haematologica
Journal of the Ferrata Storti Foundation

Transforming lymphoma outcomes through international real-world data collaborations: the Global Lymphoma Registry Alliance roadmap

by Eliza A. Hawkes, Eliza Chung, Mark J. Bishton, Natacha Bolanos, Andrew Evens, Massimo Federico, Georg Hess, Matthew J. Maurer, Zoe K. McQuilten, Susan K. Parsons, Michelle L. Poon, Karin E. Smedby, Diego Villa, Lorna Warwick, Erica M. Wood, James R. Cerhan, Martine Chamuleau, Chan Y. Cheah, Christopher R. Flowers, Herves Ghesquieres, Astrid Pavlovsky, Marek Trneny, Adam Olszewski, and Tarec C. El-Galaly. Lymphoma Registry Alliance members.

Received: April 12, 2026.
Accepted: July 31, 2026.

Citation: Eliza A. Hawkes, Eliza Chung, Mark J. Bishton, Natacha Bolanos, Andrew Evens, Massimo Federico, Georg Hess, Matthew J. Maurer, Zoe K. McQuilten, Susan K. Parsons, Michelle L. Poon, Karin E. Smedby, Diego Villa, Lorna Warwick, Erica M. Wood, James R. Cerhan, Martine Chamuleau, Chan Y. Cheah, Christopher R. Flowers, Herves Ghesquieres, Astrid Pavlovsky, Marek Trneny, Adam Olszewski, and Tarec C. El-Galaly. Lymphoma Registry Alliance members. Transforming lymphoma outcomes through international real-world data collaborations: the Global Lymphoma Registry Alliance roadmap. Haematologica. 2026 Aug 13. doi: 10.3324/haematol.2026.301008 [Epub ahead of print]

Transforming lymphoma outcomes through international real-world data collaborations: the Global Lymphoma Registry Alliance roadmap

Authors:

Eliza A. Hawkes[* 1,2], Eliza Chung[*1], Mark J. Bishton[3-5], Natacha Bolanos[6], Andrew Evens[7], Massimo Federico[8], Georg Hess[9], Matthew J. Maurer[10], Zoe K. McQuilten[1], Susan K. Parsons[11], Michelle L. Poon[12], Karin E. Smedby[13], Diego Villa[14], Lorna Warwick[6], Erica M. Wood[1], James R. Cerhan[10], Martine Chamuleau[15], Chan Y. Cheah[16], Christopher R. Flowers[17], Herve Ghesquieres[18], Astrid Pavlosky[19], Marek Trneny[20], Adam Olszewski[#21], Tarec C. El-Galaly[# 22,23] on behalf of Lymphoma Registry Alliance members

1. School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia.
2. Olivia Newton-John Cancer Research Institute at Austin Health, Melbourne, Victoria, Australia.
3. University of Nottingham, Nottingham, United Kingdom.
4. Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom.
5. National Disease Registration Service, NHS-England.
6. Lymphoma Coalition, Canada.
7. Division of Blood Disorders, Rutgers Cancer Institute New Jersey, New Brunswick, NJ, USA.
8. CHIMOMO Department, University of Modena and Reggio Emilia, Modena, Italy.
9. Department of Haematology and Oncology, University of Mainz, Germany; German Lymphoma Alliance & European Mantle Cell Lymphoma Network.
10. Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA.
11. Division of Hematology/Oncology and Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA.
12. Department of Hematology Oncology, National University Cancer Institute Singapore, National University Hospital, Singapore.
13. Divison of Clinical Epidemiology, Department of Medicine Solna, Karolinska Institutet, and Department of Hematology, Karolinska University Hospital, Stockholm, Sweden.
14. BC Cancer Centre for Lymphoid Cancer and University of British Columbia, Vancouver, British Columbia, Canada.
15. Amsterdam UMC, location VUMC, Cancer Center Amsterdam, the Netherlands.
16. Department of Haematology, Sir Charles Gairdner Hospital and Division of Internal Medicine, University of Western Australia, Perth, Western Australia, Australia.
17. Department of Lymphoma/Myeloma, Division of Cancer Medicine, MD Anderson Cancer Center, Houston, TX, USA.
18. Department of Hematology, Lyon Sud Hospital Hospices Civils de Lyon, Pierre-Benite, France.
19. Centro de Hematologia Pavlovsky/ Fundaleu, Buenos Aires, Argentina.
20. Charles University, General Hospital, Prague, Czech Rebublic.
21. Brown University Health Cancer Institute at Rhode Island Hospital and The Miriam Hospital, RI, USA.
22. Departments of Hematology, Clinical Epidemiology, and Molecular Medicine, Aarhus University Hospital, Denmark.
23. Department of Clinical Medicine, Aarhus University, Denmark.

*Contributed equally as first authors.
# Contributed equally as senior authors.

On behalf of the Global Lymphoma Registry Alliance members
Corresponding Author: Prof Eliza A Hawkes, eliza.hawkes@onjcri.org.au

Running head: A Global Lymphoma Registry Alliance to leverage real-world data

Data-sharing statement: N/A

Funding: The LyRA has been generously funded by Lymphoma Coalition.

Prof Hawkes is supported by an Australian NHMRC Emerging Leader fellowship.

Ms Chung acknowledges support from the Australian Government Research Training Program Scholarship [DOI: https://doi.org/10.82133/C42F-K220].

Drs. Cerhan, Flowers, and Maurer were supported by U.S. National Cancer Institute grant U01 CA195568.

Drs. Parsons and Evens are supported by U.S. National Institute of Health grant R01 CA262265.

Authors' disclosures:

EAH: Research funding (paid to institution): Roche, Bristol Myers Squibb, Merck KgA, Astra Zeneca, TG therapeutics and Merck. Consultant or advisory roles (*Paid to institution): Roche*, Merck Sharpe & Dohme*, Astra Zeneca*, Gilead, Antengene*, Novartis*, Regeneron, Janssen*, Specialised Therapeutics*, Sobi*, GSK*Travel expenses: Astra Zeneca, Genmab, Abbvie

MJB: Consulting/advisory/honoraria Janssen, Roche, Gilead, Recordati, Abbvie, Eli Lilly, Abbvie, Astra Zeneca. Speakers bureau Incyte, Recordati. Research funding Roche. Travel/accommodation/expenses Roche

MJM: Research funding from BMS and Roche/Genentech.

MLP: Advisory board for Regeneron/ PI for Regeneron. PI for Genmab studies. Research support from Beigene.

KES: Honoraria from Abbvie and Incyte

DV: Research funding (to the institution) from Roche, AstraZeneca, BeOne. Honoraria and advisory boards from Janssen, AstraZeneca, BeOne, Abbvie, Roche, BMS/Celgene, Kite/Gilead, Incyte, Kyowa Kirin, ONO Therapeutics.

JRC: Research funding utilizing the LEO cohort from Genentech, Genmab and BMS.

MC: research funding BMS, Gilead, AbbVie GenMab. Advisory AbbVie, Byondis, BMS

CYC: consulting/advisory/honoraria – Roche, Janssen, Gilead, AstraZenecca, Lilly, Beigene, Menarini, Dizal, Abbvie, Genmab, Sobi, CRISPR Therapeutics. BMS, Regeneron. Speaker's bureau - Janssen, Astrazenecca, Beigene, Genmab, Abbvie, Roche, MSD. Research funding – BMS, Roche, Abbvie; MSD, Lilly. Travel expenses - Lilly, Beigene

HG: REALYSA cohort was financially supported by Roche, Takeda, Janssen, Amgen, Celgene-Bristol Myers Squibb, Astra-Zeneca, AbbVie, GSK. HG declares following competing interests: consultancy, Roche, BMS, Takeda; honoraria, Gilead, Roche, BMS, Abbvie, Takeda.

EC, NB, AE, MF, GH, ZKM, SKP, LW, EMW, SKP, CRF, AP, MT, AO, TCEG: no disclosures

Authors' contributions:

EAH, EC, TCEG, AO conceptualized and wrote the manauscript. MJB, NB, AE, MF, GH, MJM, ZKM, SKP, MLP, KES, DV, LW, EMW, JRC, MC, CYC, CRF, HG, AP, MT reviewed and edited the manuscript. EAH, EC, NB, ZKM, LW, and EMW acquired funding. All authors approved the manuscript.

