The Rare Disease Gazette

Conversations with the world's experts about rare disease

Issue #31 | October 2025

HUMANS & AI
COLLABORATING TO SOLVE RARE DISEASE CHALLENGES

FONDATION IPSEN PRESS

Editorial

by **James A. Levine**
MD, PhD, Président, Fondation Ipsen
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 healthcare. 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.

The Rare Disease Gazette is a magazine that broadcasts these discussions.

James Levine

FONDATION IPSEN

DON'T MISS!

The Conversation

Experts of the month: Erika Berg, Ph.D, hosts a conversation with world's experts about rare disease: Humans and AI Collaborating to Solve Rare Disease Challenges: Opportunities and Pitfalls

The Conversation

Experts of the month

Danton Char, M.D.
(Stanford University, Stanford, CA)
Tim Guilliams, Ph.D.
(Healx, Cambridge, England)
Sherri Rose, Ph.D.
(Stanford University, Stanford, CA)
Chunhua Weng, Ph.D.
(Columbia University, New York, NY)

Erika Gebel Berg, Ph.D.
(Science/AAAS, Washington, DC; moderator)

"I believe AI can, on the one hand, empower the rare disease communities, while also empowering the more common disease communities and the general public. In doing so, it can help free up more time within healthcare systems so that nurses and doctors can then focus on the more complex cases, many of which are in the rare disease space."

– Tim Guilliams

Erika Berg (host):
Today we will explore how AI is being integrated into rare disease care and research, discussing its applications in analyzing medical records to assist in diagnoses, accelerating drug discovery, and evaluating insurance claims. We will also address the potential pitfalls of relying on AI, such as algorithmic bias, data privacy concerns, and the need for human oversight to ensure accuracy and fairness. The discussion is going to examine the delicate balance between leveraging AI's power and ensuring that the human element remains central to decision making in rare disease care. I would now like to take the opportunity to welcome a brilliant panel and give each of them a chance to say hello and introduce themselves.

Chunhua Weng:
I am a faculty member in the Department of Biomedical Informatics at Columbia University. My training background is in computer science and health informatics. I have been studying medical records, particularly electronic medical records, and how to leverage them for clinical phenotyping, cohort identification, and knowledge discovery. We have been looking at how to leverage nuanced, large-scale phenotypes in medical records that can be used to drive gene prioritization and rare disease diagnosis.

Sherri Rose:
I am a professor of health policy and director of the Health Policy Data Science Lab at Stanford University. My main research focuses on developing and integrating statistical machine learning and AI approaches to improve human health, including methods for rare outcomes. I have a PhD in biostatistics. Within health policy, I work on algorithms in healthcare, including risk adjustment and health program evaluation. Additionally, I teach a course on reproducible research.

Tim Guilliams:
I am the co-founder and CEO of Healx. We are an AI-enabled rare disease biotech company from Cambridge in the UK. I am also the co-founder of a rare disease charity called the Cambridge Rare Disease Network; I started those two organizations when finishing my PhD at the chemistry department at Cambridge University.

Danton Char:
I am one of the pediatric cardiac anesthesiologists at Lucile Packard Children's Hospital at Stanford. I provide anesthesia for children and young adults with heart disease. My early research looked at the implementation of whole genome sequencing into the care of critically ill children and the ethical issues that came up with that. I have become a medical ethicist in my research and through that work I have begun to look at big data ethical issues which led to my work on AI. I am now part of Stanford's oversight team for all the AI being considered for clinical rollout across Stanford Healthcare.

Erika Berg (host):
You have all engaged with AI in different domains of clinical research in the healthcare world. How recently have you seen this technology, which still feels new but has become mainstream, and what have been the drivers of this mainstream adoption?

Chunhua Weng:
It has been an exciting time over the past year and a half. We have seen a lot of enthusiasm towards generative AI, ChatGPT, and other large language models. They are being used to extract human phenotype concepts from medical records and then converting these phenotype concepts into ontologies, such as the Human Phenotype Ontology. This enables semantic interoperability across the medical records and other knowledge bases, allowing us to analyze the phenotype-genotype associations. This is very exciting, because right now the ability to use large language models to extract and normalize these phenotypic concepts is reaching a new level. They are achieving better accuracy and better reasoning abilities. We see the enthusiasm among the trainees as well as clinicians and clinical researchers. Everybody is excited about exploring the new opportunities to leverage this new technology for phenotype-driven rare disease diagnosis. In my own lab we are also developing algorithms for phenotype-driven genetic test recommendations for rare disease patients. As many people are aware, it is a long odyssey for rare disease patients to reach a diagnosis. Along the way, they need to go through many genetic tests, many of which are not relevant and can be very expensive. Some of them may not be covered by their health insurance. We look at the rich information in the medical records, particularly the phenotype concepts in their genetic counseling notes and other clinical notes

and find out they are actually very valuable for giving recommendations on what genetic tests will be the most relevant for the phenotypic profile of these patients. Our next step involves looking at the economic analysis to see the economic value of these phenotype-driven genetic test recommendations. It is very exciting and we are very happy and optimistic about this new era of using AI for rare disease diagnosis and genetic test recommendations.

