



Biorepositories For Global Rare Disease Research: A Narrative Review

Maria Rosa Pellico^{1,2} · Jessica Day^{3,4,5} · Meera Shah⁶ · Belina Y. Yi⁷ · Lesley Ann Saketkoo^{8,9,10} · Christian Lood¹¹ · Latika Gupta^{12,13,14,15}

Accepted: 30 April 2025
© The Author(s) 2025

Abstract

Purpose of this Review Rare diseases, although individually infrequent, collectively impact a substantial number of people. Collaborative translational research using biospecimens is essential for advancing our understanding of the diverse characteristics and pathophysiology of rare diseases. Biobanks play a pivotal role in this endeavor by collecting, processing, transporting, and storing biospecimens, thereby serving as invaluable resources for medical research. In this review, we explore currently available biobanks, with a specific focus on those dedicated to rare rheumatic diseases. We also examine accessible best practice guidelines for establishing and maintaining high-quality biobanks, discuss the limitations and propose future directions for enhancing biobanking efforts in rare disease research.

Recent Findings Advances in molecular and genomic technologies have expanded the role of biobanks, enhancing biomarker discovery and precision medicine. However, despite growth in biobanking capabilities, key challenges persist concerning ethics, interoperability, and biospecimen exchange, prompting active responses by various regulatory and governing bodies.

Summary Biobanking has transformed rare disease research. Strengthening national and international collaborations is essential for driving progress in this field and accelerating the development of novel therapeutic and precision medicine approaches.

Keywords Biobank · Biospecimen · Rare disease · Connective tissue disease · Myopathies · Myositis

Latika Gupta and Christian Lood are co-seniors.

✉ Latika Gupta
drlatikagupta@gmail.com

¹ Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy

² Division of Rheumatology, ASST Gaetani Pini-CTO, Milan, Italy

³ Department of Rheumatology, Royal Melbourne Hospital, Parkville, Australia

⁴ Inflammation Division, Walter and Eliza Hall Institute of Medical Research, Parkville, Australia

⁵ Department of Medical Biology, The University of Melbourne, Parkville, Australia

⁶ Department of Rheumatology, Indraprastha Apollo Hospital, New Delhi, India

⁷ Division of Pediatric Allergy, Immunology, and Rheumatology, Johns Hopkins University, Baltimore, MD, USA

⁸ New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center, New Orleans, LA, USA

⁹ Myositis International Health & Research Collaborative Alliance (MIHRA), Baltimore, USA

¹⁰ LSU and Tulane University Schools of Medicine, New Orleans, LA, USA

¹¹ Division of Rheumatology, University of Washington, Seattle, WA, USA

¹² Department of Rheumatology, Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK

¹³ School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, UK

¹⁴ Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Centre for Musculoskeletal Research, The University of Manchester, Manchester, UK

¹⁵ Francis Crick Institute, London, UK

Introduction

Rare diseases (RD) are defined as those affecting fewer than 1 in 2000 individuals in the community as per The European Commission or fewer than 200,000 people in the United States according to the National Organization for Rare Disorders (NORD) [1]. Several systemic connective tissue diseases in rheumatology, including idiopathic inflammatory myopathies (IIM), systemic lupus erythematosus (SLE), and systemic sclerosis meet these criteria [2]. Investigating these conditions is particularly challenging due to their low prevalence and clinical heterogeneity.

A major obstacle in studying rare diseases is assembling sufficiently large and diverse patient cohorts, given the scarcity of specialized research centers. While meaningful scientific discoveries have been made within single institutions, multi-center collaboration enhances the robustness and generalizability of findings. One of the most significant hurdles in translational research on rare diseases is limited access to infrastructure supporting biospecimen collection, storage and analysis.

Biobanks play a crucial role in overcoming these challenges. By providing centralized, facilitating access across sites, biobanks enable broader collaboration and improve statistical power. This broader access can lead to a more comprehensive understanding of rare systemic rheumatic diseases which increases the likelihood of identifying novel biomarkers that are accurate and representative of the wider patient populations, aiding in early diagnosis and treatment [3].

