

EXPERIENCE REPORT

The Global academic research organization network: Data sharing to cure diseases and enable learning health systems

Masanori Fukushima¹ | Christopher Austin² | Norihiro Sato³ | Tatsuya Maruyama⁴ | Eileen Navarro⁵ | Mitra Rocca⁵ | Jacques Demotes⁶ | Melissa Haendel⁷ | Samuel L. Volchenbom⁸ | Matthew Cowperthwaite⁹ | Jonathan C. Silverstein¹⁰ | Chris Webb¹¹ | Ida Sim¹² | Marianne Chase¹³ | John Speakman¹⁴ | Erika Augustine¹⁵ | Daniel E. Ford¹⁶ | Rebecca Kush^{17,18}

¹Translational Research Center for Medical Innovation, Foundation for Biomedical Research and Innovation, Kobe, Japan

²National Center for Advancing Translational Sciences, NIH, Bethesda, MD, USA

³Hokkaido University Hospital, Hokkaido, Japan

⁴Japan Agency for Medical Research and Development, Keio University Hospital, Tokyo, Japan

⁵U.S. Food and Drug Administration, Silver Spring, MD, USA

⁶European Clinical Research Infrastructure Network, Paris, France

⁷Center for Data to Health, Oregon Health & Sciences University, Portland, OR, USA

⁸Center for Research Informatics, University of Chicago, Chicago, IL

⁹St. David's HealthCare Office of Research, Austin, TX, USA

¹⁰University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

¹¹University of Texas at Austin Dell Medical School, Austin, TX, USA

¹²University of California San Francisco, San Francisco, CA, USA

¹³Neurological Clinical Research Institute, Massachusetts General Hospital, Boston, MA, USA

¹⁴NYU Langone Health, New York, NY, USA

¹⁵University of Rochester Medical Center, Rochester, NY, USA

¹⁶Johns Hopkins School of Medicine, Baltimore, MD, USA

¹⁷Translational Research Center for Medical Innovation, Kobe, Japan

Abstract

Introduction: Global data sharing is essential. This is the premise of the Academic Research Organization (ARO) Council, which was initiated in Japan in 2013 and has since been expanding throughout Asia and into Europe and the United States. The volume of data is growing exponentially, providing not only challenges but also the clear opportunity to understand and treat diseases in ways not previously considered. Harnessing the knowledge within the data in a successful way can provide researchers and clinicians with new ideas for therapies while avoiding repeats of failed experiments. This knowledge transfer from research into clinical care is at the heart of a learning health system.

Methods: The ARO Council wishes to form a worldwide complementary system for the benefit of all patients and investigators, catalyzing more efficient and innovative medical research processes. Thus, they have organized Global ARO Network Workshops to bring interested parties together, focusing on the aspects necessary to make such a global effort successful. One such workshop was held in Austin, Texas, in November 2017. Representatives from Japan, Taiwan, Singapore, Europe, and the United States reported on their efforts to encourage data sharing and to use research to inform care through learning health systems.

Results: This experience report summarizes presentations and discussions at the Global ARO Network Workshop held in November 2017 in Austin, TX, with representatives from Japan, Korea, Singapore, Taiwan, Europe, and the United States. Themes and recommendations to progress their efforts are explored. Standardization and harmonization are at the heart of these discussions to enable data sharing. In addition, the transformation of clinical research processes through disruptive innovation, while ensuring integrity and ethics, will be key to achieving the ARO Council goal to overcome diseases such that people not only live longer but also are healthier and happier as they age.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2018 The Authors. Learning Health Systems published by Wiley Periodicals, Inc. on behalf of the University of Michigan

¹⁸Catalysis and Elligo Health Research, Austin, TX, USA

Correspondence

R. Kush, Catalysis Research for Health, Executive; Clinical Data Interchange Standards Consortium, USA.
Email: rkush@catalysisresearch.com

Conclusions: The achievement of global learning health systems will require further exploration, consensus-building, funding aligned with incentives for data sharing, standardization, harmonization, and actions that support global interests for the benefit of patients.

KEYWORDS

data sharing, global, harmonization, registry, standards

1 | BACKGROUND

Global data sharing is essential. This is the premise of the Academic Research Organization (ARO) Council, which was initiated in Japan in 2013. The volume of data is growing exponentially, providing not only challenges but also the clear opportunity to understand and treat diseases in ways not previously considered. Harnessing the knowledge within the data in a successful way can provide researchers and clinicians with new ideas for therapies while avoiding repeats of failed experiments. In addition to data sharing, the transformation of clinical research processes through disruptive innovation, while ensuring integrity and ethics, will be key to achieving the ARO Council goal to overcome intractable diseases such that people not only live longer but also are healthier and happier as they age.

Over the past 5 years, the ARO Council has been expanding from Japan into Asia (adding Singapore, Taiwan, and Korea) and now includes leaders from academic institutions in Europe and the United States. The establishment of Japan's Agency for Medical Research and Development (AMED) in 2015 provided a new infrastructure for innovation and encouragement for the ARO Council. Annual Global ARO Workshops have taken place with themes focusing on standardization and harmonization of clinical research, a new paradigm of medical science based upon data sharing, and real-world data (RWD) and disruptive innovation. The goal is to form a worldwide complementary system for the benefit of all patients and investigators, catalyzing more efficient and innovative medical research processes. Standardization and harmonization are essential to achieve this goal with global data sharing at its core.

Key messages and academic research initiatives reported at the Global ARO Council Workshop (held in Austin, TX, in November 2017) are summarized herein. From participants in the November 2017 Global ARO Network Workshop, projects related to a global learning health system (LHS) in Asia (Japan, Taiwan, Singapore, and Korea), Europe, and the United States are described from the perspectives of government, regulators, and AROs. Themes are explored in further depth to make recommendations for potential pathways towards global LHSs. Prof Norihiro Sato opened by emphasizing that we can accelerate the cycle of an LHS through harmonization, standardization, and global data sharing. The closing discussion moderated by Professor Masanori Fukushima and Dr Chris Austin further develops the laudable goals, approach, and future vision of the Global ARO Council and provides emerging themes from the workshop participants.