Sherri Rose:
I think it depends on the research area you work in and what we consider AI. If we are focusing on things like generative AI and large language models (LLMs), over the last few years we have seen a lot of progress in these tools, but a lot of what is claimed to be AI is not what we would have previously called AI. I think that is important to highlight. We used to refer to some of these tools, those that are not large language models, as statistical machine learning or simply machine learning. I think if we are bundling all these tools together, we need to remember the limits of some of these tools, that we need to be aware of, and that we have been here before. So, some people who are newer to the AI space were not here for the previous hype cycles. I think that is something that we should really think about, especially when we are looking at healthcare data, because I have been working in this space for a long time and everything starts with, "AI is going to revolutionize healthcare," and we have been here before.

"On the clinical side, we have not seen much of a dramatic change in patient-focused care from AI tools. What we have largely seen is patients conducting AI research at home regarding clinical conditions and patients coming into clinic with information provided by AI tools, including ChatGPT."

Danton Char:
I think I have a guarded optimism. On the clinical side, we have not seen much of a dramatic change in patient-focused care from AI tools. Largely what we have seen is patients doing AI research at home about clinical conditions and coming in with information provided by AI tools, including ChatGPT. That is a mixed bag, because one of the things we must do is sort out which of the information they have received at home is relevant to the care that they are pursuing, and that is a challenge. I also agree with Sherri that we went through all of this before with the electronic health record (EHR). There is a great deal of excitement, and I would say many clinicians say that we still deliver healthcare and we have the EHR. I have concerns that it will be the same with AI: that we deliver healthcare and we have AI-enabled EHR tools.

"What is happening right now is that in drug discovery, the current paradigm is called the target-based paradigm and typically involves a single target. Biology, however, is incredibly complex, with 22,000 genes and more than 60,000 proteins. So, there is an opportunity to now use AI, but also other statistical methods and computational biology, to understand the rare disease biology in a different way."

Tim Guilliams:
We are focused on drug discovery and development. From our perspective, we really see an opportunity and already an impact to understand rare disease biology in a more complex way, where you do not have to dumb down the rare disease biology to a single protein target, for example. What is happening right now is that in drug discovery, the current paradigm is called the target-based paradigm and typically involves a single target. Biology, however, is incredibly complex, with 22,000 genes and more than 60,000 proteins. So, there is an opportunity to now use AI, but also other statistical methods and computational biology, to understand the rare disease biology in a different way. We can then predict the right, in our case, small molecules, that work on that disease and can effectively restore your disease state to a healthy state, but more at a systems level instead of a single target pathway protein level. I think that is where we see some major breakthroughs, at least preclinically. We now have drugs in Phase 2 clinical trials, and I think within perhaps three to five years, we will have the first AI-discovered drugs reaching approval and the market. So, we are still a few years away from real patient impact.

Erika Berg (host):
Sherri, you mentioned the hype cycle earlier and the significant enthusiasm about what AI can deliver. Are you at all concerned that this enthusiasm and these rapid developments in AI solutions may be outpacing the evaluation of risks and ethical considerations?

"When we look at the proliferation of AI articles in healthcare, a recent review found that only 5% of the LLM studies actually used real patient care data. So, this is a significant problem if so many of the studies are looking at small sample sizes with structured text vignettes."

Sherri Rose:
Absolutely. One of the things that I feel very strongly about is that more rigorous evaluations of these tools are needed, which include looking at effectiveness, risks, and ethics. This was previously the mandate, and that was when we were only focusing on whether it was effective. Now, a lot of the time we are not even doing that. Fifteen years ago, when I would develop new machine learning algorithms, the most common refrain was that no one will ever use this. The bar for proving the tool was worthwhile was extremely high, as it should be. I think one idea that I would like to ground in the conversation is that just because a tool exists does not mean that it will necessarily improve health or enhance decision making. Our evaluations really need to focus on multiple evaluation metrics, external data sources, generalizability concerns, as well as data and algorithmic bias evaluation. When we look at the proliferation of AI articles in healthcare, a recent review found that only 5% of the

LLM studies used real patient care data. So, this is a significant problem if so many of the studies are looking at small sample sizes with structured text vignettes. A common response that I get to these types of criticisms is that the tools will get better over time. I like to then reply by saying, "Then why would we not wait until they demonstrate superior performance?" Again, that used to be a basic bar, and it is so important in healthcare for us to have a lot of confidence in the performance of these tools.