However, despite their importance, biobanks dedicated to rare systemic rheumatic diseases remain limited, with most concentrated in developed countries and major urban centers. This geographical clustering restricts the diversity of patient populations, potentially limiting the representativeness of disease characteristics in global research. Moreover, there is a lack of evidence-based, standardized, and widely accessible protocols for handling biospecimens in rare systemic rheumatic disease research, further complicating progress in the field. These challenges are particularly pronounced for early-career researchers, especially those at smaller and emerging and lower-income countries, who may face substantial barriers in accessing the infrastructure and resources needed for biobanking and biospecimen research.

In this narrative review, we examine the current landscape of collaborative biobanks dedicated to rare systemic rheumatic diseases, addressing unmet needs, international standards and unique challenges in the field.

Methods

This review employed a structured, narrative approach to the literature, incorporating elements of systematic review methodology without adhering to the formal protocols.

Literature Search Strategy

A comprehensive literature search was conducted to identify relevant studies and publications on biobanks in rare rheumatic diseases, covering articles from January 2011 to April 2024. The search strategy was designed to capture a wide range of research articles, reviews, guidelines, and reports, with a focus on the establishment, operation, and challenges of biobanks, particularly in relation to rare rheumatic diseases. The search was performed across multiple electronic databases, including PubMed, Scopus, Web of Science, and Google Scholar.

The search terms used included combinations of the following keywords: “biobank”, “biorepository”, “bio-specimen”, “rheumatic diseases”, “rare diseases”, “rheumatology”, “standard operating procedures”, “ethics”, “multi-center”, “collection methods”, “sample storage”, “data sharing”, and “reproducibility”. Boolean operators (AND, OR) were employed to refine the search results [4].

Inclusion and Exclusion Criteria

Articles were included if they focused on biobanks related to rheumatic diseases, with particular emphasis on rare or orphan diseases within the rheumatology field. Studies discussing the establishment, management, or challenges of biobanks, including ethical, operational, and scientific aspects, were considered relevant. Publications providing information on biospecimen collection, processing, storage, and data sharing practices, especially those highlighting multi-center biobank studies, were also included.

Exclusion criteria included articles that did not specifically address biobanks in the context of rheumatic diseases, studies that focused on common rheumatic diseases without addressing issues unique to rare diseases and publications not available in full-text or those published in languages other than English.

Data Extraction and Synthesis

The identified articles were screened by title and abstract to assess their relevance to the objectives of this review. Full-text versions of potentially relevant studies were retrieved and reviewed in detail. Data extracted included the type of biobank, sample types, geographic and institutional coverage, ethical considerations, and challenges related to the biobanking of rare rheumatic diseases. Emphasis was placed on identifying common themes, challenges, and best practices that could inform the development of biobanks in resource-limited settings.

The results of the selected studies were synthesized narratively, highlighting key findings, gaps in knowledge, and areas for future research. A thematic approach was used to

organize the data, focusing on the ethical, operational, and scientific challenges faced by biobanks in the context of rare rheumatic diseases. Additionally, the review examined examples of successful biobanks and their contributions to advancing research in this field.

Defining Biobanks and Their Role

Biobanks collect, store, and process biospecimens to facilitate research on diverse diseases. These repositories integrate biological samples with clinical data, enabling large-scale studies and providing insights into population health and disease mechanisms. Biobanks can be broadly categorized into population-based and disease-oriented resources, each serving distinct research needs [5].

Population-based biobanks, such as the UK Biobank, tend to operate on a large scale, collecting biological samples from volunteers, often without specific inclusion or exclusion criteria. These biobanks enable large-scale genetic studies and development of risk assessment models [6]. Another notable example is the Swedish Malmö Diet Cancer Study, which has followed 30,000 participants for over 30 years, enabling important discoveries regarding risk factors for a variety of diseases including cancer, cardiovascular disease and rheumatic conditions like giant cell arteritis and rheumatoid arthritis [7]. Similarly, the China Kadoorie Biobank, with over 500,000 participants, provides regionally anchored, longitudinal data that has supported investigations into autoimmune disease risk factors in East Asian population [8]. The primary aim of population-based biobanks is to investigate the interplay between genetic susceptibility and environmental exposures in disease development, combining biospecimen research with comprehensive clinical and demographic data.

