Project Data Sphere®

2021 | BUILDING CONSENSUS ACCELERATING RESEARCH

THE FUTURE OF CLINICAL TRIAL DATA SHARING

FDA - PDS SYMPOSIUM X SUMMARY

"In an ideal world, each set of study results would be made available to inform future research, clinical or policy decisions."

– Dr. Robert Califf
PharmD, MSc, Associate Director of Pharmacoepidemiology at the FDA

Will not technology, is the biggest barrier to data sharing within the life sciences community. That was one consensus shared across experts speaking Sept. 23 at a symposium sponsored by the U.S. Food and Drug Administration and the non-profit Project Data Sphere®. With over 180 attendees representing Industry, Government, Non-Profits, Academic Institutions, and Healthcare Providers, it's clear that clinical trial data sharing is a "Hot" topic and it's one with a history of known challenges. Throughout this symposium, the challenges and barriers were largely set aside and 'Routes to Yes' and sustainable solutions were the focus. Data sharing is a thorny issue and as with every complex endeavor, it takes small, focused steps to build foundational and lasting change. The shared recommendation was that you need commitment from the top of your organization, start small and elevate patients as partners in clinical research and development because patients have both the will and a voice to gain agency over their data.

This was the tenth in a series of symposia co-sponsored by the FDA and Project Data Sphere since 2015 to address timely topics in cancer research and drug development. In her welcoming remarks, Donna Rivera, PharmD, MSc, Associate Director of Pharmacopidemiology at the FDA encouraged bold thinking and a focus

> **"If we don't have our own folks starting with a really high bar, we can't expect others to do so."**
>
> **– Dr. Najat Khan,**
> *PhD, Chief Data Science Officer at Janssen R&D*

on modern solutions to problems around data sharing. Bill Louv, PhD, President of Project Data Sphere® challenged everyone to identify specific actions to move data sharing forward. The discussions were broad ranging, with a major emphasis on how to advance sharing of individual patient data (IPD) from clinical trials in oncology. This is a complex ecosystem, requiring a tremendous amount collaboration, incentivization, and investment. "The enemy is the disease, and the competition is against the disease, and the goal is to put people in a better position to achieve better health." Dr. Harlan Krumholz, MD, SM, Harold H. Hines, Jr. Professor of Medicine at the Yale School of Medicine.

How the scientific community can evolve to enable broad data access.

There is a vast amount of underutilized clinical trial data. Why is mobilizing these resources so challenging? Technology is not the limiting factor for broad sharing of IPD according to keynote speaker, Dr. Robert Califf, MD, MACC, Head of Clinical Policy and Strategy at Verily Life Sciences and Google Health. Instead, it is the way we think about and put guardrails around data that is limiting. This point was echoed by Professor Steven Kern, PhD, Deputy Director, Quantitative Sciences at the Bill & Melinda Gates Foundation. Protection of patient privacy is a commonly cited reason not to share IPD. However, several speakers described the ways de-identification can mitigate this risk. Dr. Ned Sharpless, MD, Director of the NCI described how NCI has successfully deployed privacy-protecting technologies, such as synthetic data, hashing and matching, and gated cloud-only data enclaves, to permit responsible sharing of IPD including genomic data. The natural conservatism of industry will need to

> **"Data sharing requires a collaborative community focused on specific a priori objectives, with fit for purpose data and research questions that support transperancy, reproducibility and validity through methods, harmonization, and analysis."**
>
> **– Donna Rivera,**
> *PharmD, MSc, Associate Director of Pharmacoepidemiology at the FDA*

be overcome to advance data sharing. Califf and Frank Rockhold, PhD, Professor of Biostatistics and Bioinformatics at Duke Clinical Research Institute, Duke University Medical Center recalled the precedent of the launch and expansion of ClinicalTrials.gov, and concerns among some in industry that the requirement to register trials and provide summary data would be detrimental to their business. These concerns have not been realized.