Erika Berg (host):

Danton, from more of a clinical perspective, I am curious about your thoughts on this. Are we moving too quickly and are we evaluating these tools appropriately?

> ***"I think we have now seen several real-world implementations of tools that we envisioned being very helpful to healthcare delivery, yet we struggle to come up with a model where we can pay for them or pay for their use."***

Danton Char:

I agree with what Sherri said. I also would call out something that Chunhua had said earlier, which is that we are still working very hard to prove the economic value of these tools. I think we have now seen several real-world implementations of tools that we envisioned being very helpful to healthcare delivery, yet we struggle to come up with a model where we can pay for them or pay for their use. One example would be predictive tools to screen the poorly accessed patient populations for serious diseases, such as diabetic retinopathy. Although there are societal costs saved by preventing disease, it is very difficult to convince health systems to pay for tools for patient populations that are very low reimbursement. This is particularly pronounced if the tools are extremely expensive and the accuracy required to give meaningful care through these tools is very expensive to deliver.

Erika Berg (host):

I feel like we have given a good overview of the current landscape, I would now like to talk more specifically about AI in the rare disease space. Chunhua, your team has been exploring opportunities to apply deep learning and other AI methods to facilitate early and accurate diagnosis of rare diseases, which is a big problem in the rare disease space. What are some of the exciting opportunities you see there, particularly as early-life genomic sequencing becomes more prevalent, and where are we with respect to these tools being used in the clinical decision support system?

Chunhua Weng:

We have been working on this in collaboration with two experts in rare disease research. The first one is Dr. Kai Wang at the University of Pennsylvania and the Children's Hospital of Philadelphia, and the second expert is Dr. Wendy Chung at Boston Children's Hospital who specializes in rare diseases. For almost one decade now, we have been looking into large-scale EHR data and developing a serious algorithm for clinical phenotype extraction from the notes, or EHR narratives. We then use these extracted phenotypes to map them to Human Phenotype Ontology and then to look at how they can be useful for gene prioritization and matching a patient's phenotypic profile to known cases in the literature or in published genotype-phenotype knowledge bases to help with the diagnosis. We also use this extracted phenotype to facilitate a genetic test recommendation, as I mentioned earlier. So far, we have been evaluating this algorithm in a lab setting, essentially using benchmark data sets that are available within the research community. Ideally, what we want to evaluate is the impact on patient diagnosis: how many patients we can diagnose and how much we can shorten the time to diagnosis.

> ***"With the near-universal adoption of EHRs, the rich phenotypic data they contain will be extremely valuable."***

I think the major barrier there is what the implementation science people have been talking about extensively. For the algorithm to be able to run within the current EHR ecosystem it will involve a significant undertaking. While we work in the academic setting as researchers, EHR systems are usually owned and managed by hospitals and healthcare facilities. For the algorithm to run in a clinical practice setting, we have a long way to go. We have been looking into integrating informatics and AI algorithms into a clinical setting since 2011, and we feel like we have not made a lot of progress. There is still a lot of work that remains to be done. Despite these challenges, however, we remain optimistic. With the near-universal adoption of EHRs, the rich phenotypic data they contain will be extremely valuable. In addition, the newer algorithms, such as federated learning, are likely to play an increasingly greater role. It will enable multiple institutions to collaborate with each other without compromising patient privacy, while still protecting patient data security. This approach enables collaborative rare disease phenotype discovery. One example of our recent work is OARD, which stands for Open Annotation for Rare Disease. In this project, we parsed all the EHR data of about 6.5 million patients at Columbia University, spanning multiple decades, and including a majority of the clinical notes. We also incorporated data from all the patients at the Children's Hospital of Philadelphia. By leveraging these two institutions, we were able to conduct cross validation and external evaluation and build a knowledge graph showing the associations between phenotypes and rare diseases. We are hoping that more initiatives like this will emerge, allowing for the contribution of more data for rare disease research. Through this work we have identified a lot of novel phenotype-rare disease associations that were previously unknown in the literature. The next important step is to develop robust verification methods to determine which of the newly identified associations are truly meaningful and can help us to support earlier and more accurate diagnosis.

Erika Berg (host):

Tim, I would like to talk about Healx, which is somewhat unusual as a drug company, as you are principally focused on targeting rare genetic disorders. Can you talk about how you and your colleagues at Healx are leveraging AI as a tool to accelerate discovery and development of drugs specifically for rare disease?

"From a company perspective, whenever we start a new program, we establish partnerships with rare disease foundations and charities in that specific area. This is because they are the true rare disease experts: they understand the unmet needs, they help us access missing data models, and they help us generate the required data that we need in our algorithms."