Abstract

Lymphomas comprise a complex and heterogeneous group of malignancies which pose challenges in understanding their epidemiology, pathobiology, treatment responses and long-term outcomes. Evolving diagnostic classification and fast-paced therapy development compound these challenges. Robust real-world data (RWD) collection and analysis using clinical registries can contribute significantly to address gaps in understanding of practice variation and provide evidence for health technology assessments.

However, to maximize the impact of lymphoma registries, and those in other diseases, there is a compelling need for global collaboration, data harmonization and automated integration between registries and other large datasets. Technologies that enable safer data sharing are already available, but historical legal frameworks and evolving privacy concerns are not keeping pace, undermining their intended purpose and limiting the full potential of available high-quality RWD to improve patient care.

This White Paper written by the Global Lymphoma Registry Alliance (LyRA) discusses the importance and value of lymphoma registries for different stakeholders as well as benefits of forming a global alliance of the registry network. An alliance such as LyRA serves both academic endeavors and public interest through collaboration between patient and community organizations, policy-makers, regulatory authorities, industry and others seeking to use RWD. Bringing these stakeholders together and raising awareness more broadly will facilitate timely clinical trial result contextualization and innovation in public-private collaborations on novel trial emulations and designs, including external comparator cohorts. The LyRA leadership propose strategies for overcoming barriers to facilitate these key collaborations towards improving patient outcomes on a global scale.

Introduction

Lymphoma care is entering a new era characterized by accelerated regulatory approvals, rapidly escalating treatment costs, and persistent global inequities in access to modern therapies. Increasingly, new drugs receive accelerated approval on the basis of single-arm studies conducted in highly selected populations, often relying on surrogate endpoints and limited follow-up. Patients, clinicians, payers, and policymakers are therefore making high-stakes decisions during health technology assessments (HTA) with incomplete evidence regarding long-term effectiveness, comparative value, safety, and equitable access. Real-world data (RWD) are now considered an essential complement to clinical trials, providing pre- and post-marketing outcome data from populations treated in the real world **(Figure 1)**. In this environment, high-quality RWD defined here as data relating to patient health or experience or care delivery collected outside the highly controlled clinical trials context[1], are no longer optional adjuncts to clinical research—they are essential infrastructure for responsible clinical decision-making, regulatory oversight, and sustainable healthcare planning **(Figure 1)**.[2]

Registries are purpose-designed RWD repositories that inform disease trends, treatment effectiveness, toxicity, and long-term survivorship which collectively generate real world evidence (RWE). Some lymphoma registries have been systematically collecting, storing, and analysing such data prospectively for decades. Registries serve multiple purposes: some are intended purely for research, while others fulfill administrative or quality assurance functions, like monitoring population incidence, care implementation or resource utilization.[3, 4,5-9] The ability of individual registries to support research is determined by design, goals, data custodianship, stakeholder composition, governance structures for data access, and funding models.

The importance of registries is increasingly recognized internationally by stakeholders in academia, government, regulatory, industry, and patient advocacy, leading to expansion in their size and number.[10] However, the currently fragmented registry landscape needs harmonization to establish their full potential in generating comprehensive applicable knowledge that improves patient care. Existing data sharing regulations are restrictive and burdensome, creating technical and administrative data transfer barriers critical to collaboration. This reality does not serve patients' or societal interests.

Here, we provide a roadmap for worldwide registry collaboration and integration through the foundation of the Global Lymphoma Registry Alliance (LyRA). The authors serve on the LyRA steering committee, comprise leaders in lymphoma real-world research, registry science, a network of patient organizations, and lead their respective own registries. The roadmap should not be viewed as an aspiration to establish a single global registry. Rather, it is directive from academic experts and patient organizations to engage in dialogue with regulatory authorities, ethicists, broader patient communities, HTA bodies, and industry. The intended audience for this White Paper are all these stakeholders, with the hope of increasing mutual understanding of the needs and enhance collaboration to advance quality and use of RWD with the ultimate goal to improve individual patient care with a global perspective, promote sustainable healthcare systems, and strengthen regulatory decisions for innovative therapies.

Benefits of forming a global alliance for lymphoma-specific registries

The LyRA has identified 49 existing lymphoma-based registries (including research consortia and long-term cohort studies) in 35 jurisdictions around the world, with significant variation in methods, data management, disease scope, funding and governance models (**Figure 2, Supplementary Table 1**). The LyRA has adopted an inclusive approach and welcomes systematic real-world datasets on lymphoma to participate. The identified registries and databases become members of the alliance when their leadership, for example the Principal Investigator, expressed interest.

Many individual registries and internationally established consortia have provided invaluable insights into lymphoma epidemiology and outcomes. However, such evidence is most commonly regional or national, but infrequently international. Registry inclusion criteria and coverage are typically disease-based but may also focus on procedures or products not limited to lymphoma, for example, chimeric antigen receptor T-cell

therapy (CAR-T) and transplant registries[11, 12] or national population-based pan-cancer registries. Consequently, they can lack data fields of importance to lymphoma clinicians and researchers, as disease-specific variables are not available or sufficiently granular. Further, data that require prospective collection directly from patients (e.g., patient-reported outcomes) are often not available.

The fragmented registry sector undermines a global view of lymphoma epidemiology and outcomes and hinders registries from meeting established quality thresholds.[13] These limitations are amplified for lymphoma due to considerable diagnostic and therapeutic complexity. Ensuring global faithful representation of the same clinical entity in lymphoma is required. Furthermore in rare lymphomas, timely analyses of therapeutic access, treatment algorithms or evolving diagnostic criteria require multinational collaborations to achieve sufficient sample sizes.

RWD quality remains a priority for registries whilst also balancing sustainability to collect long-term follow-up data for large populations in this environment with limited resources. The lack of clear international guidance for data collection and management in lymphoma research is suboptimal. Misaligned fields across registries disrupts data interoperability and lead to unwanted wastage of research data in analyses.

LyRA recognizes these systemic issues and offers opportunity to advance the sector, not through creation of a centralized international registry, but through streamlined knowledge sharing and expert consensus, setting quality benchmarks and core data elements that meet minimum stakeholder standards, harmonizing governance and data sharing processes. A collaboration framework offers significant advantages as different registries can contribute different types of data (e.g. clinical data, genomics, patient-reported outcomes, late effects). These can then be integrated (including with clinical trials data) to analyze a variety of questions, with the understanding that not every registry can contribute to every analysis, but the data aggregation allows, for example, different registries to contribute to specific components of a multi-state model.

The time is now for global registry collaboration and alignment in lymphoma

Lymphoma exemplifies the challenges of data integration relevant to many diseases. It is the most common hematologic malignancy and encompasses over 80 heterogeneous subtypes.[14] Many subtypes are rare, with limited understanding of their natural history and few randomized trials guiding treatment decisions. Substantial geographic and social disparities exist in lymphoma epidemiology, clinical care, and outcomes. Existing datasets underrepresent large parts of the world, despite the endemic nature of certain subtypes (e.g., Burkitt lymphoma in sub-Saharan Africa[15], NK/T-cell lymphoma in Asia[16] and adult T-cell leukemia/lymphoma in central and south America). At the same time, access to transformative therapies such as CAR-T and bispecific antibodies remains highly variable across jurisdictions.

The highly competitive drug development environment results in approvals frequently based on non-randomized studies conducted with decidedly selective patient populations, surrogate endpoints, limited follow-up and dosing strategies that identify maximum tolerated doses, rather than minimum effective dosing.[17] This framework amplifies knowledge gaps and global variation in care. RWD have thus become essential for data-informed clinical decisions. RWD also inform trial design, measure effect of implementation and generalizability of trial results, and evaluate long-term safety for indications where randomized trials are not feasible. Their value increases as countries need to prioritize limited healthcare resources in the context of increasing costs. The Food and Drug Administration and the European Medicines Agency also now recognized RWD as synthetic control arms in clinical trials and as a valid data source for drug and device applications,[18] although implementation has been slow.