2 | JAPAN AGENCY FOR MEDICAL RESEARCH AND DEVELOPMENT

Dr Maruyama presented on what has been referred to as Japan's National Institutes of Health ("NIH"). The Japan AMED resulted from a Japan Revitalization Strategy launched in 2013 by Prime Minister Shinzo Abe to accelerate the development of practical applications of innovative medical technologies. The Headquarters for Healthcare Policy (HHP) was organized to promote integrated medical research and development from basic science to practical application and to enhance economic growth and the quality of medical care by promoting the creation and overseas expansion of new activities within industry that relate to health care and medical care. Related efforts focused on enhancing safety measures for drugs and medical devices and establishing development and regulation of regenerative medical products. These broad steps were based upon two key observations and strategic goals within Japan: (1) the decreasing rate of births along with an increasing overall life expectancy emphasized the need for healthy longevity and (2) a disconnect between basic medical research and clinical research needed to be addressed.

AMED was funded by reallocating funds from multiple existing agencies. AMED now promotes translational and clinical research through a number of collaborative initiatives. The Health and Medical Strategy calls for Japan to conduct innovative medical research and development from preclinical/basic science conducted by academia through regulatory approvals of new therapies and to demonstrate leadership in collaborative international research. They are keenly interested in playing a leading role in building such systems, supporting core centers and hospitals to collaborate in international clinical research. They work with regulators, industry, and academia to fund and promote global research to improve health care.

3 | US NIH NATIONAL CENTER FOR ADVANCING TRANSLATIONAL SCIENCES—CATALYZING TRANSLATIONAL INNOVATION

Dr Christopher Austin stated, in response to an opening comment about music being "the orderly placement of silence between the notes," that we currently have noise and dissonance rather than silence, in the United States. "We pay people to be dissonant and inharmonious" through incentives that are inadvertently in the wrong places. He commented that Japan has a different approach in that

they tie data sharing and budgets together; data sharing is required for funding to be awarded.

The NIH National Center for Advancing Translational Sciences (NCATS) has a mission to enhance the development, testing, and implementation of diagnostics and therapeutics across a wide range of human diseases and conditions. Collaboration with other government agencies, academia industry, and nonprofit organizations is essential for NCATS. One of the NCATS initiatives is the Clinical and Translational Science Awards (CTSA), which Dr Austin described. The NCATS CTSA Program Hubs innovate locally and collaborate regionally and nationally in projects such as informatics, training, streamlining processes, community engagement, and understudied populations. Data sharing is essential to CTSA success. Cross-CTSA initiatives, such as the Trial Innovation Network (TIN), solve systematic problems that limit the efficiency and effectiveness of clinical translational science.

The TIN leverages the expertise and experience across CTSA centers to develop resources and capabilities that can benefit all centers, such as an institutional review board (IRB) reliance platform, electronic health record (EHR)-based patient recruitment methods, standard master agreements, and contracts. These are intended to reduce time spent on these activities such that more science can be performed. Focus has been on encouraging innovative scientific hypotheses and models, novel study designs, patient-reported outcomes (PRO) endpoints, and a “learning clinical research system.”

Specific research networks mentioned as examples of data sharing were the Rare Diseases Clinical Research Network (RDCRN), the Undiagnosed Disease Network (UDN), and the Patient-Centered Outcomes Research Network (PCORNet). These involve common data center and patient advocacy groups. He also referenced the All of Us program, which is in the process of recruiting one million patients, has an extensible data platform, and is an example of personalized medicine. Dr Austin stated that it seems there is agreement that standards are important, but not on which standards to adopt.

A recent cross-CTSA NCATS grant award that deserves special mention is for the creation of a Center for Data to Health (CD2H), which is led by Dr Melissa Haendel. Dr Haendel reported specifically on her work in this area in a presentation that is summarized later in this article.

Dr Rob Califf (Food and Drug Administration [FDA] Commissioner in 2016 and well-respected academic researcher from Duke) has been quoted as saying “Clinical trials in this country take too long, cost too much, and too often don't give us the answers we need to take care of our patients. Other than that, the system works great.” Even the time from funding approval to study start can take 2 years, and a study can take over 10 years. The inability to recruit patients actually affects a large percentage of trials, while the length of time to recruit can make the trial eventually irrelevant because the underlying basic science may move more quickly. This results in unnecessary loss of patient lives, along with investigator careers and information that could improve health, thus delaying the advancement of medicine and LHSs.

4 | US FOOD AND DRUG ADMINISTRATION

Dr Eileen Navarro spoke about the value of harmonization and standards for FDA. Global clinical research studies are important,

and harmonization can reduce timelines and costs for such trials.¹ Global harmonization efforts such as those of the International Council on Harmonization (ICH)² have involved global regulators and the regulated industry for many years. There are now harmonized global guidelines that encompass many aspects of research, including protocol development, statistical analysis, and an electronic common technical document (eCTD) for regulatory submissions. One example of an ICH product that has been harmonized by the United States, Europe, and Japan is “Guidance for Industry: Providing Regulatory Submissions in Electronic Format—Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications.”³ The use of the MedDRA⁴ terminology coding system for adverse events in clinical research has also been agreed by the United States, Europe, and Japan through the ICH. See Table 1 for other harmonized guidelines developed by ICH that are referenced globally for regulated clinical research programs for development of new therapies.