Disease-oriented biobanks, in contrast, focus on specific conditions and enable investigations into pathogenesis, biomarker discovery, diagnostics, long-term patient monitoring and identification of potential therapeutic strategies. These biobanks provide high-quality biospecimens linked to detailed clinical information. By integrating a wide array of data from numerous biological samples, research groups can develop large-scale research projects and gather information on the studied population. A key example is the Systemic Lupus Erythematosus International Collaborating Clinics (SLICC), an international repository collecting DNA, plasma, and serum samples on SLE [9].

Operational Challenges in Biobanking

While biobanks are essential for medical research, their operation is often hindered by infrastructure limitations, governance issues, and lack of standardization. Effective

and efficient biobanking requires rigorous biospecimen collection, processing, storage, distribution and clinical annotation, but inconsistencies in protocols across institutions can compromise sample quality and data integrity. Additionally, many biobanks operate in isolation, leading to fragmentation, reduplication of efforts and limited data sharing.

To better understand and address these challenges, Shickle et al. classified biobank networks into six operational categories—‘storage’, ‘bring-and-share’, ‘catalogue’, ‘partnership’, ‘contribution’, and ‘expertise’—based on their operational and governance structures rather than their research focus. For example, storage networks primarily offer centralized infrastructure, developed to provide cost-effective shared storage solutions, while expertise networks focus on knowledge-sharing and standardization rather than sample exchange. This non-mutually exclusive, governance- and operational-focused classification provides a framework for improving biobanking networking and identifying solutions from an operational perspective [10, 11].

Cross-departmental collaboration, such as partnerships with pathology or clinical biochemistry laboratories, has been shown to improve sample collection accuracy and overall biobank quality, as exemplified by the Danish Rheumatologic Biobanks [12]. However, variability in sample processing and the absence of harmonized protocols across different biobanks remain major concerns. Differences in collection, handling and storage can introduce confounding factors, potentially compromising the validity of findings due to uncontrolled preanalytical variables. Additionally, accurate and standardized clinical data is also vital for linking molecular profiles to disease outcomes [13, 14].

To mitigate these challenges, international efforts have focused on establishing standards for sample collection, storage, and distribution. The International Society for Biological and Environmental Repositories (ISBER) provides comprehensive guidance through its Best Practices for Repositories, emphasizing risk and quality management, ethical and social considerations, and uniform specimen handling [15].

Similarly, the International Standards Organization (ISO) has developed ISO 20387:2018, an international standard specifically for biobanking. This certification sets globally recognized quality and competency requirements for biorepositories, covering sample collection, preparation, data management, storage, and distribution [16]. Institutions adhering to this standard demonstrate high-quality biobanking practices, fostering international collaboration and improving research reproducibility. These efforts by ISBER and ISO have been instrumental in synchronizing biobanking practices globally.

The field of oncology also provides a successful model for standardization, with The National Cancer Institute Best Practices for Biospecimen Resources serving as a benchmark for consistent biospecimen handling [17].

Establishing and Maintaining Disease-Specific Biobanks in Rheumatology: Logistical Considerations

Biospecimen Collection and Handling

The collection and storage of biospecimens for biobanking has historically lacked standardized practices, particularly in the pre-analytical steps [18]. Human biological samples encompass a wide variety of materials, such as tissues, organs, DNA or RNA extracts, blood, bodily fluids, cell lines, cell suspensions, and plasma. Figure 1 shows the steps involved in liquid and solid biospecimen collection, processing, and storage. The standardization of operating procedures directly influences the quality and usability of samples in downstream research. Standard operating procedures for sample procurement, processing, and storage should adhere to the best practices recommended by the ISBER [19]. Adherence to these procedures ensures sample integrity for downstream research applications.

Pre-Analytical Errors and Biosample Bias

Pre-analytical variables refer to any and all procedures that occur during sample collection, prior to sample analysis. The pre-analytics phase consists of two phases namely the pre-acquisition phase (the sample is not under the supervision and control of biobank personnel) and the acquisition phase (the sample comes under the supervision and control of biobank personnel), and are major sources of errors in laboratory diagnostics [18]. This includes patient identification, physical sample collection, sample transportation to the testing site and sample preparation. Variability in pre-analytic factors has significant impact on the reproducibility of results, especially in sensitive analyses such as metabolomics, proteomics, and biomarker research for rare rheumatic diseases [20].