Meanwhile, emergent new information technologies (IT) permit significant opportunities to capture, share and analyse lymphoma RWD while maintaining quality and security. Advances in automated electronic health record (EHR) data extraction, enhanced storage capacity and new sharing capabilities allow handling of large and complex datasets. Artificial intelligence and machine learning tools can produce sophisticated predictive modelling when supplied large training datasets, but their performance depends fundamentally on data

integrity, fit for purpose, and representativeness. Data privacy restrictions intended to protect patients now represent the most significant barrier to fully harnessing these methodological and technical innovations and may eventually do the opposite.

An international approach to addressing barriers of RWD collection and use

Barriers to successful registry collaborations are multifactorial and widely recognized across disease contexts, with our readily available solutions summarised in **Table 1**. We provide our goals with respect to these barriers and potential solutions **(Figure 3).**

***LyRA goal:* Achieving synergy in governance frameworks and stakeholder input via balanced leadership**

Successful RWD collaboration relies on effective engagement of diverse stakeholders including clinicians, health authorities, funding bodies, industry and policy-makers. The most important stakeholders are patients, both existing (who contribute their data), as well as future, whose management will rely on RWD availability for RWE generation. Historically the absence of input from people with lived experience in decision-making pertaining to risks and benefits of RWD research led to increasingly divergent views of patient benefit between clinician, investigators and authorities such as ethics committees and data protection bodies. Future progress requires formal inclusion of people with lived experience to provide important perspectives on clinical value and privacy risks as well as ethical considerations and levels of data protection required.

All stakeholders are vital to successful RWE generation from global RWD, but have different views and concerns. In addition to ensuring privacy concerns from all levels are acknowledged and addressed through appropriate operational safeguards, complex variation in consent must be considered; many quality or public health registries do not operationally embed explicit informed consent, yet many research governance structures require this to access data. As RWD is non-interventional, carries minimal direct risks for patients, yet LyRA members have identified substantial administrative and contractual burdens arising from privacy concerns as key barriers to collaboration. This leads to unethical waste of opportunities to improve patient care **(Table 1).** By bringing together stakeholders, more rapid progress can be made in balancing the needs and requirements of all while overcoming barriers to accessing and using high-quality RWD.

LyRA solutions: strong leadership structure, stakeholder engagement, and adaptable framework of governance to foster collaboration

Our approach: LyRA will amplify the value of individual efforts by promoting collaboration, harmonization, and shared standards through a community of stakeholders, including clinicians, data managers, researchers and patient advocacy partners. By aligning methodologies and definitions across international registries whenever possible, LyRA aims to enhance data interoperability, enabling meaningful comparisons of clinical characteristics, treatment patterns, and patient outcomes across diverse populations and health systems. The LyRA structure aims to expand to other interest groups such as imaging experts and laboratory scientists incorporating or contributing to RWE generation to provide a resource for the broader medical and research communities and the general public **(Table 1).**

Leadership structure: We deliberately incorporate a geographically diverse and experienced steering committee with representation from disease area registry experts, patient advocacy groups, and methodologists (**Supplementary Table 2).**

Communication: Our multimedia communication is purposeful with a public website that has member functions for secure document sharing, social media presence, regular stakeholder meetings to garner feedback from the group and determine future directions **(Supplementary Table 3).**

Harnessing collective expertise: LyRA will harness the expertise of its members along with patient voices to understand, and where possible, align or synergise individual registry governance structures and data sharing agreements to foster collaboration among stakeholders.

Governance: LyRA has created an overarching governance plan to rigorously assess feasibility of proposals from external or member organizations and distribute to the alliance participating member registries. Individual registries can then "opt in" to projects that meet their own local quality benchmarks and regulatory requirements. Frameworks for research data agreements will be developed across the alliance by a working group including content experts and lawyers familiar with data privacy and protection laws, common and civil laws and neutral laws (e.g. Singaporean and Swiss laws), to reflect individual project type-for example grouped data sharing versus extracted individual data. This leadership structure will provide governance for review and approval of data requests from different entities e.g. academic, non-profit or industry, and guidance for permission of various levels of access for external parties requesting summary data (such as paid requests).

LyRA's governance and operational structures will undergo continuous evaluation and improvement by implementing globally agreed mechanisms for ongoing evaluation of alliance activities, monitoring of outcomes, and adaptation of strategies to address emerging challenges and opportunities.

Patient voice: LyRA deliberately embeds the patient voice at every level through strong partnership with the Lymphoma Coalition – a worldwide network of almost 100 patient organisations from more than 55 countries dedicated wholly or in part to supporting individuals affected by lymphoma. LyRA strives to include patient representatives on all leadership committees, initatives and working groups. LyRA will also advocate for more formal inclusion of patient perpectives in regulatory decisions, as well as assessment of new RWD studies.

LyRA goal: Minimising data fragmentation and variation through data standardization

Regulatory and government frameworks provide general recommendations to support quality in RWD initiatives and registries **(Supplementary Table 4)**, but no detailed guidance on disease-specific harmonization and lymphoma-specific common standards currently exist. Fragmentation and critical data gaps also arise in clinical trials which often track only selected events during long-term follow-up. Registries can address these gaps by providing long-term outcomes, rare or late adverse events, end-of-life care and development of prognostic models that perform reliably across diverse real-world populations **(Table 1)**.

LyRA solutions: Establishment of common standards and data harmonization

The initial assessment performed within LyRA has shown significant variation in operations, consent model, data collection and sharing, and capacity for benchmarking.[19] Examples of studies which successfully harmonized data items across large, disconnected datasets to generate clinically relevant outputs include the HoLISTIC (Hodgkin Lymphoma International STudy for Individual Care) Consortium,[20] which integrated data from 8 international clinical trials spanning two decades to generate a novel clinical prediction model, or the FLIPI-24 collaboration which achieved a similar objective using 10 observational cohorts in follicular lymphoma.[21]

LyRA will build harmonization frameworks through full registry data mapping and expert consensus to create core lymphoma-specific data sets for most critical variables across the network. This core dataset will form a template for research projects and to support development of new registries.

LyRA will build on core datasets and available data via global structured surveys and consensus models to create agreed CQIs.[22] This information will be used to develop consensus-based protocols and guidelines for data collection, storage, and sharing across participating registries globally.

LyRA also aims to play a pivotal role in supporting under-resourced environments which lack comprehensive registries and provide optimized approaches to new registries, educational resources and training programs to

promote best practices in registry operations. Our website will incorporate secure member areas for sharing templated documents and the investigator network, so new registries can build or share existing infrastructure.

These issues and planned solutions on a timeline are summarized in **Table 2.**

***LyRA goal:* Ensure interoperability of technological infrastructure**

Combining inter-jurisdictional patient-level data across the bespoke independent IT systems of different registries is challenging. These are rarely designed to link or stream data to external platforms. Even with similar individual variables, data entry and coding differences can generate semantic variations that requires recoding and introduces potential error. For example, a single lymphoma diagnosis can currently be entered using one of four distict diagnostic frameworks[14, 23-25] which are not always interchangeable.[26, 27] Likewise, follow-up and ascertainment of time-dependent endpoints vary, including inconsistent collection of response, relapse systemic anti-cancer therapy data, or transformation.[28] Moreover, piecemeal collection of detailed cytogenetic, molecular and epigenetic abnormalities can be extremely challenging to harmonize.