Tim Guilliams:
We are an AI-enabled rare disease biotech and there are two important components to our approach. Basically, we are rare first and AI first, but we think about AI in a broader sense. What is so great about this panel today is that everyone is trying to use AI to do something positive for the world, such as helping rare disease patients. The opportunity to use the latest tools to have a positive impact cannot be overstated. From a company perspective, whenever we start a new program, we establish partnerships with rare disease foundations and charities in that specific area. This is because they are the true rare disease experts: they understand the unmet needs, they help us access missing data models, and they help us generate the required data that we need in our algorithms. We also hire rare disease parents to help drive projects forward. I think this is incredibly valuable, and I will share an example in a moment. From an AI, machine learning or computational biology perspective, we are focused on first understanding the complex rare disease biology. Typically, there is no validated target for a rare disease. As a result, you cannot simply apply the traditional target-based approach without a validated target. The other thing that we do is focus on the chemistry, so trying to match the right small molecule to a newly understood biology. We aim to develop therapeutic breakthroughs, not just a "me too" drug aimed at a target that already exists. I would like to share a quick example. One of our Phase 2 clinical programs focuses on neurofibromatosis Type 1 (NF1), which is a rare tumor predisposition syndrome. Patients develop benign tumors that grow along the nerves and require long-term treatment. Currently, the approved treatments are MEK (mitogen-activated protein kinase kinase) inhibitors, and while that is fantastic for patients to have that option, MEK inhibitors come with significant side effects. For long-term use this can be suboptimal for the patients. As a result, patients, clinicians, and families typically try to avoid taking these drugs until it is necessary as the trade-off between side effects and the clinical benefits for a number of years is difficult to justify.

"This is what we mean when we think about applying AI to drug discovery. It is really about finding therapeutic breakthroughs, new mechanisms, and first-in-class solutions that will make a difference to patients."

When we started the program in 2019, and we are now at phase 2 in 2025, we partnered with the Children's Tumor Foundation and what we heard from them was that the NF1 community was looking for a treatment that would be safe for chronic, long-term use and that was not a kinase inhibitor or a traditional cancer drug. Using our AI tools and creative innovative pharmacologists, we set out to find a molecule with a strong safety profile and minimal side effects, and that also had antitumor and anticancer properties. We also aimed for nerve-specific accumulation, because these tumors grow along the nerves. Our phase 1 data demonstrated an excellent safety profile, which is fantastic. Preclinically, we showed that this small molecule accumulates specifically around the nerve, with a 90-hour half-life around the nerve, compared to a two-hour half-life in the rest of the body. So, it clears rapidly from everywhere else but accumulates in the nerve. It also has very strong anticancer and antitumor properties which reached maximum efficacy in the preclinical models. We would not have achieved this without, first of all, the Children Tumor Foundation and the NF Clinical Trial Consortium, who are incredibly committed and dedicated to finding solutions for patients. We also would not have found this if we had followed the traditional target-based approach. We therefore use AI to look at transcriptomic, phosphoproteomic, and metabolomic data to try and identify first-in-class mechanisms, as well as molecules with a higher affinity for membrane potential, which is important if you want the molecule to reach the nerve. This could potentially lead to a treatment paradigm shift for patients, allowing a potential treatment to be initiated earlier, supporting tumor prevention or secondary tumor prevention, and offering an option that is safe for chronic use. This is what we mean when we think about applying AI to drug discovery. It is about finding therapeutic breakthroughs, new mechanisms, and first-in-class solutions that will make a difference to patients. I want to thank the Children's Tumor Foundation, The NF Clinical Trial Consortium and everyone in the team. It is a very exciting time as we are now in phase 2, and the trial is going well so far.

Erika Berg (host):
Sherri, there has been a lot of discussion about the use of AI to expedite healthcare coverage decisions. One example is in the context of prior authorization. You have looked at both the potential advantages as well as opportunities for harm and abuse. What is your current perspective on the appropriate use of AI by payers, and what are your foremost concerns?

Sherri Rose:
This is an example that is quite dissimilar from what Tim was just talking about as here there is a very primary concern about harms and abuse. We really need humans as a bulwark against error for these algorithms. We have seen class action lawsuits against insurers over algorithms that are designed to restrict post-acute care as well as wrongful denials of prior authorization requests. This has been discussed at the level of congressional committees, for example. A challenge here is that the algorithms currently in use are not performing well. This is true whether the approach uses AI or a checklist for prior authorization. When we think about machine learning and AI systems that rely on these data types and the information being used to make these decisions, often a complex tool will not necessarily be better. Sometimes it is the additional data that is included that might potentially improve predictions. However, this is not always the case. Then, there is the issue of having the same data recorded relatively uniformly for patients, which is typically not the case. So, this question is an AI question, but it is also a transparency question. When we think about patients, how is the

decision being made? AI can make that part worse because it is not understood what the basis is for the denial, and a lot of these AI-based denials are being overturned.