Dr Navarro used the analogy of shipping containers to illustrate the value of standards for regulators. The shipping container, a simple steel box, has a standard size that can be loaded onto semis, ships, and planes. This has revolutionized shipping, making it “the biggest enabler of globalisation.”⁵ The shipping container (a simple steel box 8 ft. wide, 8 ft. 6 in. high, and 40 ft. long) can be loaded onto a semitruck, a train,

TABLE 1 Harmonized regulatory guidelines by the International Council on Harmonization (ICH)

Harmonized regulatory guidelines	
SAFETY	
Carcinogenicity	Biotechnology products
Genotoxicity	Pharmacology
Toxicokinetics, pharmacokinetics	Immunotoxicology
Reproductive toxicology	Nonclinical evaluation of cancer drugs
Toxicity testing	Photosafety
EFFICACY	
Clinical study reports	Statistical principles for clinical trials
Dose-response	Evaluation by therapeutic category
Ethnic factors, special populations	Risk benefit
Clinical trials	Pharmacogenomics
QUALITY	
Stability	GMP
Analytical validation	Pharmaceutical development
Impurities	Quality risk management
Pharmacopeias	Pharmaceutical quality systems
Biotechnology product specifications	Manufacture of drug substances
MULTIDISCIPLINARY	
MedDRA terminology	Common technical document (CTD) and electronic CTD
Electronic standards	Drug dictionaries: Standards and data elements
Nonclinical safety reports	Gene therapy genotoxic impurities

<https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.

a ship, and/or a plane and moved readily around the globe using these different forms of transportation. This is a lesson for data sharing.

A sharing environment can streamline global research. Examples provided are the common protocol template (CPT). Developed by an FDA/NIH team^{6,7} concurrently with the TransCelerate CPT⁸ initiative, these two templates have now been harmonized such that they apply for industry/regulated trials as well as for academic/NIH trials. Protocol standardization is a key opportunity to being able to plan for required standards downstream. The protocol elements comprising the study design (ie, epochs, arms, and elements) are all required for regulatory submissions so that the reviewers can have the necessary context to comprehend the associated research results. This standard template has been key in ensuring that key endpoints are aligned with protocol objectives and provides an opportunity to critically review essential data requirements, the statistical analysis plan, and integrity of the study design. IRBs/ethics committees welcome protocols in standard formats to facilitate their reviews and to ensure that the necessary protocol components are addressed. This template is now being “technology-enabled” to facilitate reuse of components for trial registration (ct.gov, WHO ICTRP, and EudraCT), in the analysis plan and in the clinical study report or publication and results sharing. “Content/metadata sharing” complements data sharing and facilitates interpretation of the data.

Dr Navarro listed five core reasons on why standards matter: Standards (1) bridge the evidence spectrum, (2) facilitate transparent research, (3) facilitate efficient reviews, (4) facilitate data reuse, and (5) matter to patients. A number of examples were discussed, including the importance of being able to answer questions such as whether a drug has varied effects on individuals in differently populations (ie, women, men, children, or various races).⁹ She emphasized that nonproprietary data standards encourage innovation. FDA now maintains a Data Standards Catalog¹⁰ indicating standards they expect to be employed for regulatory submissions to the US FDA for new product approvals. Japan’s analogous regulatory authority (PMDA) also provides similar resources that are largely aligned with FDA on the selection of and requirements around standards. Europe’s EMA does not yet require that the data be included in eSubmissions; however, they are aligned on requirements around trial registration and related protocol metadata.

Mitra Rocca of FDA spoke on the topic of RWD. RWD includes “data derived from electronic health records, claims and billing data, data from product and disease registries, patient-generated data, including in home-use settings, and data gathered from other sources that can inform on health status such as mobile devices.” Real-world evidence (RWE) is “the clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD.”¹¹ The FDA wishes to maximize the opportunities to incorporate data/evidence from settings that more closely reflect clinical practice into the regulatory decision-making process. Additional goals are to increase population diversity and to improve efficiencies such as identification of patients/populations and reduction of data reentry or transcription, which is time-consuming and negatively impacts data quality. There are expectations in the 21st Century Cures Act in the United States for FDA to establish a program to evaluate the use of RWD to generate evidence. Global initiatives for organizations such

as AMED, Europe’s Innovative Medicines Initiative (IMI), and the ARO Council also reflect their interest in such objectives.

One project that is designed to explore the use of RWE derived from delivery of health care in routine clinical settings is a Common Data Model (CDM) Harmonization project funded by the PCOR Trust Fund and led by Ms Rocca of FDA. This project will harmonize varying CDMs (ie, Observational Medical Outcomes Partnership [OMOP]/ODHSI, PCORNet, Sentinel, and Informatics for Integrating Biology with the Bedside [i2b2]) with the Biomedical Research Integrated Domain Group (BRIDG) model. The BRIDG model will serve as the intermediary model; it is a global *Clinical Data Interchange Standards Consortium* (CDISC), Health Level Seven (HL7), and International Standards Organization (ISO) standard with the broad scope of protocol-driven research, including genomic data. The goal is to build a data infrastructure for conducting PCOR using RWD/observational data derived from the delivery of health care in routine clinical settings. This will be achieved by developing and applying an automated extract, transform, and load (ETL) process to harmonize several CDMs with each other, leveraging open standards and controlled terminologies to advance PCOR (Figure below). The CDM Harmonization project Use Case will be a pharmacovigilance assessment in the area of oncology. Testing this Use Case will be a number of different data partners, including AROs that have spent time and resources changing data into different formats to accommodate the various networks that have developed varying CDMs rather than using an existing standard. The long-term goals for this project are to

1. leverage the developed infrastructure to address additional use cases;
2. share the developed methods, tools, and resources with interested stakeholders;
3. enhance regulatory decisions by providing FDA reviewers with access to a larger network of RWD data; and
4. enable a set of standard data representations that can be used to define relevant data models for an LHS.

5 | EUROPEAN CLINICAL RESEARCH INFRASTRUCTURE NETWORK

Dr Jacques Demotes, Director of the European Clinical Research Infrastructure Network (ECRIN), presented on the ECRIN Data Center Certification process, the Clinical Research Initiative for Global Health (CRIGH), and the Consensus Document on Principles of Sharing Individual Patient Data from Clinical Trials.¹² ECRIN is a not-for-profit intergovernmental organization in Europe that supports the conduct of multinational clinical trials. ECRIN has spearheaded numerous projects within Europe designed to build infrastructure for clinical research. They have developed a set of criteria, including 129 requirements, to certify data centers and have conducted such certifications, including in Asia and Africa as well as Europe. CRIGH includes projects in the areas of (1) infrastructure and funding, (2) global core

competencies, (3) research ethics, (4) patient involvement, (5) comparative effectiveness research, and (6) regulatory awareness.