***LyRA solutions:* Data collection and management**

The minimal Common Oncology Data Elements (mCODE)[29] provides a framework for core oncology data elements to be built into any EHR, encompassing standardized terminology for patient, disease, treatment, and outcome variables. Universal adoption across LyRA member registries would facilitate future data abstraction and aggregation across heterogeneous systems and reduce inconsistencies.[29] Additionally, adherence to FAIR data principles (Findable, Accessible, Interoperable, and Reusable) would enhance the discoverability and usability of data for collaborative research.[30] However, individual investment in mCODE, FAIR, and other interoperable IT systems within each registry is unlikely to occur. Leveraging advanced technologies such as safe harbors and federated data analytic systems that enable computation to occur locally while sharing only aggregated results, can facilitate data integration whilst maintaining both data integrity and privacy according to registry-specific requirements and overcomes complex legal requirements. Establishing such research platforms requires significant upfront and continuous investments. An interim simpler version for collaboration would be sharing of statistical coding and adapting the code locally to the individual registries. This approach has been successfully used in several discovery-validation studies in lymphoma like the BLIPI for Burkitt lymphoma and R-IPI for relapsed/refractory diffuse large B-cell lymphoma.[31, 32]

Trustworthy data analyses requires a consolidated framework with a rigorous statistical basis that provides a standardised process for both data access and output quality assurance. LyRA governance committees, comprising multidisciplinary expertise in registry science, will act as an unbiased bodies to evaluate the scientific merit of suggested research projects and access requests. As demand grows for advanced analytical outputs, including AI-based models, systems must ensure compatibility with machine learning tools. A long-term goal is the adoption of trusted research environments (TRE), allowing approved investigators secure access to anonymized multi-source data. Shared cloud workspaces, with adequate cybersecurity, can provide access to research data, code libraries and analytics tools. To safeguard patient confidentiality, outputs will be subject to semi-automated and/or human review prior to release/access, ensuring that individual-level data remain protected within the system.

***LyRA goal:* Address privacy and ethical considerations**

Protecting patient privacy and ensuring adherence to ethical standards when sharing and combining sensitive health data are key LyRA priorities. Regulatory frameworks such as the European General Data Protection Regulation (GDPR) and US Health Insurance Portability and Accountability Act (HIPAA) must be rigorously followed. However, researchers and registry custodians face susbstantial variation in regulatory interpretations. While individual-level disease trajectory data are useful, de-identification can substantially mitigate risks if performed prior to data sharing and analysis. These include removing direct identifiers as well

as transformation of specific dates through consistent computational shifting or replacement with time-to-event variables.

Risks associated with sharing sensitive health information should not be underestimated, however, an open discussion is needed at all levels from local ethics committees to regulatory and government bodies about consequences of overly restrictive data-sharing regulations. The current landscape is especially detrimental for patients with rare diseases, where clinical management is based on expert consensus rather than robust evidence. This is due to analyses lacking sufficienty large, granular datasets, resulting partially from data sharing difficulties. The consequences are potentially serious. At the patient level, suboptimal management and inferior outcomes result. At the societal level, diminished data-deriven healthcare leads to resource waste on ineffective therapies or procedures. LyRA argues for a more nuanced and balanced discussion on data use that addresses privacy risk while explicitly considering ethical consequences of not using available data. A discussion that acknowledges and potentially quantifies the potential harm to future patients when medical practice cannot be improved through responsible learning from existing RWD and ensure these are fully understood by relevant authorities.

***LyRA solutions:* Global data sharing and integration**

LyRA will create and adopt agreed principles aligned within existing ethical and legal frameworks while advocating for practical solutions that facilitate responsible generation of RWE. This will include clear data sharing and integration guidelines and protocols that comply with regulatory requirements and uphold patient confidentiality.

i. The group will review and expand what works now in ethical and secure data sharing.
ii. LyRa aims to improve global transparency of data sharing processes between members by reviewing existing public material from key member registries, e.g. the Lymphoma Epidemiology of Outcomes (LEO),[9] and encouraging registries to publish data access processes.
iii. LyRA will work towards robust synergized approvals and aligned frameworks that permit secure and ethical data sharing, including developments of master documents and standardized data sharing agreements.
iv. LyRA will engage in health policy discussions in close collaboration with patient organisations via Lymphoma Coalition, industry and regulators, in particular to highlight the serious clinical consequences for future patients of not using available data to learn from past practice
v. Ultimately, LyRA will develop guidance on importance, relevance and risks assessments of new RWD studies in partnership with our partner patient organisations, industry collaborators and regulators. This guidance will provide frameworks to support researchers in negotiatations with local authorities. This will include carefully balanced considerations on ethical obligations to improve clinical practice with existing data against privacy risks in order to inform authorities, industry and regulators of a more patient-focused and holistic view of benefit/risk related to data access and usage.

Enhancing the value of registries for patients, clinicians, researchers, and policy makers

For RWD end-users including researchers, patient organisations, industry and regulatory bodies, LyRA's coordinated international framework can transform fragmented data into a powerful, reusable, and evolving resource. Standardized variables and data models will facilitate high-quality dataset integration and cross-country studies of prognosis, survivorship, and health equity. Investigators establishing new registries will benefit from LyRA's shared infrastructure, templates, and governance models.

For the broader ecosystem of industry, regulatory agencies, patient organisations and policy makers, LyRA will generate high-fidelity, globally comparable data to identify unmet needs and opportunities to reduce disparities in clinical care. LyRA will prioritize frameworks that enable patient-generated and registry data

merging fit for use by regulatory decision makers. Industry partners can use these insights to inform development of innovative therapies, identify populations for clinical trials, and guide access initiatives. LyRA aims to catalyze interventions to close global gaps in lymphoma care by tracking consequences of variable access to newer treatments such as CAR-T, bispecific antibodies and antibody-drug conjugates that have a major impact on lymphoma outcomes.

Ultimately, a robust, harmonized registry network will reinforce a cycle of improvement: better data enabling better science, better science informing better clinical care, and better care inspiring continued investment. Demonstrating the collective value of shared data, LyRA will attract sustained funding and participation, ensuring the registry's long-term viability. Most importantly, patients, as the central stakeholders, will benefit from accelerated translation of discoveries into equitable, high-quality lymphoma care worldwide, and gain access to high quality data to inform the advocacy endeavors.

Conclusion

In a time of fast-paced clinical development and increasingly strained health care systems, registries are an invaluable resource to advance lymphoma knowledge, improve patient outcomes and enable data-driven, sustainable healthcare planning. LyRA, a global alliance for lymphoma registries, represents a transformative approach to advance research, clinical care, and policy development in the field through collaboration between registries. By uniting efforts, sharing resources, and leveraging collective expertise, stakeholders can accelerate progress towards understanding the complexities of lymphoma and improving patient outcomes on a global scale. Overcoming challenges through collaborative solutions, engagement in policy discussions, and adopting technical solutions will be essential first steps for LyRA in realizing the full potential of data integration and fostering innovation in lymphoma research and care delivery worldwide. The next steps will require close dialogue with a broad range of stakeholders, including more patient communities, legal experts, pharmaceutical industry, and regulatory bodies, in order to achieve tangible, important progress. Nevertheless, we remain optimistic that advances can be made in several areas relevant for better opportunities for RWD research, since we are united by a common goal that is both indisputable and uncontroversial: whatever we can learn from past patients with lymphoma must be used to improve care for the next patients.

