Data Governance and Social Challenges to Sharing Clinical Data for Research: Problems & Solutions

This is the first workshop in a two-workshop series. Each workshop will cover similar topics from two perspectives: Workshop 1 will be focused on data governance and social challenges associated with clinical data sharing, and Workshop 2 will focus on technical challenges for these topics (e.g., data standards, data quality, software for record linkage). "Clinical data" refers to electronic health records and other real-world data related to healthcare.[1]

Duration: Workshop 1 is expected to last 1.5 days **over** July 13–14.

Goals:

Overarching Goal: To develop actionable recommendations and inform NLM's strategic plan by identifying high-impact data governance strategies and the roles NLM can play in enabling responsible and innovative clinical data sharing for research.

Goal 1: Compare Data Sharing Models

Examine how different clinical data sharing approaches (distributed networks, repository-mediated sharing, centralized platforms) implement data governance, protections, and assess the tradeoffs of each approach.

Goal 2: Surface Critical Governance and Social Challenges

Identify the most consequential data governance (e.g., consent, legal, policy) and social challenges limiting clinical data sharing across existing models, including challenges related to stakeholder engagement (e.g., patients, provider organizations, researchers, and funders).

Goal 3: Identify Solutions and Opportunities for Alignment

Highlight areas where incompatible policies, access controls, consent frameworks, or institutional practices create friction and explore approaches to improve compatibility, interoperability, and data reuse. Solutions can include common policies and frameworks, recommendations, best practices and identify significant gaps that NLM should focus on.

[1] https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

Day 1 – July 13, 2026

9:00 – 9:30 AM	**Introductions / Kickoff**	Workshop Goals & Scope • NIH definition of data sharing • Introduction to the three models for sharing clinical data for research **Speakers**: Workshop Co-Chairs: Lisa Federer, Kin Wah Fung, Valerie Cotton
9:30 – 10:15 AM	**Session 1A: Overview Presentation Lightning Talks** (10 min/each) Topics: ▪ Key data governance challenges (e.g., legal, consent, policy) ▪ Access process ▪ Lessons learned	**Models of Clinical Data Sharing – Lightning Talks** **Model 1 ("distributed"):** Academic institutions, hospitals and clinics allow shared access to clinical data in a **distributed** network • **Effort & Speaker**: OHDSI (Hua Xu, Yale University) • **Effort & Speaker**: PCORnet (Nik Koscielniak) **Model 2 ("repositories"):** Researchers extract clinical data and share through established **data repositories** • **Effort & Speaker**: DATA COUNTS (Susan Gregurick, NIH ODSS and Shannon Silkensen, NCI) • **Effort & Speaker**: NHGRI AnVIL & eMERGE (Robert Carroll & Anna Lewis)
10:15-10:30 AM	**Break**	
10:30-11:00 AM	**Session 1B: Overview Lightning Talks continued**	**Model 3 ("centralized")**: **Centralized** systems specifically designed to integrate clinical data for research use • **Effort & Speaker**: All of Us (John Giannini, NIH) • **Effort & Speaker**: N3C (Christopher Dillon, NCATS) • **Effort & Speaker**: CMS ResDAC /VRDC (TBD)
11:00-11:45 AM	**Session 2: Panel Discussion with all Session 1 Speakers** **Lessons learned from specific challenges**	***** the following topics will be split between the 2 Panel Sessions on Day 1 & Day 2***** **Moderated by:** Lisa Federer How each speaker/model addresses these challenges: • Data access processes – pros/cons of each • Balancing data sharing & data protections

		• Consent or lack of consent for research/data sharing • Hesitation to share the data • Fostering transparency and reproducibility • Other policy/legal or social barriers • Enclave requirements and impacts of system interoperability • Splitting multi-modal data from one study across repositories • Risks/challenges when applying AI/ML cross the data lifecycle **If time**: questions will be opened to the audience (in-person & virtual)
11:45 AM – 12:45 PM	**Lunch**	
12:45 – 1:30 PM	**Session 3: User Perspectives** (7 min/each) Users/researchers describe their experiences using, accessing, combining data, or submitting data for Session 1 efforts & models	Users/researchers present • **Effort & Speaker**: PCORnet/REACHnet (Model 1) (Tom Carton) • **Effort & Speaker**: TBD (Model 2) (TBD) • **Effort & Speaker**: N3C (Model 3) Emily Pfaff • **Effort & Speaker**: All of Us (Model 3) (Julie Holm) **Q&A with the audience moderated by:** Lauren Oliveira Hashiguchi (17 min)
1:30 – 1:45 PM	**Breakout Session Introduction** (only for onsite attendees)	**Instructions by:** Kimberly Thomas - Groups move to breakout areas, assigned based on preferences collected at registration - Each breakout group will discuss one Clinical Data Sharing Model
1:45 – 2:30 PM	**Session 4A: Breakout** Logistics: NIH planning	**Questions for each group**: - *Should NLM support this model – why or why not?* - *What is NLM's role in implementing or facilitating this model?*

	committee to serve as break out group moderators, each group will designate notetakers, each group assigns someone to report out. Session will not be videocast.	
2:30 – 2:45 PM	**Break to Reconvene in Auditorium**	
2:45 – 3:45 PM	**Session 4B: Breakout Groups Report Out** *(10 min each, assumes 6 groups)*	**Introduction by:** Kimberly Thomas - One person from each group reports out
3:45 – 3:4:00 PM	**Break**	
4:00 – 4:45 PM	**Session 5: Presentation + Discussion**	**Leveraging TEFCA's federal data exchange/sharing infrastructure to share EHR data with researchers** **Speaker**: Jawanna Henry, ASTP/ONC **Discussion with audience, moderated by**: Kin Wah Fung
4:45 – 5:00 PM	**Wrap up Day 1 Recap**	Workshop Co-Chairs: Lisa Federer, Kin Wah Fung, Valerie Cotton

Day 2 – July 14, 2026

9:00 – 9:15 AM	**Welcome & Goals for Day 2**	**Introductions**: Workshop Co-Chair: Lisa Federer
9:15 – 10:00 AM	**Session 6:**	**Policies and frameworks that shape clinical data sharing strategy** • **Framework & Speaker:** Clinical Care Data Research Principles (Dina Paltoo, NLM) (15 min) • **Framework & Speaker:** Controlled access data sharing in repositories (Cheryl Jacobs, OSP) (15 min)

		• **Q&A with the audience moderated by**: Lisa Federer (15 min)
10:00 – 10:15 AM	**Break**	
10:15 – 11:00 AM	**Session 7: Continued discussion with all Session 1 Speakers**	*****Continue from Day 1 Panel Session***** **Moderated by:** Valerie Cotton Continue discussion from Day 1/Session 2 – consider topics raised by other sessions (including breakouts) **If time**: questions will be opened to the audience (in-person & virtual)
11:00 – 11:45 AM	**Session 8: Data governance for participant-level linkage across multiple data sources**	Data governance frameworks for sharing, using, and linking participant-level data for research • **Framework & Speaker:** Frameworks for Streamlining Decisions about Data Linkage, Rebecca Rosen and Elizabeth Clerkin, NICHD ODSS (15 min) • **Framework & Speaker:** A Proof-of-Concept Model to Streamline Processes for Data Access & Linkage", Heather K. Basehore, NCI (15 min) • **Q&A with the audience moderated by**: Valerie Cotton (15 min)
11:45 – 12:00 PM	**Wrap up Day 2**	NLM: Reflections and next steps