One of the recent efforts of ECRIN involved a multiyear project to build global consensus around a set of 10 data sharing principles with associated specific recommendations. The resulting consensus document was published in December 2017 and covers this topic in more depth than prior efforts have dared to address. ECRIN is a core supporter and facilitator of the ARO Council's activities. Additional ECRIN activities are detailed on their website.¹³

6 | GLOBAL AROS

The remainder of the Workshop was dedicated to learning more about activities within AROs and research institutes, based in Asia and across the United States, that could align with the ARO Council's goals and promote global expansion of LHSs. Three areas emphasized were (1) supporting academic and pharma-sponsored research (infrastructure, modernization, and the research enterprise); (2) data sharing (methods, tools/platforms, registries, and data commons); and (3) standards and tools to bridge the chasm between basic research and clinical research.

From Asia, presentations by Dr Chung Y. Hsu from China Medical University Hospital (Taiwan), Dr Ueng-Cheng Yang from National Yang-Ming University (Taiwan), Shu Ling Lee from Singapore Clinical Research Institute (SCRI), Dr In-Jin Jang from Seoul National University Hospital (Korea), and Dr Norihiro Sato a vice president of ARO Council from Hokkaido University (Japan) were delivered.

Dr Chung Y. Hsu described the Taiwan Stroke Registry (TSR), a stroke research network, the Taiwan Clinical Trial Consortium for Stroke, and the Stroke Biosignature Project. TSR has network of 60 hospitals in Taiwan, and the consortium is contributing to improve

the effect of preventive measures and the development of safer and more effective therapeutic strategies. TSR also takes the advantages of sophisticated big data analytics based on the big database from National Health Insurance covering more than 99% of the population in Taiwan.

Dr Ueng-Cheng Yang also presented their efforts that have been made to harmonize and standardize their operation to facilitate multicentered trials by using Clinical Informatics and Management System (CIMS) that was developed at NIH and improved in Taiwan. He also spoke about their preparation of obtaining accreditation of Data Center by ECRIN to improve data quality.

Shu Ling Lee presented their challenges to adopt and meet CDISC standard in SCRI, which is the only and national ARO in Singapore. SCRI and TRI are providing similar clinical research services in their respective countries.

Dr In-Jin Jang presented approaches of Seoul National University Hospital such as adopting CDISC standards, preparing ECRIN Data Center Certification to be obtained in 2019, and facilitating investigator-initiated trials to enhance the clinical research capability as a leading hospital in clinical research in Korea.

Dr Norihiro Sato presented the strategy for global clinical trials by Japan ARO Council as a vice president. History of national TR projects for medical innovation and achievement by utilizing ARO Network in Japan was also described. The Asian ARO Network was constructed as the first phase of the formation of multinational network towards the common goal as Figure 1 shows.

The Asia ARO Network has been formed among Korea, Taiwan, Singapore, and Japan and is cooperatively seeking to realize this common goal. CDISC standards implementation in Asian ARO Network was agreed among Asian institutes. Likewise, the ARO Network is targeting to be ECRIN certified data center to promote standardization and organize multinational clinical trials.

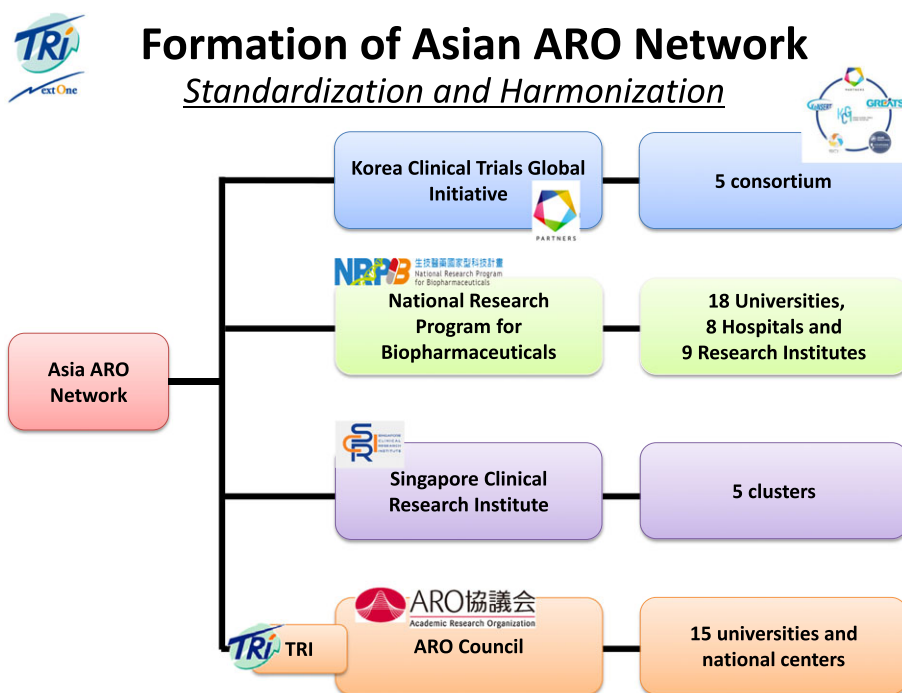


FIGURE 1 The Asia Academic Research Organization (ARO) Network

Concurrently, ARO Council is focusing to form disease-specific consortia in global countries (Figure 2). For each specific disease, a registry has been constructed, often in each country, or sometimes collaboratively. Registries should be combined, which is believed not to be a difficult task. An initial step will be catalog creation of the registry to enable sharing of the greatest common measures of those data that have been collected across the registries. For such data sharing, a good leading example, which has already taken a concrete approach towards standardization and harmonization, is the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, for which there has been excellent cooperation among research groups.¹⁴

The next milestone will be to initiate clinical trials concurrently and to obtain regulatory approvals simultaneously for academia-originated candidate products within 5 years. Plans are in progress to expand this framework to the United States and EU. This November ARO Network meeting in the United States was a first step in the direction of the vision to conduct global clinical trials and concurrent regulatory approvals of academia-developed innovation within 10 years.