Disclaimer: The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the regulatory agencies, ethical committess or other organizations with which the authors are affiliated**.**

References

1. National Institute for Health and Care Excellence. UK NICE Real World Data. 2022 ; Available from: https://www.nice.org.uk/corporate/ecd9/chapter/introduction-to-real-world-evidence-in-nice-decision-making#what-is-real-world-data (Accessed on January 30, 2026)
2. El-Galaly TC, Cheah CY, Villa D. Real world data as a key element in precision medicine for lymphoid malignancies: potentials and pitfalls. Br J Haematol. 2019;186(3):409-419.
3. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743-1748.
4. Mainz J. Defining and classifying clinical indicators for quality improvement. Int J Qual Health Care. 2003;15(6):523-530.
5. Ghesquières H, Rossi C, Cherblanc F, et al. A French multicentric prospective prognostic cohort with epidemiological, clinical, biological and treatment information to improve knowledge on lymphoma patients: study protocol of the "REal world dAta in LYmphoma and survival in adults" (REALYSA) cohort. BMC Public Health. 2021;21(1):432.
6. Henson KE, Elliss-Brookes L, Coupland VH, et al. Data Resource Profile: National Cancer Registration Dataset in England. Int J Epidemiol. 2019;49(1):16-16h.
7. Lymphoma Related Diseases Registry Investigators. Improving outcomes for patients with lymphoma: design and development of the Australian and New Zealand Lymphoma and Related Diseases Registry. BMC Med Res Methodol. 2022;22(1):266.
8. Chang Y-S, Tuan Y-H, Gao J-W, et al. Impact of comorbidities on patients with Waldenström macroglobulinemia in Taiwan: a multicenter study. JCO Glob Oncol. 2025;11:e2400576.
9. Cerhan JR, Maurer MJ, Link BK, et al. The lymphoma epidemiology of outcomes cohort study: design, baseline characteristics, and early outcomes. Am J Hematol. 2024;99(3):408-421.
10. Tunaru F, Robinson DE, MacDougall A, Carpenter L. Use of real-world data and real-world evidence in NICE (UK) health technology appraisals of new therapeutics in oncology: a systematic review. BMJ Open. 2025;15(9):e102769.
11. Guerrettaz RD, Spellman SR, Page KM. Role of the CIBMTR biorepository and registry in precision transplantation research. Bone Marrow Transplant. 2025;60(2):161-164.
12. The European Society for Blood and Marrow Transplantation. The EBMT Registry. 2026 Available from: https://www.ebmt.org/registry/ebmt-registry (accessed on January 30, 2026)
13. Niemeyer A, Regierer AC, Hartz T, et al. Development of a maturity model for medical patient registries - a community approach. Stud Health Technol Inform. 2024;316:1262-1263.
14. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of haematolymphoid tumours: lymphoid neoplasms. Leukemia. 2022;36(7):1720-1748.
15. Mbulaiteye SM, Devesa SS. Burkitt lymphoma incidence in five continents. Hemato. 2022;3(3):434-453.
16. Kwong Y-L, Zhang H, Wang X, Tse E. Epidemiology of mature T-cell and NK-cell neoplasms: east and west. Lancet Reg Health West Pac. 2025;62:101646.
17. Lasch F, Carvalho JRB, Pothet C. Demonstration of major therapeutic advantage from a review of EU conditional marketing authorizations in oncology and hematology. Clin Pharmacol Ther. 2025;117(4):1098-1105.
18. US Food and Drug Administration. FDA Eliminates Major Barrier to Using Real-World Evidence in Drug and Device Application Reviews. 2025 15th December 2025 Available from: https://www.fda.gov/news-events/press-announcements/fda-eliminates-major-barrier-using-real-world-evidence-drug-and-device-application-reviews (accessed on January 30, 2026)
19. Chung E, Bishton M, Cherblanc F, et al. Mapping the methodology & attributes of global lymphoma registries: A Lymphoma Registry Alliance (LyRA) initiative. EHA Library. 2026;4208679:PS2128.

20. Rodday AM, Parsons SK, Upshaw JN, et al. The advanced-stage Hodgkin lymphoma international prognostic index: development and validation of a clinical prediction model from the HoLISTIC Consortium. J Clin Oncol. 2022;41(11):2076-2086.
21. Maurer MJ, Prochazka VK, El-Galaly TC, et al. FLIPI24: a modern prognostic model and clinical trial enrichment tool for newly diagnosed follicular lymphoma. J Clin Oncol. 2026;44(2):117-128.
22. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. Qual Saf Health Care. 2002;11(4):358-364.
23. World Health Organization. ICD-11: International classification of diseases (11th revision). 2019 Available from: https://icd.who.int/ (Accessed on January 30, 2026).
24. International Health Terminology Standards Organisation. Overview of SNOMED-CT. 2016. Available from: https://www.nlm.nih.gov/healthit/snomedct/snomed_overview.html (Accessed on January 30, 2026).
25. Campo E, Jaffe ES, Cook JR, et al. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. Blood. 2022;140(11):1229-1253.
26. Ko BS, Chen LJ, Huang HH, et al. Subtype-specific epidemiology of lymphoid malignancies in Taiwan compared to Japan and the United States, 2002-2012. Cancer Med. 2018;7(11):5820-5831.
27. Abro B, Maurer MJ, Habermann TM, et al. Real-world impact of differences in the WHO and ICC classifications of non-Hodgkin lymphoma: a LEO cohort study analysis. Blood. 2024;144(19):2063-2066.
28. Weir HK, Johnson CJ, Mariotto AB, et al. Evaluation of North American Association of Central Cancer Registries' (NAACCR) data for use in population-based cancer survival studies. J Natl Cancer Inst Monogr. 2014;2014(49):198-209.
29. Osterman TJ, Terry M, Miller RS. Improving cancer data interoperability: the promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. JCO Clin Cancer Inform. 2020;4:993-1001.
30. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016;3:160018.
31. Evens AM, Danilov A, Jagadeesh D, et al. Burkitt lymphoma in the modern era: real-world outcomes and prognostication across 30 US cancer centers. Blood. 2021;137(3):374-386.
32. Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. Blood. 2006;109(5):1857-1861.

Table 1: Problems and potential impact within registry sector plus solutions proposed by the Global Lymphoma Registry Alliance.