The natural conservatism of industry will need to be overcome to advance data sharing. Califf and Frank Rockhold, PhD, Professor of Biostatistics and Bioinformatics at Duke Clinical Research Institute, Duke University Medical Center recalled the precedent of the launch and expansion of ClinicalTrials.gov, and concerns among some in industry that the requirement to register trials and provide summary data would be detrimental to their business. These concerns have not been realized. Although, as Dr. Harlan Krumholz, MD, SM, Harold H. Hines, Jr. Professor of Medicine at Yale School of Medicine

> **"PATIENTS HAVE INCREDIBLE CAPACITY TO HELP RESEARCH EFFORTS AND ARE CLAMORING TO DRIVE IMPACT, PATIENT INVOLVEMENT IS RIPE FOR DISRUPTION AND EACH OF US CAN PLAY A ROLE BY THINKING CREATIVELY TO MAXIMIZE THEIR CONTRIBUTIONS. THIS CAN BE THROUGH THEIR VOICE DRIVING POLICY, IT CAN BE THROUGH TANGIBLE MEASURES SUCH AS ACQUIRING AND SHARING THEIR SAMPLES AND DATA, AND IT CAN BE THROUGH PARTICIPATING IN EFFORTS LIKE OURS (COUNT ME IN) TO DRIVE THE ABILITY TO GENERATE MASSIVE AMOUNTS OF DATA TO BE SHARED."**

– Corrie Painter,
Corrie Painter, PhD,
Deputy Director, Count Me In,
Broad Institute

pointed out, compliance with ClinicalTrials.gov reporting is not optimal so the precedent set is sub-ideal.

The panelists and speakers didn't spend this symposium rehashing these known hurdles and perceived barriers to sharing. Throughout the session they emphasized ways to overcome these challenges. Najat Khan, PhD, Chief Data Science Officer at Janssen R&D, suggested a few key steps including how to think about sharing patient level data. Johnson & Johnson was instrumental and a proactive partner for Yale's YODA (Yale University Open Data Access) initiative and they are committed to the cause of advancing science by demonstrating the value in making these data available through responsibly answering questions that add value to patients. Khan shared examples of the systems and technology that must be in place -- consolidated platforms, audit trails for all work, a protocol and SOP to support this.

There is an opportunity to shift the cultural norm in the clinical trial ecosystem and give trial participants more agency over their data. Several speakers described the potential for patient-driven data sharing. This is an important mechanism, but patients need an infrastructure to support this. There are instances such as AllStripes where this is actively happening for rare tumors, as Ms. Khan pointed out. Krumholz emphasized how important it is to bring participants to the table, remove the use of the term 'subjects'and replace it with 'partners'. Dr. David Fajgenbaum, MD, MBA, MSc, Assistant Professor of Medicine at the University of Pennsylvania shared examples where his patient advocacy group, Castleman Disease Collaborative Network, were directly capturing information from
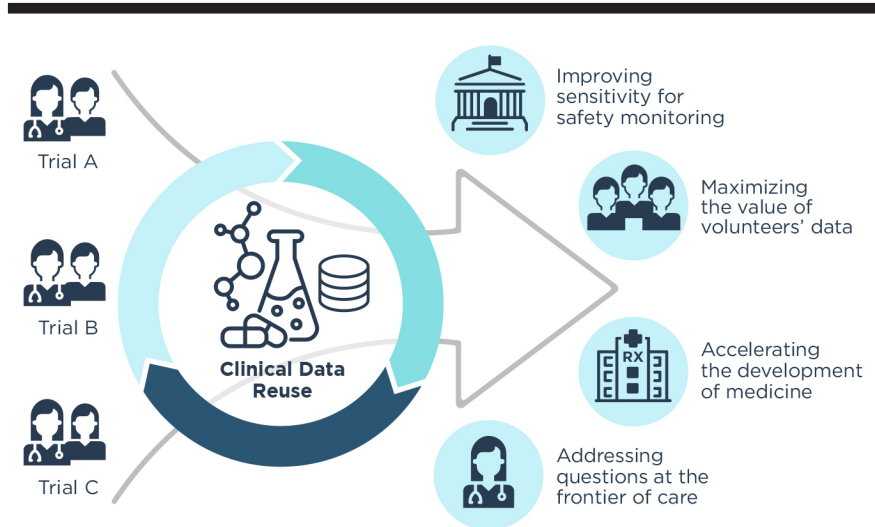
patients and working around some of what's historically been bureaucratic delays or red tape. Does access to real-world IPD need to be democratized? Dr. Atul Butte, MD, PhD, Director of the Bakar Computational Health Sciences Institute at the University of California at San Francisco, asked panelists to consider the impact of entry of third-party electronic health record data aggregators into the biomedical ecosystem. Krumholz commented that this is of concern because it creates friction in the ecosystem, turning real-world IPD into a commercial product and limiting data transparency. Khan pointed out that in the reality of the current ecosystem the commercial data aggregators serve an important role because they are a source for large, structured data sets that are not otherwise available. Kern concurred, "I would argue that they help in the sense that they are creating new ways that we can combine the data.It reinforces this idea that it's not a technology problem we have here. We're really struggling with a problem around governance and will." Ultimately, he foresees an evolution to a market based on which entity provides the best insights and knowledge rather than limiting access and creating barriers to making data available.