Erika Berg (host):
Danton, I was hoping you could speak about AI as an opportunity to deliver greater value for the rare disease community versus the general population. Does AI have some special opportunity in the rare disease community to improve care, compared to more general healthcare concerns?

"I think that there are a number of qualitative outcomes that communities have been quite excited about, namely, the way social media and AI-driven social media has allowed rare disease communities to connect with each other and find support. I think that those communities have developed strength from that connection, along with a stronger sense of identity and mutual support."

Danton Char:
I do not know if I have any uniquely ethical challenges that would be inherent to that, but I think it will depend on the quality of the data available for a given rare disease community and how well a tool can analyze that data, and what it is able to deliver. I think that there are several qualitative outcomes that communities have been quite excited about, namely, the way social media and AI-driven social media has allowed rare disease communities to connect with each other and find support. I think that those communities have developed strength from that connection, along with a stronger sense of identity and mutual support. I also think that we have all encountered downsides of these social media connections too. There have been some strong instances of misinformation being conveyed within communities, particularly those that were a little more vulnerable from a healthcare perspective. That is something that probably needs to be addressed.

"I believe AI can, on the one hand, empower the rare disease communities, while also empowering the more common disease communities and the general public. In doing so, it can help free up more time within healthcare systems so that nurses and doctors can then focus on the more complex cases, many of which are in the rare disease space."

Tim Guilliams:
From our perspective, because we work very closely with foundations, charities, patients and their families, we think AI can really empower people to start understanding their disease in a different way, potentially receive a faster diagnosis and ultimately bring a sense of agency back to patients, families and foundations. It can also help them begin to use some of the tools to discover new biology and potential molecules that could be used for their disease. Through these partnerships, we see that they come with suggestions for potential drugs that could work. I think that there is a democratization of innovation for these populations that have been largely ignored for a very long time, and it has the potential to make a huge impact. Another aspect that I believe could be helpful, although I am not a medical doctor myself, but I have a mother who is a doctor, is of the ability to free up expert clinicians from managing the general public and more common diseases. Validated AI tools that are sufficiently reliable can help point people in the right direction, allowing expert clinicians and nurses to focus on the more complex cases. I believe AI can, on the one hand, empower the rare disease communities, while also empowering the more common disease communities and the general public. In doing so, it can help free up more time within healthcare systems so that nurses and doctors can then focus on the more complex cases, many of which are in the rare disease space.

Erika Berg (host):
Danton, you had mentioned earlier that the use of AI in the healthcare setting is still slow to be adopted and that you are seeing a lot on the patient side with patients using AI to explore their conditions. How much enthusiasm or skepticism are you seeing among the broader clinical community (physicians and healthcare providers) around the adoption of AI, both in the clinic and in the patient setting.

"A vision of the future is that AI tools could potentially be providing that surveillance function, with your clinicians acting as alert responders. So, when the AI determines an anomaly, a clinician would step in for deeper surveillance."

Danton Char:
I think, like all of society, it is mixed. There are some people who are very enthusiastic about the potential for AI. There are also several areas of strain. There are some visions of changing the way labor functions within healthcare that could be AI-enabled. Right now, your main point of surveillance is your clinician, your bedside nurse or the doctor you see who is assessing you and determining the status of your health. A vision of the future is that AI tools could potentially be providing that surveillance function, with your clinicians acting as alert responders. So, when the AI determines an anomaly, a clinician would step in for deeper surveillance. That gives rise to several potential challenges that we have not yet worked out. I think Sherri alluded to it already. We do not really know that we have any metrics to validate how effective these tools are, or how good those alerts really are. We must look at other high-stake sister industries for comparison. They are not the same, and history does not always repeat itself, but as the saying goes, it often rhymes. We have certainly seen this within the aviation industry, for example, with the Boeing MCAS (Maneuvering Characteristics Augmentation System) question. There has been a big narrative about the challenges Boeing has faced, but one of

the challenges that came out of that discussion was whether pilots had adequate expertise to determine when the tool was malfunctioning, to turn it off, and to reassert control over the aircraft. There are all these concerns about deskilling that come up with changing a workflow. It also touches the very nature of labor itself, with questions and concern about how these tools will affect the job, the reasons why people enter the profession, the altruism that motivates them, and the stability of the profession.

Erika Berg (host):

AI models are only as good as the data that is used to construct them, and they require validation. Chunhua, what kinds of data resources are proving most powerful in enabling the type of work you are doing, or other clinical uses of AI, particularly in the domain of rare diseases? Are you seeing increased efforts to establish these data resources specifically for AI applications?