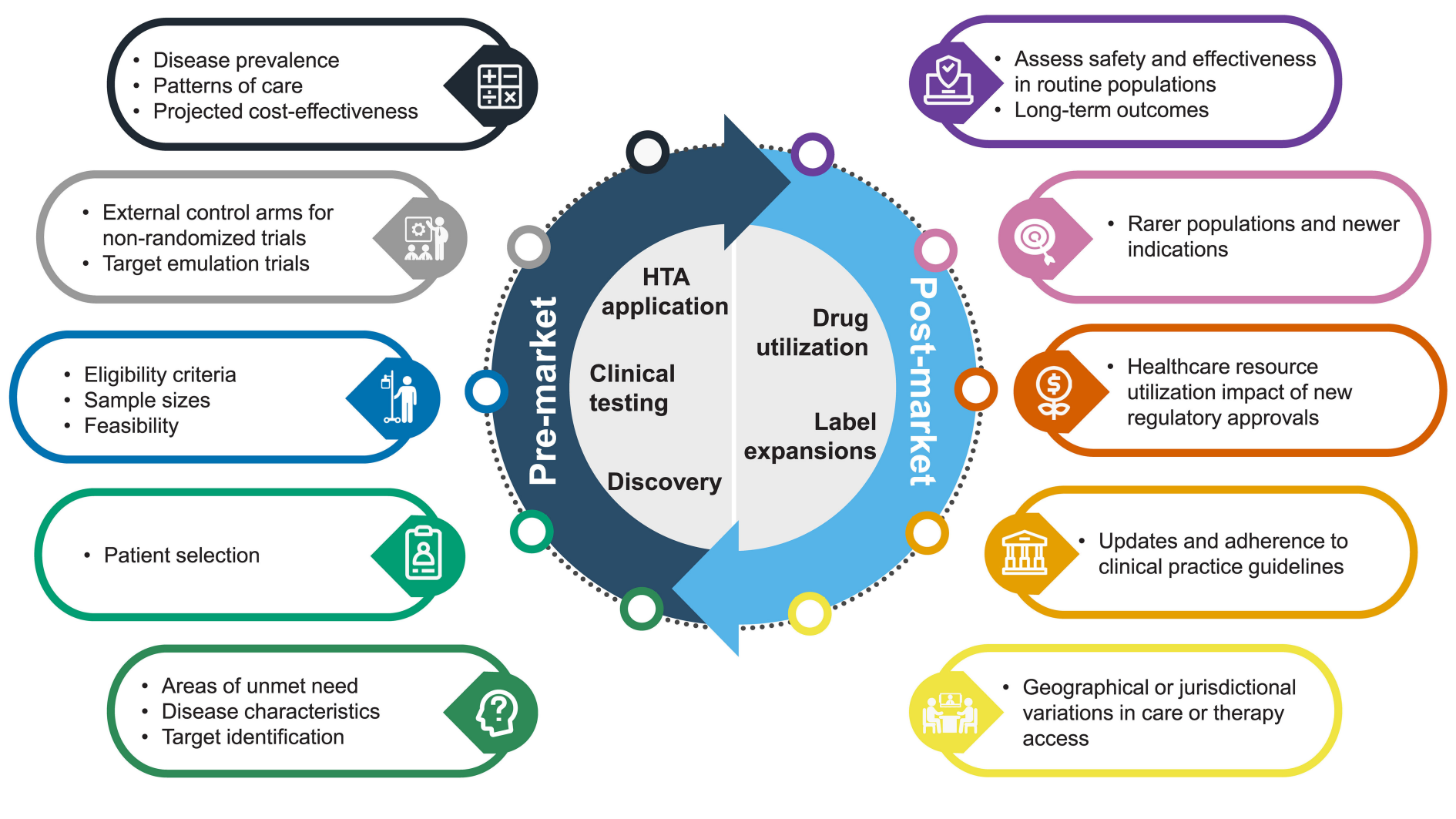
Problem	Stakeholders perpective	Potential impact	Proposed solution
Variations in data management, statistical rigor and analytical approaches	**Researchers, Regulators:** Data dictionaries lacking standardized design, formatting, and approaches to handling biases and missing data; Varied resourcing, statistical expertise, data granularity and management	*Aggravating reduced data completeness and quality *Incompatible and incomplete data fields across registries delay or prohibit dataset merges	*Enhanced and harmonized data quality and scope *Enhanced statistical power *Improved generalizability through data-merging opportunities and increased patient diversity
Data sharing and access challenges	**Researchers, Industry, Regulators:** *Complex, varied legal and/or regulatory requirements to collect/share/access data for research regardless of level of data anonymization; Government datasets not built for research/regulatory use requiring infrastructure and expertise to adapt	*High admin burdent but essential to ensure privacy, ethical data use- risk/benefit of sharing to regulators, funders, industry and patient care must be considered. *Existing onerous processes for agreement execution is time consuming. Lead to burdens too high for resources available.	*Resource optimisation from sharing information and expertise in data infrastructure, methodology and conduct reduces redundancy, coding and legal documents; maximizes efficiency and outputs *Uniting stakeholders strengthens advocacy for funding, regulatory support, patient engagement, and policy changes that benefit all patients.
Wastage and duplication in effort in research and administration	**Researchers:**Data collected in local databases because relevant data fields in national or established mature registries of all lymphomas unavailable. **Patients:** Consent to multiple projects of similar purpose & delayed care improvements	*Funding competition for similar infrastructure with rare targeted funding; funding gaps for registries and data repositories. *Data from multiple sources due poor linkages between datasets; excess/duplicate patient data collection is unethical; inconvenience/psychological distress.	*Harmonised funding models *Shared resources/procedures *Minimisaton of patient burden
Lack of routine collection on patient reported outcome data	**Researchers, Patients, Regulators:** Patient outcome and experience data uncommon in clinical trials; increasingly considered as essential in evaluating therapeutic efficacy, value and reimbursement	*Lacking patient feedback on treatment tolerance and adherence; Limits opportunities for clinicians to respond to the patient experience; Comparison of registry data on standard therapy in comparison with clinical trials limited by lack of patient-reported data.	*Involvement of Lymphoma Coalition in the governance of LyRA to embed patient voice on exploring the best practice and to inform aspects of patient management. *Explore methods to harmonise and combine patient generated data with registry datasets
Lack of existing collaboration between and with registries	**Researchers:** Insufficient compatibility of research administration and interoperability of datasets hinder robust partnership of registries, limiting opportunities for synergies.	*Registry infrastructure and network not being utilized to its potential for international benchmarking and large-scaled research projects	*Collaborative research initiatives promote faster discovery of novel biomarkers, predictive models and innovation; acts as platforms for clinical trials *Comprehensive patient characteristics, care patterns offering more complete, real-time patient data supporting robust, informed clinical decision making.
Lack of stable and long-term funding	**Regulators:** Gaps may exist in robust data, i.e. disease epidemiology, death indices; morbidity endpoints and treatment patterns **Researchers, Patients**: RWD analyses from registry data can be incomplete due to funding gaps and underresourcing for analyses	*Compromise in quality data informing patient care *Administrative value to regulators and policy makers of large-scale data is reduced by gaps in funding	*Coordinated approach to registry synergy means aligned, efficient data collection and usage will reduce duplication and waste; higher value datasets available to industry and regulators. *Global education and partnership between data custodians, researchers, regulators, policy makers and industry to promote and secure sustained support for registry activity.

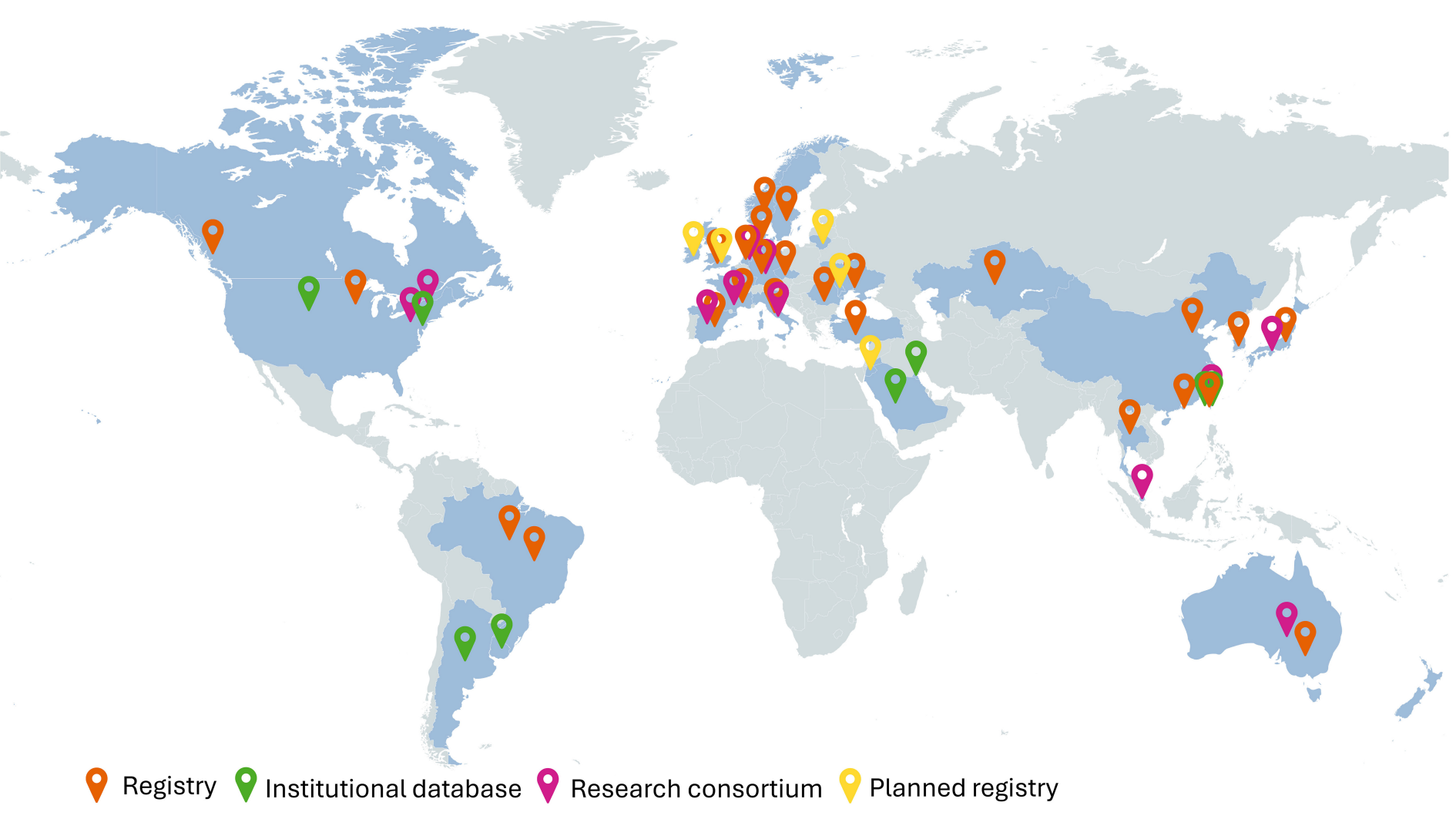
Table 2: Plans for global data integration.

