia: the dream to win the gold medal in 2024

Listen to our podcasts  
series on health



(in French)

### Webinar: Intelligent detection and diagnosis of rare diseases

Listen to our webinars  
AAAS/ Science\_Fondation Ipsen



***Webinars:***



***Podcasts:***



***Books:***



Fondation Ipsen  
65, quai Georges Gorse  
92 650 Boulogne-Billancourt Cedex  
France  
**[www.fondation-ipsen.org](http://www.fondation-ipsen.org)**



**Contact:**  
[fondation@ipsen.com](mailto:fondation@ipsen.com)