Topics that emanated from representatives of US academic institutions and related clinical research endeavors centered around (1) infrastructure, modernization, and development of a research enterprise to support various types of research, including academic and biopharmaceutical and government research; (2) data sharing and promoting standardization and harmonization to enable effective aggregation and interpretation of data; (3) planning for data sharing from the start; and (4) developing resources and common data repositories to support cross-organizational research.

Presenters from AROs and related research organizations based in the United States included Dr Daniel Ford (Johns Hopkins), Marianne Chase (Neurological Clinical Research Institute [NCRI], Massachusetts General Hospital/Harvard), Dr Erika Augustine (University of

Rochester), Dr Jonathan Silverstein (University of Pittsburgh), John Speakman (New York University), Dr Matthew Cowperthwaite and Dr Chris Webb (University of Texas Dell Medical School), Dr Ida Sim (University of California, San Francisco, and Vivli), Dr Sam Volchenboum (University of Chicago), and Dr Melissa Haendel (Oregon Health & Science University).

Dr Daniel Ford of Johns Hopkins spoke about the need for academic medical centers involved in research to be able to support principal investigators who are involved (both at sites and at the university) in research that is commercially, not-for-profit, or federally funded. Dr Ford emphasized that, despite the progress made in using EHR data for research, there will not be efficiencies in accessing data for clinical research unless there is a similar infrastructure across the various studies. Without a common way to access research information, study start-up is delayed, start-up costs are higher, additional training is needed, and error rates are higher. In addition to the need for standardization and infrastructure harmonization, it is important to make the effort to integrate research into usual clinical practices while developing efficient and safe methods to access and share data in multicenter research studies.

John Speakman (New York University Langone Medical Center) concurred with the need for the central IT function of the medical center to support multiple different medical departments in their research interests and to ensure "pharma-grade" quality and regulatory compliance. These include study design, setup and management, data management (large scale and traditional), and data analytics. The NYU Langone Research IT group is well on their way to implementing the aspiration for their DataCore services. They have employed CDISC experts to focus on the standardization and harmonization activities. Specifically, the aspirations for the NYU DataCore Clinical Research Data Management include standardization of data format (ie, CDISC), standardization of reports, DSMC reports, a case report form library,

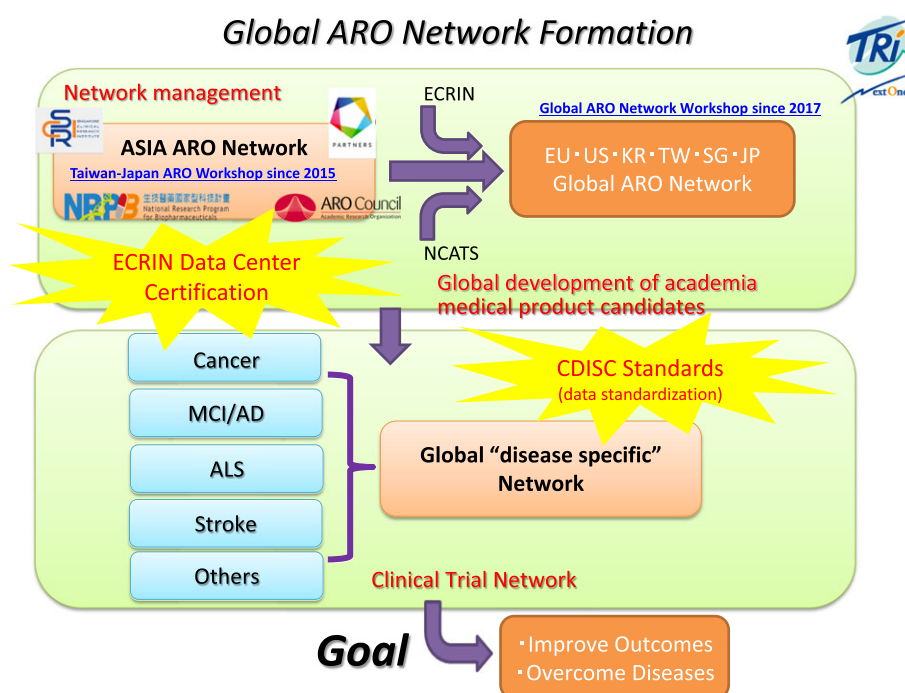


FIGURE 2 Academic Research Organization (ARO) Network focus on disease-specific registries to enable multinational clinical trials

centralized oversight (administration) of study databases, expert back-office support, 21CFR11 compliance, process flow for IT study build, and technology (including electronic data capture [EDC] and EHR integration).

Marianne Chase, from the NCRI/Massachusetts General Hospital,¹⁵ Harvard, presented on the academic leadership, methods, capabilities, and technology tools that have been developed and are being employed by the NCRI, which supports both the NIH-funded NeuroNEXT Network (25 sites in the United States specializing in neurology) and the NEALS Network (119 sites around the world involved in research on ALS). Tools include platforms that are compliant with regulations and enable EDC, regulatory document management, and clinical trial management in addition to PRO. NCRI also has a biorepository and registries for these therapeutic areas.

Dr Matthew Cowperthwaite spoke about research within St. David's HealthCare, which is based in Austin, TX, operates eight acute care hospitals and more than 100 clinical sites across Texas. St. David's is a partnership between Hospital Corporation of America (HCA) Healthcare, the largest hospital network in the United States; St. David's Foundation, which is the largest charitable not-for-profit foundation in Central Texas; and the Georgetown Health Foundation. St. David's maintains large clinical trials programs in cardiology (electrophysiology and interventional cardiology) and neuroscience. Across the HCA enterprise, research strengths include oncology, cardiology, neuroscience, and infection prevention. Sarah Cannon, the Global Cancer Institute of HCA, acts as the Site Management Organization (SMO) for oncology research sites and physicians in the United States and the United Kingdom. Sarah Cannon also operates a large Clinical Research Organization (CRO) to run studies for industry partners and also functions as an SMO for cardiology research in certain HCA divisions. HCA is highly focused on data standardization (normalizing data from 177 hospitals in an enterprise data warehouse), building and aligning collaborative networks of physicians and executive leaders, and improving processes such as communication channels, study start-up processes, corporate policies, and QA programs. HCA is a Sentinel Data Partner on the CDM Harmonization project described previously.