Pre-analytical factors crucial for sample integrity include:

1. Tube type: Additives impact analyte composition
2. Clotting time: Affects coagulation and biomolecule release
3. Transport/incubation time: Influences enzymatic activity
4. Transport temperature: Affects protein and metabolite stability
5. Storage conditions: Impact long-term sample preservation
6. Processing time: Critical for metabolite profile accuracy

Biosamples should be labeled pseudonymously and stored in aliquots. For long-term aliquoting, it is recommended

to store samples from an individual in multiple 0.5- or 1.0-mL aliquots. It is essential to minimize resampling or freeze–thaw events, both of which have been shown to impact downstream analysis [20, 21]. Examples of key pre-analytical factors and their impact on laboratory diagnostics across various assays and sample types are provided in Table 1.

The Challenge and Importance of Complex Biospecimens

Another important concept in biobanking is the management of complex biospecimens. These samples are characterized by features like linkage to other samples from the same donor across different locations or time points, offering invaluable insights into longitudinal changes in disease and gene-environment interactions. However, they are challenging to acquire and manage, particularly when combining smaller cohorts. These biospecimens are essential in health research, especially within biobanks established for longitudinal research studies [29]. The integrated registry-biobank approach offers a promising solution, enabling systematic tracking and management of complex biospecimens and associated longitudinal data across different stages and locations [30].

Sample Data Recording:

Accurate and standardised recording of sample-related data is critical for biobank functionality. The ISBER advocates for the implementation of harmonized data elements such as specimen type, fixation/stabilization methods, study type and mass/volume, in addition to various other essential annotations. Such standardization enhances biospecimen traceability and database functionality. Key initiatives such as BRISQ (Biospecimen Reporting for Improved Study Quality), SPREC (Standard PREanalytical Code) and MIABIS (Minimum Information About Biobank Data Sharing) have been devised to ensure standardization of data [31–33].