"More recently, in the research community, we have also been looking into multimodal AI, which combines genetic data, clinical data, and imaging data to support better diagnosis."

Chunhua Weng:

Yes, I think multiple data types are valuable for developing AI solutions for rare diseases. I mentioned the large-scale EHR data, the medical records data, and particularly the nuanced narrative notes in the EHRs. More recently, in the research community, we have also been looking into multimodal AI, which combines genetic data, clinical data, and imaging data to support better diagnosis. For example, for many rare disease patients, facial pictures and the images of their facial features have proven to be very valuable, when combined with other types of their data, for diagnosis. Dr. Kai Wang and my lab have been collaborating in this space, and our early results show some promise in combining these data for diagnostic purposes.

Another important data type is from social media, which Tim and others already mentioned earlier. In the NIH funded Biomedical Data Translator Program a couple of years ago, we saw great success in analyzing posts from rare disease patient community discussion forums. Using this information, they were able to identify potential treatments for some rare disease patients, even before they received a diagnosis, simply by comparing the similarities of their phenotypes and diseases and then reasoning over possible treatment options. Another data source would be the case reports that already exist in PubMed, which is the publicly available literature database for medical research. We have been looking across all of these data types. Recently we also saw an announcement at NIH about plans to create a comprehensive integrated database to support rare disease research. So that is very exciting. As I mentioned, rare disease research often suffers from its small sample sizes. Very often, each institution or organization may only have a limited number of cases. With the maturation of federated learning methodologies, however, I think the trend is moving toward enabling network-wide collaboration. This allows multiple institutions to leverage federated learning, while keeping their data in their own database and behind their own firewall, yet still being able to share and aggregate evidence to enable collaborative knowledge discovery and treatment discovery for rare diseases. So that is another exciting new trend. Also, within the informatics community, we are still working on addressing data missingness and data bias. Clinical data bias really needs to be based on the better understanding of the healthcare process biases. For example, my collaborator Dr. Wendy Chung shares with me that, very often, data missingness depends on the patients who are sitting in front of you. If the patient is an immigrant and English is not their native language, they may be more reserved, and the doctors or clinicians may be afraid to ask a lot of questions or collect a lot of information. This can influence how much information and how much phenotypic data is collected from different patients. All these issues need to be considered when we think about data quality and data missingness. I live in Manhattan, where data fragmentation is another big issue. Patients move around more frequently than many other places. How do we bring together data scattered across different hospitals and databases to build a complete patient profile? These issues still require a lot of careful work.

Erika Berg (host):

Tim, where are you getting your data from? Is it from similar sources or what is feeding into your algorithms?

"In rare diseases, the challenge is basically data scarcity, data gaps and, in some cases, the quality of the data. It is therefore incredibly important to map this out. When we start a program, we always start with a data landscape analysis. What is the data available, what is the quality of the data, and what is missing?"

Tim Guilliams:

I would first like to echo what Chunhua just shared. In rare diseases, the challenge is basically data scarcity, data gaps and, in some cases, the quality of the data. It is therefore incredibly important to map this out. When we start a program, we always start with a data landscape analysis. What is the data available, what is the quality of the data, and what is missing? Then I go back to the team and to the foundations and say, "If you had a magic wand, what data would you generate?" There is always a long wish list, and we must make trade-offs and choose the areas that will give us the greatest insights as we are here to find novel rare disease mechanisms and then small molecules that can act on those mechanisms. Because we do not live in a very data-rich world, we need to make the most of what there is and avoid approaching the problem with blinders on, for example, assuming that because everyone else working on NF1 is developing a kinase inhibitor, we need to do more of the same. Instead, once we have this data landscape and analysis completed, we can try and analyze all of it and see if any of it can provide new insights or reveal something novel. Then, we try to observe patterns across data types. For NF1, for example, this is how we arrived at our breakthrough. We also make a deliberate effort to learn from places where there is more data, namely the common diseases. We use algorithms

to try and derive insights from common diseases and see how this can help us fill in data gaps from rare diseases where we are missing data and cannot generate it. The challenge here is a bit different than within the common disease space. From a biological perspective, the richest sources of insight come from your single cell transcriptomic and proteomic data, so those very rich biological data sets, ideally from patient derived samples or cell lines, which are hard to obtain. This is where clinicians and foundations can play a critical role in helping to generate, collect, and share these data, and it is often where we see the biggest breakthroughs. All the other data types that Chunhua mentioned are also extremely useful. The real-world evidence, the phenotypic data, and there are a number of other data types.