Regulatory guidance around data sharing needs to be clear and consistent. Of concern, ambiguity may arise around European Union regulations. While the General Data Protection Regulation provides strong privacy protection for individuals' information, including health information, with significant financial penalties for infringement, the European Medicines Agency is working toward achieving greater sharing of IPD from clinical trials used to support drug approvals.

"Routes to Yes": Balancing incentives to drive data sharing

Throughout the session there was no argument around the incentives to share and belief that sharing advances scientific knowledge and delivers on the partnership with trial participants When data are shared and findings are reproduced, or replicated on aggregated data, confidence in the original findings increases and impact is enhanced. Aggregated data may support more robust findings, particularly in settings where the patient population is relatively small such as in a rare disease or a biomarker -identified subgroup. Secondary research may support entirely new findings about disease. Journals and funders typically require data sharing as conditions of publication and funding, respectively, although the importance of this incentive is unclear given well-documented poor compliance. Rockhold related how Duke University has adopted a forward thinking approach by formally assigning academic credit for data sharing. When Duke faculty share data and can document that other researchers use it, it can be factored into promotion decisions.

Panel #1 discussed concerns surrounding new analyses performed on shared data. There should be full transparency, where access to IPD

Trial A

Trial B

Trial C

Clinical Data Reuse

Improving sensitivity for safety monitoring

Maximizing the value of volunteers' data

Accelerating the development of medicine

Addressing questions at the frontier of care

could be contingent on sharing methods and code. Krumholz reminded the audience that rogue analyses already exist and have since the beginning of time, so any concern about introducing it by way of sharing more IPD should be tempered. For example, there are cases of primary researchers improperly analyzing IPD and secondary researchers improperly using summary trial data. Krumholz shared a perspective of embracing different methods and approaches, "making data available offers the ability for others to go in and check claims, creates trust". Butte reassured the audience that there is a way to get through all these threats and rogue analysis.

Additional measures are available

to drive more data sharing. Kern shared that his organization in principle favors incentives over mandates because the quality of the data shared will be better. The Gates Foundation has shifted their support model largely to collaborate on the front end of trial planning and prospectively ask for data-sharing plans and partners with grant recipients in developing a data infrastructure. Kern also said some systems allow anyone to replicate or rerun analysis which allows for checks and balances, and this fosters trust and complete transparency. Khan and Yili Pritchett, PhD, Vice President and Head of Biometrics at G1 Therapeutics, described how successful data sharing by industry can be accomplished by taking an incremental approach. Khan shared how starting with a small data-sharing pilot can demonstrate to internal stakeholders both the value and ability to mitigate risk. Pritchett described sharing clinical trial data sets in a limited fashion, not including experimental arms and exploratory endpoints though not all panelists echoed this as the path to full transparency. It was evident, that across industry and academic institutions, large and small, through implementing processes

"IRBs and advocates and patients who care about cancer progress have to realize there's a fundamental tension between privacy and research, Privacy is good and research is good and somewhere as a society we have to draw a line about what sort of research we allow in the absence of consent, we have to be able to do that."

– Dr. Ned Sharpless,
MD, Director of the NCI

> **"We need to get rid of the word subjects, start thinking about partners and these people who are participating."**
>
> **– Dr. Harlan Krumholz,**
> *MD, SM, Harold H. Hines, Jr. Professor of Medicine at Yale School of Medicine*

and close collaboration we can achieve greater sharing and adopt this way of thinking.