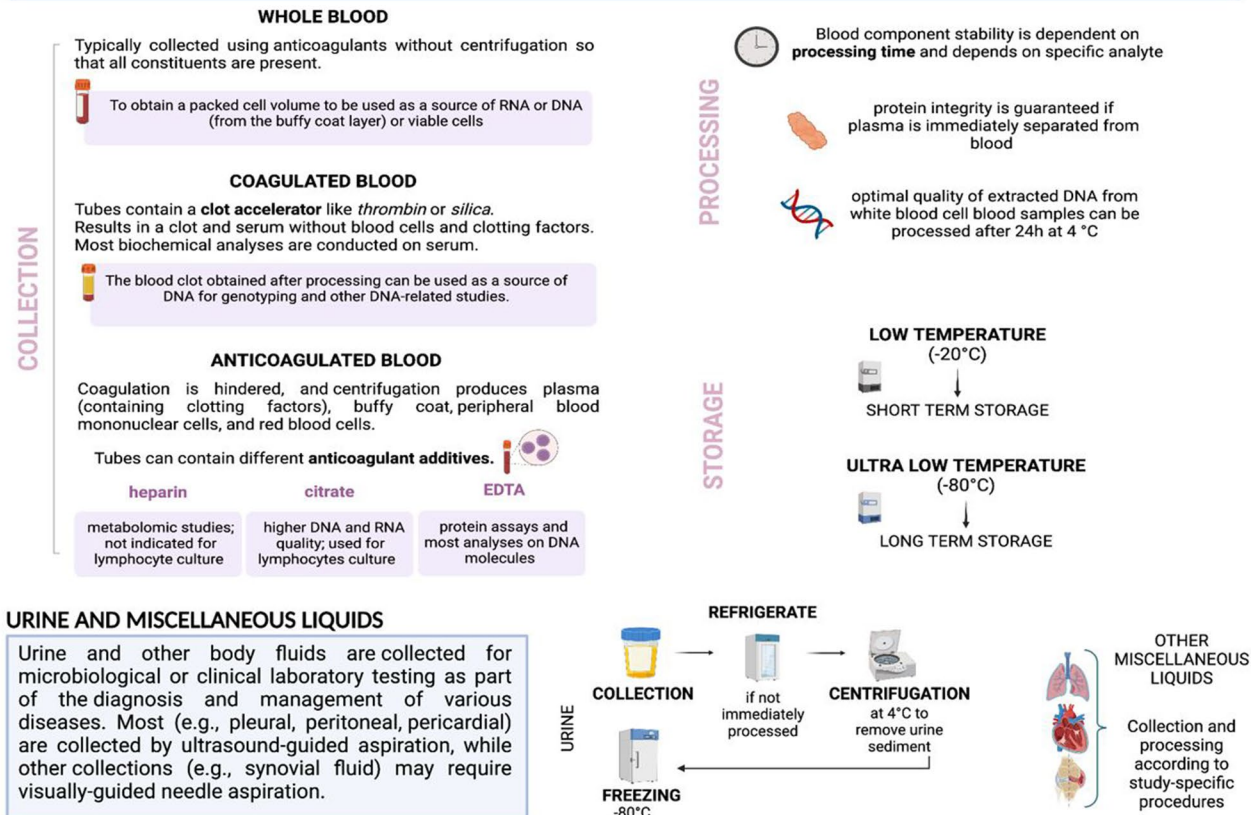
Clinical Data Acquisition

The acquisition of clinical data is a cornerstone of effective biobanking, necessitating meticulous documentation of participant identifiers, demographic parameters, medical histories, family history, diagnoses, therapeutic interventions, and diagnostic outcomes [34]. Standardization of data collection methodologies is imperative to facilitate interoperability and cross-platform data exchange in research endeavours. Comprehensive documentation should encompass an

LIQUID BIOSPECIMEN

BLOOD SAMPLES

The collection of blood should be conducted according to a **fit-for-purpose** procedure that takes into account specific constraints of the collection process (e.g., geographic distances between collection and processing sites, low-resource settings). It should consider the intended downstream use, particularly for cellular, metabolic, and/or genomic analyses, which may require unique collection tubes, processing procedures, and storage conditions that should be determined before blood collection (Matzke et al., 2019).



SOLID BIOSPECIMEN

Strict collaboration with pathology laboratories is essential to control pre-analytical variables such as cold ischemia time and time to fixation, particularly for FFPE (Formalin-Fixed Paraffin-Embedded) tissue specimens.

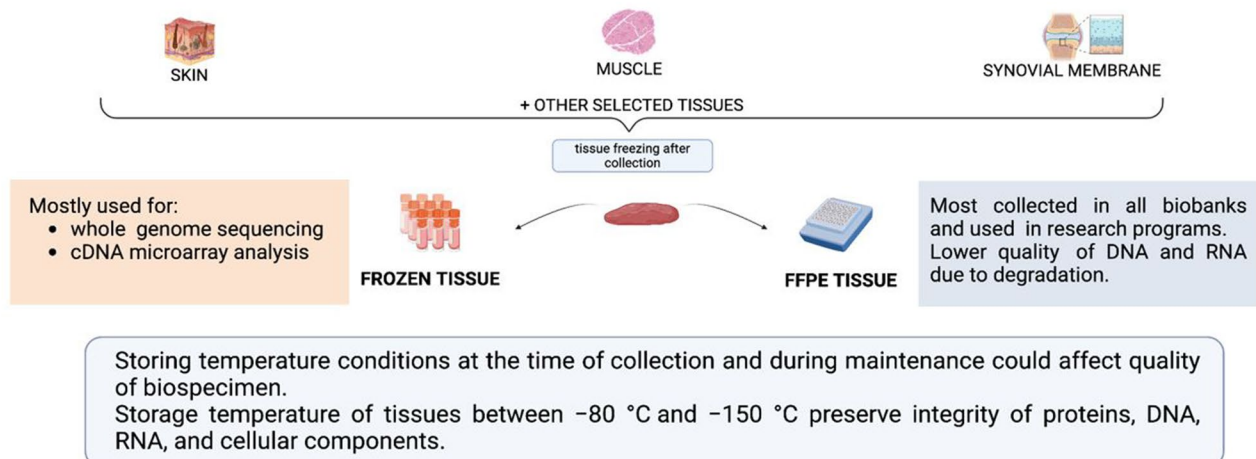


Fig. 1 Overview of Liquid and Solid Biospecimen Collection, Processing, and Storage