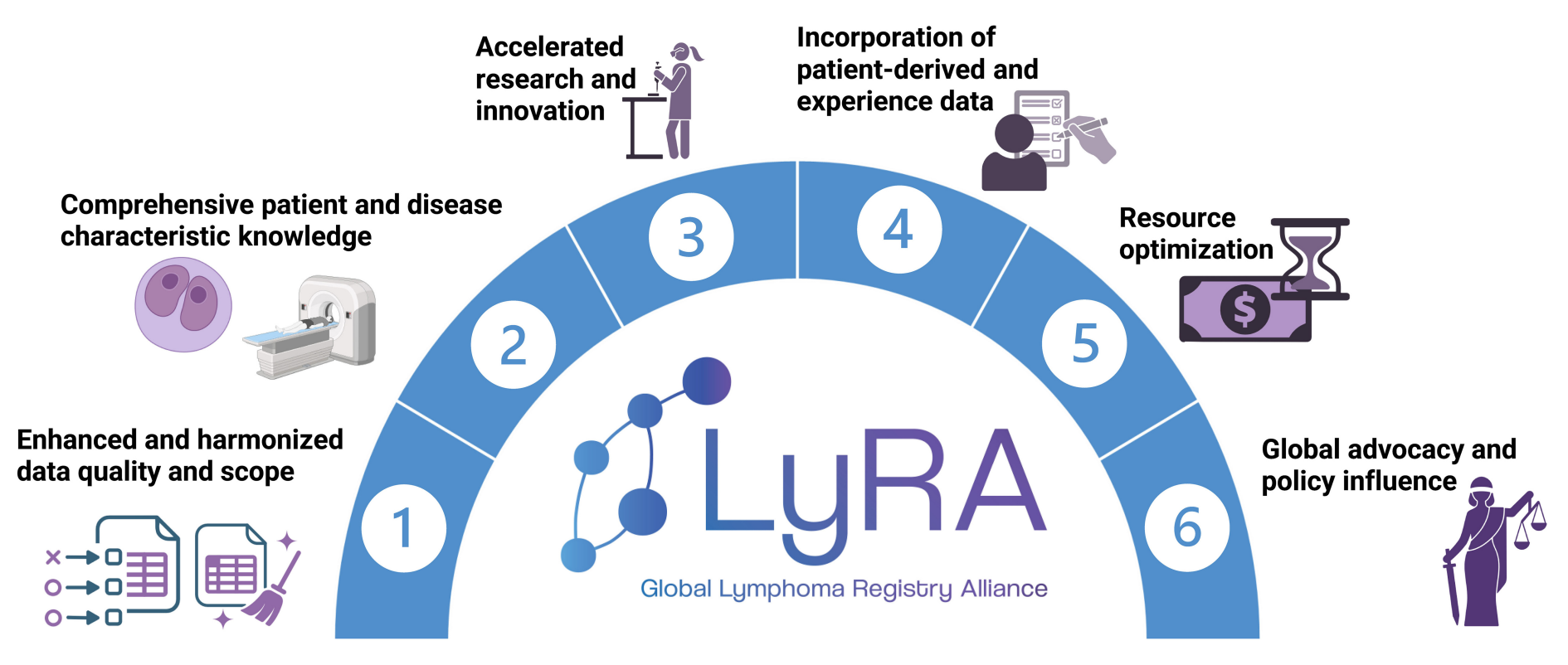
Problem	Immediate action	Short-term action	Long-term solution
Registries collect data with variable specifications	Collect existing dictionaries of main clinicipathologic variables across participating registries	Create a harmonized map of main clinicopathologic variables important for lymphoma research and practice	Optimize the variable specification to maximize relevance and facilitate consistent collection
Diagnostic patterns of lymphoma vary across settings	Identify methods of recording of lymphoma diagnosis in participating registries	Create a hierarchical diagnostic pattern conforming to the current WHO and ICC frameworks	Optimize diagnostic data collection to maximize diagnostic accuracy in registries from diverse resource settings
Registries collect data for varied purposes (cancer surveillance, clinical research, outcomes research) current lack of transparency of scope and boundary of existing data	Identify and classify the purpose and data scope of currently existing registries	Create a directory for researchers of existing registries including data collection, type and processes for access	Create a minimal common dataset for LyRA registry to provide basic interoperability regardless of registry purpose. Build and deploy a harmonized framework for projects with diverse objectives
Governance, data sharing, and privacy concerns vary across jurisdictions and settings	Collect and classify relevant policies, levels of data access, and data sharing options in participating registries	Create a LyRA governance body and define minimal requirements to satisfy both LyRA and local needs	Develop federated operational and statistical methods that allow data integration without data transfer

Figure Legends

Figure 1: Global real-world data applications in the health technology assessment cycle of lymphoma.
Figure 2: Global LyRA membership.
Figure 3: Benefits of having a Global Lymphoma Registry Alliance.
Figure 4: Global Lymphoma Registry Alliance's strategic roadmap to success.


• Disease prevalence
• Patterns of care
• Projected cost-effectiveness
• External control arms for non-randomized trials
• Target emulation trials
• Eligibility criteria
• Sample sizes
• Feasibility
• Patient selection
• Areas of unmet need
• Disease characteristics
• Target identification
Pre-market
HTA application
Clinical testing
Discovery
Drug utilization
Label expansions
Post-market
• Assess safety and effectiveness in routine populations
• Long-term outcomes
• Rarer populations and newer indications
• Healthcare resource utilization impact of new regulatory approvals
• Updates and adherence to clinical practice guidelines
• Geographical or jurisdictional variations in care or therapy access


Registry
Institutional database
Research consortium
Planned registry


Accelerated research and innovation
Incorporation of patient-derived and experience data
Comprehensive patient and disease characteristic knowledge
Resource optimization
Enhanced and harmonized data quality and scope
Global advocacy and policy influence
1
2
3
4
5
6
LyRA
Global Lymphoma Registry Alliance

Outcome

- Powerful global network for advocacy
- Compatible registry system
- Aligned documentation and processes for new registries

- Improved statistical power with global datasets
- Ability to decentralize data analyses

- Global framework for secure data sharing that meets stakeholder needs

- Improved statistical power with global datasets
- Efficient generation of global, high-quality data to improve outcomes in lymphoma

Stakeholders involved

Patients/those with lived experience

Legislators & policy makers

Regulatory bodies

Industry partners

Clinicians

LyRA initiatives

- Global network of lymphoma registry experts
- Secured web platform for investigators to share contact information, data access requirements of each registry

- Data mapping
- Data common standards
- Incorporate changes in diagnosis and management

- Analysis of patient, regulatory and clinician expectations in privacy and ethics of data sharing
- Identification of technical solutions to data analysis and federated learning opportunities

- Quarterly meetings on research topics
- Global approach to securing funding

Strategic areas and goals

Information sharing

- To synergize governance frameworks and global real-world data expertise

Data standardization

- Global understanding of existing data synergy and differences
- Core agreed minimum data fields and templates

Ethics and regulatory

- To establish practical solutions for collaboration in context of ethical perspectives

Research rigor and reproducibility

- validity, reliability.
- transparency,
- efficiency, and
- Implementation of large-scale data

Supplementary appendix

Supplementary Table 1: Registry membership of Global Lymphoma Registry Alliance.