Dr Erika Augustine described the University of Rochester's Center for Health and Technology (CHeT), which has three core areas of focus: Clinical Trials Coordinating Center, Material Science Services Unit, and Clinical Research Innovation. They partner with academic groups, private groups, and industry sponsors. Examples of innovative tools they have employed include wearable sensors, smartphones, mobile apps, telemedicine, and virtual studies. Activities they are doing to modernize the way clinical trials are implemented include virtual visits, objective measures of disease including wearables and apps, disease modeling, data visualization, data-driven site selection, and risk-based monitoring. CHeT is experienced in conducting and participating in collaborative clinical research, has developed approved regulatory compounds, and focuses on programs that bring stakeholders together to advance change.

Dr Chris Webb leads the research strategy for the new Dell Medical School at the University of Texas at Austin, which has a unique opportunity to change the academic health environment to better serve society, as a new medical school associated with a major

university. Its first class of medical students matriculated in 2016. The vision, mission, and tagline ("Rethink Everything") speak to this opportunity. Dr Webb also spoke about the desire to incorporate research into "everything" and that every patient visit should be a research visit. There is keen interest within this new medical school in building an LHS as evidenced by the hiring of an experienced individual to lead this initiative. The mission is to revolution how people get and stay healthy by improving health in our community as a model for the nation, evolving new models of person-centered, multidisciplinary care that reward value, advancing the pace of research and innovation to improve health, educating leaders who transform health care, and redesigning the academic health environment to better serve society.

Dr Jonathan Silverstein from the University of Pittsburgh further described an LHS as one that gets the right care to people when they need it and then captures the results for improvement. The Institute of Medicine (now the National Academy of Medicine) held a series of workshops on the Digital Infrastructure for the Learning Health System. "We seek the development of a learning health system that is designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety, and value in health care" (US National Academies of Medicine/Institute of Medicine).¹⁶

Dr Silverstein went further to discuss service-oriented science¹⁷ and data science as a service, which led to requirements for a Data Commons¹⁸ and the definition of a new role, a Clinical Research Information Officer (CRIO), a role that is "critical to managing the interface between systems and research needs and can provide dedicated leadership and governance models charged with designing, deploying and leveraging various information resources to advance research."¹⁹ Dr Silverstein then described a number of related projects ongoing with the Research Informatics Office at University of Pittsburgh. These include building a research data warehouse and a cohort finder using EHR data, participating in the NIH/NCATS ACT Project (focused on the accrual of patients to clinical trials); NIH All of Us; PCORI-funded research; the TIES Cancer Research Network (TCRN) oncology project to share tissues and data using tools based on natural language processing; and the Pittsburgh Genome Resource Repository at the University of Pittsburgh and Pittsburgh Supercomputing Center (PSC), which provides data management and computing infrastructure to support use of national genome data resources for personalized medicine research. Dr Silverstein's presentation began by showing the relationship between obtaining new data and being able to reuse existing data based upon whether data are structured or unstructured. This challenge, summarized as collecting data consistently and completely, is at the core of being able to more efficiently use information/data to support LHSs.

Dr Ida Sim of the University of California, San Francisco, has experience in creating tools to search publications using structured data. During this workshop, she described Vivli, a new nonprofit with a mission to "promote, coordinate, and facilitate analysis across clinical research data through the creation and implementation of a sustainable global data-sharing enterprise."²⁰ Vivli consists of a user-friendly, secure, state-of-the-art cloud-based data sharing and analytics

platform that will serve the international community, including trials from any disease, country, sponsor, funder, or investigator in conducting open searches, with robust security and modern tools and technologies. Vivli makes use of the Cochrane ontology²¹ based upon PICO (Population, Intervention/Comparison, Outcomes) and provides a place for clinical research data to be shared if there is not already and agreed location for such sharing. The objective is to catalyze the evolution of clinical research towards and effective culture of data sharing and data reuse.

A diagram of the Vivli platform is shown in Figure 3.

Dr Sam Volchenbom of the University of Chicago and its Center for Research Informatics spoke of his experience developing a Cancer Data Commons with a Call to Action to have harmonized data leading to shared data, samples, and data with universal identifiers, envisioning data collection, and sharing at all stages of care, with an overarching goal of “all data from all patients at all times.” Building a Data Commons addresses current issues, including the fact that analyzing big data is time-consuming and costly and does not always provide useful results due to lack of standards. In addition, having phenotype/clinical data with universal identifiers harmonized into an easily-accessible Commons facilitates linking to genomic data and enriching the ability of researchers to utilize this important information. Based upon experience in creating a Data Commons for pediatric cancer, Dr Volchenbom shared a step-wise paradigm he has developed for this purpose.²²

Dr Melissa Haendel of the Oregon Health & Science University described the challenge in biomedicine around integrating knowledge across domains and disciplines, especially basic research and clinical care. This integration will inform clinical diagnoses, especially in cases of rare diseases when the genotype and phenotype information of a new patient can be compared with other known cases. There is a current gap (referred to by Dr Christopher Chute as the “chasm of semantic despair”) between basic science information and clinical information.

The Monarch Initiative^{23,24} is a global, translational consortium that leverages a large body of structured and integrated genetic information to provide sophisticated algorithms for phenotype comparison within and across species. Monarch provides computational tools, interactive visualization, and bioinformatics analyses based upon multiple information sources to shorten the information exchange path for clinicians and researchers. The ultimate goal is to advance personalized medicine.