Table 1 Pre-analytical factors and their impact on laboratory diagnostics across different assays and sample types

Author Year	Sample Type	Assay Type	n	Pre-analytic factor	Impact	Processing Part Without Impact
Kirk et al. 2008 [22]	Urine	ELISA	13	Time at room temperature, protease inhibitors, pH alteration	VEGF levels were significantly impacted by time at room temperature, with higher values observed between 3–7 h. Protease inhibitors and pH alteration had no significant impact	No significant impact of protease inhibitors or pH alteration on VEGF levels
Timms et al. 2007 [23]	Serum	SELDI-TOF MS	25	Clotting time, transport time, storage	Extended transport/incubation at room temperature led to proteolysis, altering protein profiles significantly	Clotting time had minimal impact when samples were on ice
Webster, MJ. 2006 [24]	Postmortem Brain	RNA and Protein Analysis	450	pH, postmortem interval, agonal state	pH was a better predictor of RNA integrity (RIN) than postmortem interval. Certain diagnostic groups had significantly lower pH values without corresponding lower RIN values, indicating potential metabolic abnormalities	RNA integrity was not significantly affected by PMI within certain ranges
Banks et al. 2005 [25]	Plasma and Serum	SELDI-TOF MS	10	Time delay before processing, tube type, anticoagulant	Time delays before processing significantly altered protein profiles, especially in serum samples. Differences between anticoagulants and tube types also affected results, with specific peaks influenced by platelet activation and clotting	Minimal impact when using citrate as an anticoagulant in plasma samples
Drake et al. 2004 [26]	Serum	MALDI-TOF MS	N/A	Blood collection tube type, tube additives	Different types of blood collection tubes released polymeric components detectable by MALDI-TOF MS, potentially interfering with the analysis of low-molecular-weight serum polypeptides	Plasma tubes showed fewer interfering components compared to serum tubes
Breit et al. 2004 [27]	Bone Marrow	Microarray, RT-PCR	N/A	RNA extraction protocols, time delays	Time delays between bone marrow aspiration and RNA extraction had significant effects on mRNA gene expression profiles, with 18.8% of probe sets showing deregulation	Immediate processing or stabilization preserved mRNA integrity
Srinivasan et al. 2002 [28]	Human Tissues	DNA Microarray, Proteomics	N/A	Fixation and tissue processing	Different fixation methods and tissue processing steps significantly impacted the content and integrity of nucleic acids, affecting downstream molecular analyses	Proper fixation and minimal processing time preserved nucleic acid integrity

accessible data dictionary or codebook, delineating data element nomenclature, definitions, data types, permissible ranges, formatting specifications, and validation protocols.

Standalone registries excel in collecting detailed clinical and epidemiological data, tracking disease progression and elucidating natural disease history, but have a limited capacity to support basic and translational research. Rare disease biobanks, on the other hand, provide the infrastructure for basic research, omics research, but typically lack comprehensive epidemiological and clinical data [30].

The shortcomings of each individual modality can be overcome by the integration of registries with biobanks. This synergy facilitates biomarker identification, gene discovery, omics-level investigations, by with linking clinical data and biological specimens [30]. Importantly, integrated registry-biobank models emerge as cost-effective, synergistic solutions, enhancing both data comprehensiveness and research capabilities, facilitating the translation of basic science into clinical applications [35]. Collaborative multi-stakeholder efforts strengthen treatment protocols, data completeness, and industry engagement, ultimately advancing research in rare diseases.

Additionally, when linking large datasets to biobanks, adherence to the RECORD (REporting of studies Conducted using Observational Routinely-collected health Data) statement will ensure consistency in data collection practices. This may strengthen the reporting and reliability of translational observational research [36].