In today's world, the algorithm itself and the analysis are less of a bottleneck. There are still things that need to be done correctly, and challenges remain, but the bigger issue is the quality of data that you can generate, rely on, and then analyze in an objective way with less bias. I think that there is some real progress happening in that space, particularly as we approach it in a multimodal way, where you can have scans, genetic data, transcriptomic data, and real-world evidence and analyze this across, what we call, the "lasagna" data stack of layers. Then whatever gives us an insight, or a breakthrough is a win. Whether it is generated by an LLM or another type of model, it does not really matter to the patients.

Erika Berg (host):
Sherri, there are many applications of data that are predicated on the analysis and interpretation of real-world medical data and other personal information. What are some of the practical concerns with respect to patient protection and privacy, and other similar issues?

"We already know that for the general population it takes maybe three pieces of data, or sometimes even fewer, from an EHR for a person to be completely identifiable. With rare disease patients, it could be a single binary flag for their health condition combined with a piece of their zip code, and then the rest of their data is identifiable. So, there are many questions around patient protection and privacy that have not yet been answered, and I think this is particularly problematic for the rare disease community."

Sherri Rose:
I will dovetail on some of the comments from Chunhua and Tim. The data is paramount, and AI and machine learning is secondary to the data. When we think about what it means for patient protection and privacy, there are several concerns we could think about. One of the topics that I know Danton has worked on is whether patients are going to be notified that an AI tool is being used? For example, some large medical centers are using AI-based transcription and are not telling patients that the visits are being recorded. They have decided that it is not required. When I sign onto Zoom and am being recorded, it notifies me and I have to click that I agree. Yet, if I am going to a clinician's office and that information is being recorded, a decision has been made in certain centers that the patient does not need to know. This raises further questions about where that data is stored and for how long. A lot of the time there are assurances that it is stored locally and then deleted. However, as we have seen with data breach after data breach, what does being deleted mean, especially for individuals in the rare disease community? We already know that for the general population it takes maybe three pieces of data, or sometimes even fewer, from an EHR for a person to be completely identifiable. With rare disease patients, it could be a single binary flag for their health condition combined with a piece of their zip code, and then the rest of their data is identifiable. So, there are many questions around patient protection and privacy that have not yet been answered, and I think this is particularly problematic for the rare disease community.

Erika Berg (host):
Danton, do you want to talk about the example that Sherri mentioned?

Danton Char:
I think this is germane to all of us. The question of what patients want to know, and need to know, about AI being implemented into their care, and whether considering AI separately is a form of AI exceptionalism. There are many things we do in healthcare and decisions we make, such as the supplies we use or the drug vendors we purchase from, that patients are not informed about, even though these choices can have a significant impact on their care. With AI, we have talked about trying to develop guiding metrics for our own institution. We are using something now, but we are watching it closely and are willing to revise it if we view it differently. If we feel that the risk of harm is high from an AI tool being involved in someone's care and/or a patient could meaningfully opt out, then we feel that informing patients makes sense. If it is very much a background AI that does not have much of a footprint and that they could not meaningfully opt out of, such as the use of AI tools for prior authorization, claims, and submissions, we do not tend to talk to patients about it. That has been consistent and driven by our patient partner panel. We work with a group of patients who are learning alongside us about AI tools and are actively engaging with these discussions. They felt that this is a reasonable first step, because there is clearly data that shows the harms of opening a fire hose of information on a patient, particularly when they are under duress. Over-disclosure can lead to panic, anxiety about small decisions, and loss of focus on the critical decisions that their minds need to be focusing on. We are trying to strike a delicate balance. I do not pretend that we have gotten it perfectly yet, but this is our initial guidance.

Erika Berg (host):
Sherri, where are we in clinical settings with respect to establishing standards and guidelines for the ethical and transparent use of patient data? The use of AI in these clinical settings sounds a bit like a patchwork, with variations from institution to institution, but are there standards?

> ***"We have some newer or updated guidelines like TRIPOD+AI for prediction algorithms, but I would say that we are very far from the consistent use of these types of standards and guidelines. This needs to be something that the community agrees that we must be doing."***

Sherri Rose:
I will reiterate some of the things that Danton was highlighting. Yes, there are institutional processes to decide when and if they would deploy machine learning and AI tools. There is the institutional governance, and sometimes there are also state and federal policies. However, when we think about the general development of machine-learning and AI algorithms, there are some standards and guidelines for researchers, but they are not widely adopted or used, and that is a significant problem. Sometimes they are adopted and enforced when they are required by either medical journals or funders. But it is definitely a patchwork. We have some newer or updated guidelines like TRIPOD+AI for prediction algorithms, but I would say that we are very far from the consistent use of these types of standards and guidelines. This needs to be something that the community agrees that we must be doing. However, that usually involves an enforcement mechanism, otherwise it does not happen.