Name	Leading jurisdictions
Registries	
BC Cancer Lymphoma Registry	Canada
Brazilian Non-Hodgkin T Cell Registry	Brazil
Brazilian Prospective Hodgkin Lymphoma registry	Brazil
CAEBV Registry	Japan
Chinese Lymphoma Registry	China
Cluj-Napoca Lymphoma Registry	Romania
Czech National Lymphoma Registry	Czech Republic
Danish National Lymphoma Registry (LYFO)	Denmark
Fondazione Italiana Linfoma (FIL)	Italy
Hong Kong Lymphoma Registry	Hong Kong
Kazakhstan Lymphoma Registry	Kazakhstan
Korean Lymphoma Registry	Korea
Lymphoma and Related Diseases Registry (LaRDR)	Australia and New Zealand
National Cancer Registration Service	England
Netherlands Cancer Registry	Netherlands
Norwegian Registry of Lymphoid Malignancies	Norway
Spanish Lymphoma Registry (RELINF)	Spain
Swedish Lymphoma Register (SLR)	Sweden
Taiwan Alliance for Hodgkin Lymphoma Outcome Study	Taiwan
Thai Lymphoma Study Group (TLSG)	Thailand
Ukrainian Lymphoma Registry (ULR)	Ukraine
Turkish Lymphoma Registry	Turkey
Prospective cohort studies	
Lymphoma Epidemiology of Outcomes (LEO) NCT02736357	United States
REal World Data in LYmphoma and Survival in Adults (REALYSA) NCT03869619	France
Institutional database	
Fundación para Combatir la Leucemia (FUNDALEU)	Argentina
Kuwait Cancer Control Center	Kuwait
King Saud University Medical City	Saudi Arabia
Memorial Sloan Kettering Cancer Centre	United States
National Taiwan University Hospital Waldenström macroglobulinemia (NTUH WM)	Taiwan
Taiwan Mantle Cell Lymphoma Network (TMCLN)	Taiwan
Unidad Académica de Hematología, Hospital de Clínicas, Facultad de Medicina, Universidad de la República (UDELAR)	Uruguay
University of Utah	United States
Research consortium	
Australasian Lymphoma Alliance	Australia and New Zealand
ENKTCL Registry	Japan
European Research Initiative on CLL	Spain
European Mantle Cell Lymphoma Registry (EMCLR)	Germany
German Lymphoma Alliance Registry	Germany
Global nLPHL One Working Group (GLOW)	United States
Hodgkin Lymphoma International Study for Individual Care (HoLISTIC)	United States

International Burkitt Lymphoma Network	Netherlands
International Lymphoma Epidemiology (InterLymph) Consortium	France
International T-cell Lymphoma Project (TCP)	Italy
Singapore Translational Cancer Consortium	Singapore
Taiwan National Health Insurance Alliance (TNHIA)	Taiwan
Registries in set up	
Chișinău Lymphoma Registry	Moldova
Vilnius Lymphoma Registry	Lithuania
King Hussein Cancer Center Amman	Jordan
United Kingdom T-Cell Registry	United Kingdom
Irish Lymphoma Registry	Ireland

Supplementary Table 2: Composition of Global Lymphoma Registry Alliance steering committee.

Names	**Jurisdictions**	**Affiliated bodies**
A/Prof Mark Bishton	England	European Mantle Cell Lymphoma Network
Prof James Cerhan	US	Lymphoma Epidemiology of Outcomes
Prof Martine Chamuleau	Netherlands	Internarional Burkitt Lymphoma Network
Prof Chan Cheah	Australia	Australasian Lymphoma Alliance
Prof Karin Ekström Smedby	Sweden	Swedish National Lymphoma Register
Prof Tarec El-Galaly	Denmark	Danish National Lymphoma Registry
Prof Andrew Evens	US	Hodgkin Lymphoma International Study for Individual Care
Prof Massimo Federico	Italy	International T-Cell Lymphoma Project International Burkitt Lymphoma Network Lymphoma Registries Project by Angela Serra Association for Cancer Research
Prof Christopher Flowers	US	Lymphoma Epidemiology of Outcomes
Prof Hervé Ghesquieres	France	REal world dAta in LYmphoma and Survival in Adults
Prof Eliza Hawkes (Founding chair)	Australia	Lymphoma and Related Diseases Registry
Prof Georg Heß	Germany	European Mantle Cell Lymphoma Network
A/Prof Matthew Maurer	US	Lymphoma Epidemiology of Outcomes
A/Prof Adam Olszewski	US	International Burkitt Lymphoma Network
Prof Susan Parsons	US	Hodgkin Lymphoma International Study for Individual Care
Dr Astrid Pavlovsky	Argentina	Fundación para Combatir la Leucemia
A/Prof Michelle Poon	Singapore	Singapore translational Cancer Consortium
Prof Marek Trněný	Czech Republic	European Mantle Cell Lymphoma Network
A/Prof Diego Villa	Canada	BC Cancer Lymphoma Registry
Ms Lorna Warwick Ms Natacha Bolanos	Canada	Lymphoma Coalition

Supplementary Table 3: Media strategies.

Category	**Details**	**Status**
Public website	Lymphomaregistryalliance.org	Active
Social media presence	LinkedIn	Active

Regular stakeholder meetings	Dedicated meetings during: American Society of Hematology Annual Meeting & Exposition: 2023, 2024 International Conference on Malignant Lymphoma: 2025 Virtual meetings: 2025/2026 European Hematology Association Congress: 2026	Ongoing

Supplementary Table 4: Examples of minimum quality standards considerations for robust RWD usage

Data collected	Data collection frameworks	Data analysis	Purpose fit
Provenance	High level of trust in systems -transparency	Appropriate study designs	Data available answers relevant questions
Reliability	Stakeholder engagement	Robust analytical methods	Evidence generation is appropriate
Missingness	Dynamic infrastructure	Machine learning models such as federated learning	Analysis fits aims

HoLISTIC methodology

HoLISTIC (Hodgkin Lymphoma International STudy for Individual Care) Consortium, established in 2018, exemplifies the possibility to build large datasets that integrate such fragmentation and gaps with adequate expertise, resources, and funding. The HOLISTIC operational team executed data use agreements (DUA) across multiple countries with specific attention to the privacy provisions introduced by the EU General Data Protection Regulation (GDPR). Additional ethics approvals from many other countries and groups were also obtained for secure transfer of detailed IPD from pivotal clinical trials across the disease spectrum for adult HL.

The HoLISTIC consortium gathered 90+ disease and methods experts to develop clinical decision models based on individual patient data (IPD) along the disease and treatment continuum from diagnosis/pre-treatment through long-term follow-up. The clinical trial database was augmented with IPD of patients captured in prominent cohorts and institutional, regional, and national HL registries that are enriched with information on non-HL outcomes. These registries and cohorts relied on diverse methods to collect and validate the incidence and severity of late effects, including linkage to claims databases, defining risk groups, based on treatment exposure.

Measures to ensure strict adherence to confidentiality provisions were in place, including removal of potential patient identifiers and conversion of dates in the datasets available for analyses. Applying established data science methods,[31] a data dictionary for each data source was created, based on case report forms, primary publications, and guidance from study statisticians. Then a common data model with a data dictionary across all sources was normalized, standardized, and harmonized into one master, annotated database. Normalization refers to the same measurement units, especially for laboratory values in prognostic scores. Standardization typically refers to uniformed definition in patient characteristics, whereas harmonization ensures that all information contained in each variable is consistent across trials and registries, clinical subgroups, and ages (including stage, prognostic scores, and response criteria).

The resulting HoLISTIC dataset is the most ambitious lymphoma database created through sharing and merging data multiple data sources and has already led important new insights, most important highly predictive prognostic model for outcomes after first-line treatment of HL. [19]

HoLISTIC utilized the best available data along the continuum in the analyses, where possible. For example, due to central adjudication and complete data, short-term risk factors and outcomes have utilized data from seminal clinical trials in the modern era. Data from registries have been used to externally validated clinical decision models. To characterize longer-term outcomes often not collected on clinical trials, individual patient-level meta-analysis has been performed, drawn from population-level data sources. The database has been built to allow for updating, based on emerging knowledge. This includes longer follow-up, inclusion of additional data elements of interest, and inclusion of additional trials/registries with unique and valuable information. Ongoing work is linking treatment exposure to late effects estimates, utilizing simulation modeling to address data scarcity and uncertainty.