Informed Consent

Informed consent in biobanking is a complex and evolving issue, balancing the need for comprehensive ethical oversight with practical research needs. A point of contention is the use of broad versus specific consent – specifically whether the provision of broad consent (allowing future, unspecified research) is ethical [37–39].

Proponents of broad consent argue that, when coupled with secure data handling, the right to withdraw participation, and stringent ethical oversight, it respects donor autonomy while enabling critical research [37, 40–42]. Critics, however, contend that broad consent undermines true informed consent since donors cannot anticipate all future uses of their biospecimens [40].

Hartanti et al. developed and piloted a broad informed consent model in Indonesia consisting of an information sheet, comprehension test, agreement sheet and exit survey. When implemented on patients, they found it to be ethically sound, understandable, and acceptable, with recommendations for improving readability and staff coordination [41]. The COVID-19 pandemic further transformed consent practices, emphasizing flexibility and recalibrating research efficacy and participant autonomy [43]. These developments

highlighted the need to balance autonomy with societal benefits and adapt ethical protocols for large-scale studies.

Another form of consent that can be implemented is the dynamic informed consent, which claims to be fully informed and involves continuous communication between the researcher and participants. This ensures participants' decision making, autonomy and real-time consent management, conferring to ethical standards [44].

Recognition of Contribution

Ensuring proper recognition of biobanking resources is essential for incentivizing data and sample sharing. However, heterogeneous practices and lack of guidelines on how bioresources should be acknowledged in scientific publications have led to inconsistencies in reporting. To address this, the Bioresource Research Impact Factor/Framework (BRIF) was introduced over a decade ago as an initiative to standardize the recognition of bioresources in research, and is currently under further development [45]. The CoBRA (Citing Bioresources in Research Articles) guideline was developed to standardize the citation of bioresources in scientific articles, recommending that each bioresource used in a study be cited in the methods section and listed as an individual reference [46]. Tools like CoBRA and The Open Journal of Bioresources enhance the traceability and visibility of bioresource use, encouraging transparency, recognition, and sharing within the research community [47].

Interoperability and Biospecimen Exchange

The premise of the Open Science movement is that sharing data, methods and knowledge accelerates progress, a principle especially valued in the rare disease communities [48]. The widely adopted FAIR principles (findable, accessible, interoperable, and reusable), initially established for data sharing in health research, should also be systematically applied to biospecimens. However, the limited exchange leads to inefficient use of research funds, reduced productivity, and hinders reproducibility efforts [49].

To address these limitations, novel data governance models have emerged to manage and share biobank data in more collaborative and responsible ways. These models seek to balance data accessibility, consistent with FAIR principles and open science, with privacy and ethical safeguards, often by distributing governance across networks rather than a single repository [50–52]. One notable example is the use of federated data networks, in which data remain at multiple distributed sites but can be queried or analyzed across those sites under a common governance framework. This approach enables researchers to gain insights from combined

data without centralizing sensitive information. For example, the National Patient-Centered Clinical Research Network (PCORnet) in the United States and the international Observational Health Data Sciences and Informatics (OHDSI) network illustrate federated governance: each participating institution keeps custody of its patient data in a standardized format, and analyses are run locally to aggregate results [53, 54].

Implementing FAIR-compliant data stewardship is challenging, necessitating expertise in domain knowledge, information technology systems, data access policies, machine-readable formats, and communication software. To address these barriers, the rare disease community in Europe has embraced the "Bring Your Own Data" (BYOD) workshop concept to help rare disease researchers learn how to make their data FAIR-compliant. Additionally, initiatives like "RDs GO FAIR" and activities of ELIXIR (the European Infrastructure for life science data), BBMRI (Better Biology Makes Reality Interesting), NIH (National Institute of Health), and EURORDIS (European Organization for Rare Diseases) amongst others are working to establish a FAIR-compliant ecosystem that enables new data analysis possibilities in an Open Science environment [48, 55].

Another significant obstacle for researchers involves locating relevant resources via cataloging services like the European Research Infrastructure on Biobanking and Biomolecular Resources (BBMRI-ERIC) Directory. Such directories, often lack sufficient metadata for researchers to accurately assess resource suitability. The BBMRI-ERIC developed the Negotiator tool, which facilitates direct negotiation of access request among over 600 participating biobanks, streamlining requests across multiple resources while enabling biobanks to manage access decisions independently [56].

