Know Before You Go: Lessons from the creation of a registry-based Federated Network

Sarah Gasman[1], Clair Blacketer[1], Federico Zazzetti [2], Ashley Orillion[1], Erika Noss[1], Anna Sheahan[2]

[1] Johnson & Johnson, Spring House, PA, USA; [2] Johnson & Johnson, Horsham, PA, USA

Background

Registry-based Federated Networks harmonize data from multiple disease-specific registries, enabling efficient collaboration and increasing the statistical power of observational research.[1] The heterogeneity of these data sources provides unique opportunities and challenges in data standardization, study design, and interpretation of results.[2] We discuss recommendations based on lessons learned from the creation of a global federated network of five registries harmonized to the OMOP Common Data Model (CDM) for the study of Systemic Lupus Erythematous (SLE). Key recommendations for future projects include establishing clear objectives for the federated network and ensuring that the patient populations and available data of partner registries align with these objectives.

Methods and Results

The following topics are important to consider when establishing a new federated network comprised of registries.

1. **Determine the purpose and scope of the federated network, including research questions of interest to key stakeholders.**

   Creating a federated network requires a significant investment of time and resources. Prior to undertaking this effort, key stakeholders should align on the specific objectives of the federated network. A broad mission statement is not sufficiently comprehensive to facilitate the necessary coordination between stakeholders and potential partners. Clearly outlining the intended scope provides a framework for collaboration and analyses within the federated network. Registries can vary substantially; well-defined priorities inform the registry recruitment process and ensure that participating registries are able to contribute to key research questions.

2. **Evaluate the founding objectives of planned partner registries and ensure patient populations align with the priorities of the federated network.**

   Registries are often created for a specific purpose; this dictates which patients are recruited to participate. Therefore, it is critical that registry objectives align with the objectives of the broader federated network. Registries enriched for patients with more severe disease or specific disease manifestations, for example, produce insights that cannot be generalized to a broader patient population. Variances in registry patient populations should be characterized and accounted for when interpreting results from studies using federated network data.[3,4] For this project, as depicted in Table 1, registry patient populations overlapped, but varied in disease severity and cohort type (open vs. closed). We managed differences in patient populations by creating clear Cohort Definitions with specific inclusion criteria.

Table 1: Variation in recruited patient populations across five registries

Registry	Patient Population
Registry 1	All patients with SLE, ongoing recruitment
Registry 2	All patients with SLE or Lupus Nephritis, ongoing recruitment
Registry 3	Patients with rheumatologic disease including SLE. Patient-reported data. Ongoing enrollment.
Registry 4	Patients with SLE and/or Lupus Nephritis, including pre-treatment. Specific patient proportion recruitment goals. Closed cohort.
Registry 5	Patients with SLE or incomplete SLE. Recruited from previous study based on disease manifestations of interest. Closed cohort.

3. Ensure data collected by registries align with the priorities of the federated network.

Data collected vary by registry, even among those studying similar patient populations. If a federated network is created to answer specific research questions, confirm that a partner registry contains the data needed to answer those questions. For example, if biomarker data is of interest to stakeholders, registries that collect such data should be recruited as partners.

Prior to conducting analyses using data from this federated network, we performed extensive database characterization, feasibility analyses, diagnostics, and quality assurance to develop a comprehensive understanding of available (and missing) data. This project's registries contain a variety of disease activity and damage measures, as shown in Table 2; comprehensive characterization enabled us to develop feasible research questions based on available data and address confounding variables in analyses.[3,4]

Table 2: Assessments available across five registries

Registry	Disease Activity Assessments	Damage Assessments	Patient-Reported Outcome Assessments
Registry 1	PGA | MEX-SLEDAI | SLEDAI	SLICC Damage Index	LupusQol | FACIT | SF-36 | WPAI
Registry 2	PGA | SLEDAI	SLICC Damage Index	SF-36
Registry 3	SLAQ	BILD	ClinHaq | EQ-5D | SF-36
Registry 4	PGA | SLEDAI	SLICC Damage Index	LupusQol | WPAI
Registry 5	PGA | SLAQ | SLEDAI	SLICC Damage Index	EQ-5D | LIT

BILD: Brief Index of Lupus Damage, **ClinHaq:** Clinical Health Assessment Questionnaire, **EQ-5D:** Europe Quality of Life, **FACIT:** Functional Assessment of Chronic Illness Therapy, **LupusQol:** Lupus Quality of Life, **PGA:** Physician Global Assessment, **SLICC Damage Index:** Systemic Lupus International Collaborating Clinics Damage Index, **SF-36:** 36-Item Short Form Survey, **SLAQ:** Systemic Lupus Activity Questionnaire, **SLEDAI:** Systemic Lupus Erythematosus Disease Activity Index, **WPAI:** Impairment of Work Activity.

4. Verify time periods covered by partner registries overlap with other registries in the federated network.

It is important to ensure that time periods covered by registries in a federated network overlap in order to successfully increase statistical power and to prevent the addition of confounding variables. Non-overlapping time periods limit the utility of longitudinal analyses and make it difficult to interpret results influenced by the calendar year.

For example, standards of care and available treatments for a disease change over time. If each registry represents entirely different treatment availability, data cannot be pooled or accurately compared. Table 3 below depicts the time periods covered by registries in this project. In our analyses of treatment utilization, we prioritized years where registries overlapped (2015 to 2022) and included context from external sources such as new drug approval dates.

Table 3: Observation time periods of participating registries

Registry	**Observation Time Period**
Registry 1	2012 to 2024
Registry 2	2013 to 2021
Registry 3	1999 to 2023
Registry 4	2019 to 2022
Registry 5	2015 to 2022

Conclusion

Registry-based federated networks can be powerful tools for evidence generation in observational research. Based on our experience from the creation of an SLE-specific, registry-based federated network, we recommend addressing the following topics to enable the more efficient creation of more effective federated networks:

1. Determine the purpose and scope of the planned federated network
2. Ensure that selected partner registries align with federated network objectives, specifically:
    a. Registry purpose
    b. Patient population
    c. Data availability
3. Verify that the time period covered by a partner registry overlaps with others in the federated network

References

1. van Baalen V, Didden E-M, Rosenberg D, et al. Increase transparency and reproducibility of real-world evidence in rare diseases through disease-specific Federated Data Networks. Pharmacoepidemiol Drug Saf. 2024;33:e5778. doi: 10.1002/pds.5778

2. Gini, Rosa, et al. “Describing Diversity of Real World Data Sources in Pharmacoepidemiologic Studies: The DIVERSE Scoping Review.” *Pharmacoepidemiology and Drug Safety*, vol. 33, no. 5, 1 May 2024, https://doi.org/10.1002/pds.5787
3. Dempster S, Schuemie, M. Chapter 19 Study Steps. In: The Book of OHDSI [Internet]. 2021. Available from: https://ohdsi.github.io/TheBookOfOhdsi/
4. Kostka K, Klebanov G, Dempster S. Chapter 20 OHDSI Network Research. In: The Book of OHDSI [Internet]. 2021. Available from: https://ohdsi.github.io/TheBookOfOhdsi/