Erika Berg (host):
We have heard a lot about AI hallucinations, or errors by AI. How reliable is it? Have you in your research experienced these sorts of hallucinations and how should the medical community think about these issues with respect to patient safety?

Chunhua Weng:
I think that to address the hallucination problem we can develop knowledge graphs and knowledge bases and then combine them with large language models. The technique we have been using is retrieval-augmented generation. In this way we can always anchor the information to the facts and to the rules in the knowledge base. We also studied some algorithm topics such as the lost-in-the-middle phenomenon. That is another phenomenon when large language models are used to rank the retrieved documents. We can develop more algorithms to potentially detect these issues and then mitigate the risk. I remain optimistic about these issues. Technical issues are relatively easier to solve and whenever we discover one we can always explore and seek new solutions to address it by leveraging the knowledge out there, including the rich knowledge in the literature, and combining it with large language models. I think the more challenging issues are really the social, ethical, legal and policy ones that have been mentioned by Sherri and Danton. Those are the bigger challenges right now.

Erika Berg (host):
Looking forward, are there domains of medical care where you think AI could have particularly powerful benefits in the future? And do you see any red lines in terms of where it should not be used?

> ***"Looking five to ten years from now, I see the potential for breakthroughs, not just incremental improvements, through foundational models applied to multi-omic biological data that help us understand the language of life."***

Tim Guilliams:
I think that if we project five or ten years ahead, what may potentially become possible is the use of foundational models that have now been trained on natural language, so English, where we have seen incredible improvements in performance and understanding. I believe one of the biggest opportunities now is to create and train foundational models that can understand biological data and chemistry, in other words, models that are biologically aware and chemistry aware. They would not help us understand the natural language of English, but rather the language of life, which is your biology. Of course, it is too early to see any meaningful impacts from this approach because a few years ago it did not even exist. That is why I believe this could potentially lead to the biggest breakthroughs in human disease understanding, in a way that we cannot comprehend ourselves. The red lines would be following the regulations and ensuring patient safety to make sure that any promising breakthroughs are agreed upon with the FDA and included in the IND packages for clinical trials. This is the balance needed to maintain patient safety. Looking five to ten years from now, I see the potential for breakthroughs, not just incremental improvements, through foundational models applied to multi-omic biological data that help us understand the language of life.

Danton Char:
I agree with Tim. I think there is a lot of optimism in what AI tools can do. I do still have reservations about the economics of making them do what they do and about translating that vision into the amount of data and funding required, and specifically whether the healthcare system will be able to afford it. From a clinical standpoint, I think the things that have the most meaningful impact on clinician-patient interactions are those that move the many current distractions that get between you and your healthcare provider into the background. AI tools that could potentially shift billing and charging tasks away from the direct clinical encounter could be very beneficial in allowing clinicians to focus more on care. What we are seeing, at least anecdotally, is that when AI tries to replace the clinical interaction itself, there is resistance. Few patients would want an automated phone tree delivering their healthcare.

Chunhua Weng:
I think that keeping humans in the loop will still be very important. We often see people on the news worrying that AI will eventually replace or destroy humans. These are scary narratives, but from my perspective, I feel that AI in healthcare has great potential and can empower patients, particularly through education, as it provides them with more information. I am also optimistic that AI can handle some simple tasks so that the more complex cases can be saved for the experts. That could be another positive change that I am looking forward to.

Sherri Rose:
I will bring a few threads together. I think it is important to carefully consider the data type, as far as the domains that could be particularly powerful in other areas, for example imaging data. These tools are often just not as well suited to categorical or

integer data types. Oftentimes these are the types of data that are used in a lot of population health research. I also want to echo what Chunhua and Tim have said: seeing reproducible and generalizable results is going to take years. This is what we saw previously with earlier machine learning algorithms, and I agree with that. I want to reiterate that the default should not be an algorithm that will solve every problem. It is often the case that it will be a much more challenging solution, especially when we are thinking about issues of access to care and social drivers of health. Simply making a small change to an algorithm, or introducing an algorithm, is not going to solve that problem. When I think about red lines, they would be poorly considered algorithms, of which there are currently many, developed in a rush to publish or to bring products from the tech world to healthcare. That should definitely be a red line.

Book # 12.31

Have your say.

Webinars:

Podcasts:

Books:

Fondation Ipsen
70, rue Balard
75015 Paris
France
www.fondation-ipsen.org

Contact:
fondation@ipsen.com

Fondation Ipsen operates under the aegis of Fondation de France – www.fondation-ipsen.org
Editing: Natasha Barr – www.caretently.com
Graphic design: Céline Colombier-Maffre
Legal deposit: March 2026