Patient Power, harnessing their voice to advance treatment

Patient centricity was a theme throughout the symposium but brought into focus during the fireside chat and closing session with representation across three advocacy groups (Count Me In, Castleman Disease Collaborative Network, and The LAM Foundation). Through advocacy and direct engagement, these collaborative networks have proven successful in navigating this clinical research ecosystem to advance science. While it still takes much longer than desired, Sharpless clearly challenged the audience to think creatively, and find ways to actively work directly with patients to drive

faster results and establish practices that allow for the maximum capacity to share. Patients have the largest stake in ensuring that their data has the greatest impact and there's little stopping the will of the patient. Advocacy centralizes the patient's voice and allows for active participation in the research and development process including driving policy change. Data sharing honors patients by maximizing the value of their participation in trials. During panel #2, Pritchett offered a clear observation that sharing IPD is a physical means to demonstrate an organization's commitment to patient-first values.

Fajgenbaum shared that patient advocates are well positioned to find paths to overcome these barriers, serving as unbiased, relentless drivers motivated by progress against disease. Patients can serve as inter-

mediaries between other stakeholders, e.g., industry and academics, and can advocate for policy changes with regulatory and funding agencies.

Patients can take an active role in data sharing by collecting samples, sharing health records, and answering surveys under research protocols that employ remote consent. This engagement overcomes the challenges of obtaining Institutional Review Board approval at individual institutions and can be particularly effective for accelerating data collection for rare diseases. As patients and leaders of advocacy organizations, Fajgenbaum who lives with Castleman disease, and Corrie Painter, PhD, Deputy Director, Count Me In, Broad Institute, an Angiosarcoma survivor, shared their own creative solutions in accessing or capturing patient data to advance research in their rare disease space. Both are deeply committed to making a positive impact for the future benefit of patients affected by these diseases through advocacy and establishing repositories of registry data.

Sue Sherman, MHA, Executive

> **"Patient advocates have the ability to cross between lanes and serve as an intermediate between a pharmaceutical company that may have some data and a researcher"**
>
> **– Dr. David Fajgenbaum,**
> *MD, MBA, MSc,*
> *Assistant Professor of Medicine at the University of Pennsylvania*

PROJECT DATA SPHERE PLATFORM BY THE NUMBERS

>200 Patient Level Datasets	**>240K** Patient Lives	**>18** Data Providers	**>115** Peer-reviewed Publications

Director and Chief Executive Officer of the LAM (lymphangioleiomyomatosis) Foundation, shared how their partnership with the LAM patient community, NIH, Clinical Researchers, and the FDA has made it possible achieve remarkable progress in treating this rare disease, including the first FDA-approved treatment for LAM, published treatment guidelines, development of a blood-based diagnostic marker reducing the need for biopsy, enrollment of 3,300 patients in a US patient database, and $29 million raised for research, education, and patient support.

Technology advances can facilitate patient-driven data and insights sharing. Sherman described a recent LAM trial in which patients overseas provided data entirely remotely, including measurements obtained by home spirometry. Andrea Slattery, a LAM patient and Director of Research at Symmetry Peak Management, advocated for harnessing technology to empower patients in data sharing, for example, streamlined apps for patient data entry and personal health dashboards with integrated health records. Slattery and Sherman challenged everyone participating in this symposium to commit to making

progress, patients are ready to continue their part and are looking to all parties involved to collaborate and drive change.

Plan for data sharing up front and practice good data stewardship throughout

What are some of the best practices for successful sharing of clinical trial IPD? David Chambers, DPhil, Deputy Director for Implementation Science at the NCI, moderated the second panel on sustainability in data sharing and suggested that the guiding principle is to "plan for sustainability so that the investment pays off overtime." To optimize efficiency and impact, data sharing should be an integral part of the clinical research plan. How data will ultimately be shared should be a key consideration in the framework established for data collection and

analysis. Future research questions are yet to be defined so implementing policies and procedures which support the broadest applied use are important things to consider. Data standards and broad patient consents are key elements. Data sharing involves substantial time, money, and expertise and being thoughtful about the approach (defining what key data elements and access models) which has the highest likelihood of supporting the research question is a key to success, bringing patients in as part of these conversations where possible. A sustained, significant commitment to data sharing at the highest level of an organization is important.

"It can't be a trend in support because it's a long process. It takes policy changes, it takes compliance, and not everyone is going to be happy about those decisions along the way. I can attest there's a lot of push back from time to time," according to Jaime Guidry Auvil, PhD, Director of the Office of Data Sharing at the NCI.