To address the need to bridge resources and activities such as what Monarch provides and achieves with clinical care and health data, NIH/NCATS has awarded a grant to Dr Haendel as the lead investigator of a collaborative opportunity to create a new CD2H as a cross-CTSA initiative.²⁵ The priorities for the new CD2H will to support a “vibrant and evolving informatics ecosystem,” including

- support and enhancement of a collaborative informatics community;
- development of Good Data Practices (GDP);
- promotion of software standards for interoperability;
- growth of collaborative innovation across informatics tools, methods, and processes;
- advancement of cutting edge biomedical research informatics;
- data science education for CTSA Program researchers; and
- novel methods and tools for the evaluation of the impact of these activities to enhance health care through data and informatics.

Dr Haendel also spoke about a number of ongoing projects and initiatives that will be leveraged to achieve the goals of the CD2H. These include the Human Phenotype Ontology (HPO),^{26,27} which captures symptoms and phenotypic findings to assist in making computational phenotype comparisons with known diseases and patients; diagnostic tools such as Exomiser²⁸; and the Global Alliance for Genomics and Health (GA4GH) Matchmaker Exchange. She described the GA4GH “phenopackets” exchange format, which enables data sharing of computable phenotype information. She indicated the value of collaborating with international organizations such as the ARO Network to catalyze a “vital translational community focused on the collaborative application of integrated genotype and phenotype data to aid human disease discovery and diagnosis.”

The CD2H project led by Dr Haendel includes other collaborators. Named in the award are Kristi Holmes, PhD, Northwestern University; Sean Mooney, PhD, University of Washington; Christopher Chute, MD, Johns Hopkins University School of Medicine; and John Wilbanks, Sage Bionetworks. The Data to Health Coordination Center grant, which supports the new center, has been awarded to these institutions, together with The Scripps Research Institute, Washington University in St. Louis, The University of Iowa, and The Jackson Laboratory. They will all work with other centers that have received NIH/NCATS CTSA.

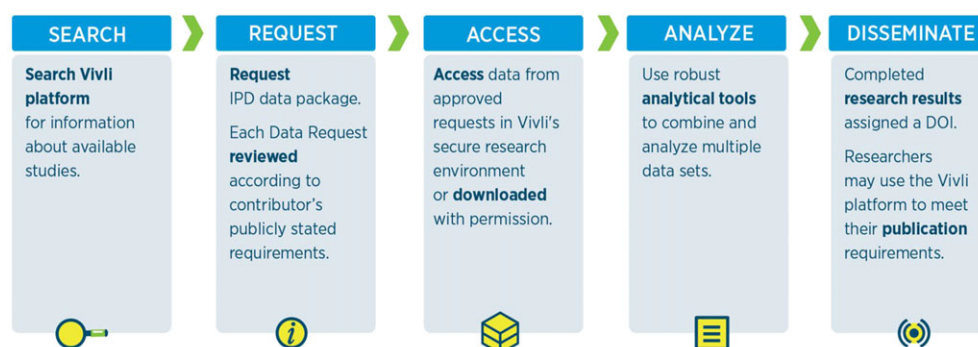


FIGURE 3 The Vivli Process

7 | WORKSHOP SUMMARY

Discussions at the end of the Global ARO Network Workshop of November 2017 were led by Dr Chris Austin (US NIH/NCATS) and Dr Masanori Fukushima (Japan's Translational Research Informatics Center, Foundation for Biomedical Research and Innovation). The focus was on the primary roadblocks to streamlining global clinical research projects. The participants listed the need for infrastructure investment, perseverance, metathesaurus/standards, convergence, and core expertise. Participants voiced their concern that this topic needs further discussion and agreement on how to alleviate the barriers. The achievement of true LHSs will require such discussion, exploration, consensus building, and funding with the necessary contingencies that this be tied to data sharing and other activities that are not currently incentivized through government funded research, at least not in the United States. Lessons can be gleaned from the IMI in the EU, ECRIN, AMED, and the ARO Council in Europe and Japan. The recent CORBEL Consensus Document on data sharing and reuse from Europe will also serve as an excellent resource.

Dr Fukushima stressed in his closing remarks the need for AROs around the world to work together, to harmonize and standardize and to ensure that data are housed in data centers that meet certification criteria. The goal towards long, healthy lives and towards treating diseases relies on global data sharing. Participants in the Global ARO Council Workshop expressed interest in continuing to be included in further discussions to facilitate the progress towards global research and true LHSs.

8 | EMERGING THEMES AND RECOMMENDATIONS FOR REALIZING THE GLOBAL ARO NETWORK VISION

Themes from the previous Global ARO Network Workshops have included Harmonization and Standardization, New Paradigm for Medical Science based on Data Sharing, and RWD and Disruptive Innovation. These broad themes resonated throughout the presentations. From a tactical perspective, an action plan has now been published by the Learning Health Community²⁹ to identify key areas that need focus in order to realize an LHS. Principles and recommendations are available for Data Sharing in the CORBEL Consensus Document.

Emerging themes and recommendations emanating from this specific research-focused ARO Network Workshop that could improve research processes and serve to better connect research with practice could be summarized as follows.

- Plan research with the end in mind and have a robust infrastructure, potentially preceding global clinical research studies with disease-specific registries for sharing data.

Understanding the objectives of the research, for example, what you would like to have in the results tables of a publication or on the label of a new therapy, the statistical requirements, and how the data should be formatted and what the key endpoints are will help in planning the study before it starts. Having an infrastructure that

supports this planning will facilitate implementation. An initial step that is proving to be very important for the Asia ARO Network, prior to initiating global clinical trials, is to create a registry and begin sharing data in this manner. Harmonization and standards can emerge when registries are combined and/or data are shared through such means; these can then be applied to the design and data collection for clinical research studies.

- Globally harmonize and encourage adoption of common standards.