As Krumholz reminded us, science and medicine are constantly evolving and making data available in the public space allows for secondary and tertiary analysis which supports further validation of methods and data quality. Technology solutions and data sharing platforms should have good governance models to mitigate risks tied to patient privacy and appropriate access. Researchers should willingly commit to sharing their findings and methods, including

> **"We can't let what's difficult prevent us from finding a solution."**
>
> **– Andrea Slattery,**
> *LAM patient, Advocate, and Director of Research at Symmetry Peak Management*

"GET COMMIT-MENT FROM THE TOP AND TREAT YOUR DATA COMMENSURATE WITH THE VALUE IT COST YOU TO OBTAIN IT"

– Frank Rockhold,
PhD, Professor of Biostatistics and Bioinformatics at Duke Clinical Research Institute

the code and tests executed against the data.   Every person speaking throughout the panel spoke from the perspective of being a good data steward and leader driving change in this space.

How will you make an impact?

It's widely known there are limiting factors to data sharing, some of these were mentioned during the symposium and questions that surfaced from the audience. There are real barriers tied to funding, concerns around privacy, and data quality. The path forward isn't magnifying the need for greater technology, while there have been interesting advancements in applying artificial intelligence to support confidence testing of patient privacy and other aspects that support efficient enablement of sharing and access, they aren't the solutions that will make us think differently about sharing patient level data. These active concerns will persist about

the data being shared, who owns the data and how these data will be used. It was echoed throughout the day that the focus needs to be on good data stewardship, including traceability that includes who owns the original data and what if any manipulations have taken place. Have mechanisms for checks and balances. There will always be unknowns, particularly with older "legacy" data sets, the collection standards vary and, in some cases, extensive manual curation may be needed to make data usable. Ultimately let's let researchers do what they do best and make new discoveries with the data. The first step is commitment, then establishing policies and procedures to enable sharing or access. Focus on those data that seem to fall within the bounds of having big impact and low risk, don't let perfect be the enemy of progress and start to look across the data you are holding even from failed or discontinued trials, where can you start making

advancements here. When you have questions or concerns think about reaching out to one or any of these leading experts and leverage patient advocacy groups to keep progress moving forward.

The FDA and Project Data Sphere were honored to have such an esteemed group of presenters in this session and are committed to continued partnership and collaborations to actively participate within the scientific community to advance sharing and ultimately cancer research. As Martin Murphy, DMedSc, PhD, Director Emeritus of the CEO Roundtable on Cancer commented in the symposium chat, "At the very least these data need to be available…it honors the patients who volunteered to become subjects in these trials." There is a need for greater awareness of this missed opportunity to advance research and failure to fulfil obligations to patients.

"I am grateful to all of the presenters and partici-pants who made Symposium X an interesting and important forum on Data Sharing. Simply put, we have an obligation to optimize the value of data collected from patients to benefit future patients. Several successes were described during the symposium and they give us confidence and inspiration that we can do more. And we must do more. The processes of data sharing do not operate at scale and are too slow. Data sharing is not keeping up with the speed of innovation in drug development. Through our collaborations, PDS is committed to catalyze improvements in policies, procedures, standards, and infrastructure to optimize data sharing.

– Bill Louv,
PhD, President of Project Data Sphere®

Thank you to all the speakers, panelists, and guests.

Bill Louv,
Project Data Sphere

Donna Rivera,
FDA

Robert Califf,
Verily Life Sciences, Google Health

Atul Butte,
UCSF

Najat Khan,
Janssen, Johnson & Johnson

Harlan Krumholz,
Yale School of Medicine

Steven Kern,
Bill and Melinda Gates Foundation

David Chambers,
NCI

Frank Rockhold,
Duke University Medical Center

Jaime Guidry Auvil,
NCI

Yili Pritchett,
G1 Therapeutics

Corrie Painter,
Count Me In, Broad Institute

Ned Sharpless,
NCI

David Fajgenbaum,
University of Pennsylvania

Susan Sherman,
The LAM Foundation

Andrea Slattery,
Symmetry Peak Management

A special thanks to our data sharing partners.

OUR MISSION

Improve outcomes for cancer patients by openly sharing data, convening world class experts, and collaborating across industry and regulators to catalyze new scientific insights that accelerate delivery of effective treatments to patients.

"CEO Roundtable on Cancer sees a future of broad sharing and access of deidentified clinical trial data as one significant pathway in making progress toward bridging the gap in access to care and the elimination of cancer. The PDS model of making this data available in an easy to access way amplifies how the ecosystem can come together to rapidly advance patient outcomes."

– MaryLisabeth Rich,
President CEO Round Table on Cancer