The more broadly adopted standards (for data, metadata, models, and terminology), the easier the sharing of data and communication of meaning along with that data. Starting a new standard (especially if it is a one-off or proprietary standard) exacerbates the issues we currently face, creating inefficiencies and increasing costs and resource needs. New standards should be extensions of existing foundational standards and not redundant standards that have already been developed.

- Integrate from beginning to end—from basic science to clinical research to health care.

This is the premise behind translational medicine, from “bench to bedside,” precision medicine, the CD2H initiative, and a number of CTSA initiatives. A number of areas still need to be addressed for this to become a global reality, but acknowledging its value and funding initiatives in this area is an important step.

- Address privacy, confidentiality, legal, and other issues around data sharing.

Recommendations in the CORBEL Consensus Document have paved the way for such issues to be addressed; however, they (especially legal issues) still use excessive time, energy, and resources that could be dedicated to research.

- Implement new technologies that are innovative and standards-based.

There is no shortage of new technologies available that could be applied to research. FDA is encouraging their use. However, many still need validation and a way to readily share data in standard formats before they can be widely adopted and applied.

- Accommodate RWD to accumulate RWE.

As with new technologies, the use of RWD for research remains a challenge. FDA has issued Guidance on this topic, and the 21st Century Cures Act calls for its use. However, the variation and implementation specificity of EHRs and moving target of health care standards around the world have been a significant barrier to obtaining high-quality RWD for research purposes.

- Improve and accelerate the sharing process from research to improve clinical care decisions.

Patients who participate in research studies do so not only for their own benefit but also for the greater good. They expect their data to be used wisely and for learnings to be shared as quickly as possible, not to be trapped in a research silo or lost due to lack of quality or standards. Computable knowledge is essential as is data sharing. Accelerating learning health cycles forms the basis for LHSs.

CONFLICT OF INTEREST

Ida Sim is a co-founder of Vivli and receives consulting fees from Vivli for technical and strategic consultation. The remaining authors affirm that they have no conflicts of interest to disclose.

REFERENCES

- Molzon JA, Giaquinto A, Lindstrom L, et al. The value and benefits of the International Conference on Harmonisation to drug regulatory authorities: advancing harmonization for better public health. *Clin Pharmacol Ther.* 2011;89(4):503-512.
- <http://www.ich.org/home.html>
- eCommon Technical Document (eCTD) <https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm333969.pdf>. Accessed April 11, 2018
- <https://www.meddra.org/>
- <http://bbc.com/news/business-38305512>. Accessed October 12, 2017
- <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-064.html>. Accessed April 13, 2018
- https://www.niams.nih.gov/sites/default/files/files/Protocol_Template_V1.0_040717.pdf. Accessed April 12, 2018
- <http://www.transceleratebiopharmainc.com/initiatives/common-protocol-template/>. Accessed April 13, 2018
- <https://www.fda.gov/Drugs/ucm412998.htm>. Accessed October 10, 2017
- Data Standards Catalog <https://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm#Catalog>
- "Real world evidence—what is it and what can it tell us?", NEJM Rachel E. Sherman, M.D., M.P.H., Steven A. Anderson, Ph.D., M.P.P., Gerald J. Dal Pan, M.D., M.H.S., Gerry W. Gray, Ph.D., Thomas Gross, M.D., M.P.H., Nina L. Hunter, Ph.D., Lisa LaVange, Ph.D., Danica Marinac-Dabic, M.D., Ph.D., Peter W. Marks, M.D., Ph.D., Melissa A. Robb, B.S.N., M.S., Jeffrey Shuren, M.D., J.D., Robert Temple, M.D., Janet Woodcock, M.D., Lilly Q. Yue, Ph.D., and Robert M. Califf, M.D. <https://www.fda.gov/downloads/drugs/developmentapprovalprocess/smallbusinessassistance/ucm572939.pdf>
- Ohmann C et al. *BMJ Open.* 2017;7:e018647. <http://bmjopen.bmj.com/cgi/content/full/bmjopen-2017-018647?ijkey=79SivGTa9igpfbN&keytype=ref>
- <http://www.ecriin.org/>
- <http://www.adni.loni.usc.edu/>
- <https://www.massgeneral.org/ncri>
- Grossmann C, Powers B, McGinnis JM (Eds). *Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care: Workshop Series Summary.* Washington (DC): National Academies Press (US); 2012.
- Foster I. Service-oriented science. *Sci AAAS.* 2005;308(5723):814.
- Grossman RL, Heath A, Murphy M, Patterson M, Wells W. A case for data commons: toward data science as a service. *Comp Sci Eng.* 2016;18(5):10-20.
- Embi PJ, Tachinardi U, Lussier Y, Starren J, Silverstein J. Integrating governance of research informatics and health care IT across an enterprise: experiences from the trenches. *AMIA Jt Summits Transl Sci Proc.* 2013 Mar 18;2013:60-62. PMID: PMC3845750
- Bierer B, Li R, Barnes M, Sim I. A global, neutral platform for sharing trial data. *NEJM.* 2016;374(25):2411-2413. <https://doi.org/10.1056/NEJMp1605348>
- <http://linkeddata.cochrane.org/pico-ontology>
- Volchenboum SL, Cox SM, Heath A, Resnick A, Cohn SL, Grossman R. Data commons to support pediatric cancer research. *Am Soc Clin Oncol Educ Book.* 2017;37:746-752. https://doi.org/10.14694/EDBK_175029. Review. PubMed PMID: 28561664.
- www.monarchinitiative.org
- The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species <https://academic.oup.com/nar/article/45/D1/D712/2605791>
- <https://ctsa.ncats.nih.gov/cd2h/home>
- <http://www.human-phenotype-ontology.org>
- <https://academic.oup.com/nar/article/45/D1/D865/2574174>
- <http://exomiser.github.io/Exomiser/>
- Learning Health Community Action Plan – LHS Journal; www.learninghealth.org

How to cite this article: Fukushima M, Austin C, Sato N, et al. The Global academic research organization network: Data sharing to cure diseases and enable learning health systems. *Learn Health Sys.* 2019;3:e10073. <https://doi.org/10.1002/lrh2.10073>