Ethical, legal and societal considerations in biobanking

The operation and collaboration of biobanks are often complicated by ethical, legal and societal challenges, which can restrict biospecimen and data sharing across jurisdictions, hinder international research efforts and impact public trust. Ensuring ethical governance and compliance with legal frameworks is therefore essential to maintain public confidence in biobanking initiatives.

One of the primary ethical concerns is the safeguarding donor privacy and rights. Transparent policies, well-defined governance structures and ethical oversight mechanisms, such as data-access committees and continuous engagement with donors and policymakers are critical for mitigating risks of misuse [57]. Building and maintaining public trust is another major challenge, especially in the era of big

data analysis and artificial intelligence. Indeed, a roundtable discussion from the 2021 ISBER annual meeting highlighted the need to improve public trust in biobanking. They found that the awareness regarding biobanking, transparency and communication, as well as community involvement in biobank was particularly limited in Asian countries compared to Australia and United States [58]. Addressing these concerns through education, engagement and transparent policy practices is essential to ensure sustainable participation in biobanking efforts.

Cross-country collaboration presents additional complexities due to differing ethical and legal frameworks, making it difficult to share biological samples and data across jurisdictions. Key barriers include inconsistent consent requirements, differing data protection laws and the absence of globally standardized protocols. The Organization for Economic Co-operation and Development (OECD), an inter-governmental organization with 38 member countries, plays a vital role in shaping the global biobanking landscape by providing guidance on ethical governance, regulatory frameworks and international collaboration [59]. The OECD Guidelines on Human Biobanks and Genetic Research Databases serves as a reference for ensuring transparency, protecting donor rights and fostering trust in collaborative biobank-based research.

Innovative data governance models, such as federated data networks, offer promising solutions for ethical and secure collaboration by allowing data to remain within local institutions while permitting controlled, query-based access across networks. This approach minimizes the need for physical data transfers, thereby reducing risks to privacy and respecting local legal frameworks [50]. However, it also raises ethical considerations around transparency, consent for secondary use, and accountability in distributed systems.

Sustainability in Biobanking

The rapid increase in biospecimen generation has placed growing pressure on biobanks, underscoring the importance of long-term sustainability. Ensuring the viability of biobanks requires financial stability, cost effective operations, patient and public engagement, standardized protocols, and interoperability. Interoperability means biospecimens and their relevant data can be shared between different biorepositories institutes or databases [60]. Regular evaluation of bioresources (specialized research infrastructures that collect, store, and manage biological specimens and their associated data) also plays a significant role in biobanking sustainability. Beyond tracking the number of biospecimens collected, these evaluations assess how effectively the resource has translated to scientific outputs and outcomes [61].

Examples of Biobanks in Rheumatology – Challenges and Achievements

Collaborative research is vital in advancing diagnostics and treatments in rare rheumatic diseases. Several international initiatives have established specialized biobanks, such as the SLICC biobank for SLE, the European Scleroderma Trials and Research Group (EUSTAR) biobank for systemic sclerosis, and the MyoCite biobank for IIM. These initiatives underscore the importance of high-quality biospecimens and standardized data in enhancing medical knowledge and patient outcomes for rare diseases [62–64]. Recent initiatives such as the RA-originated Gut Microbial Biobank (RAGMB) highlight the expanding frontier of biobanking by integrating culture-based microbiome collections, advancing our understanding of microbial contributions to autoimmune pathogenesis [65]. The EUSTAR and SLICC biobanks have highlighted the importance of large-scale data in boosting the statistical power of studies, leading to significant discoveries in disease mechanisms and treatment responses [62, 63]. Latin America's coordinated biologics registries—such as BIOBADAMEX and BIOBADABRASIL—similarly demonstrate how longitudinal data systems can enhance outcome surveillance and inform equitable biobanking strategies in resource-constrained settings [66]. The Childhood Arthritis and Rheumatology Research Alliance (CARRA) has established a biobank specifically for pediatric rheumatic diseases [67]. Impactful biobanks dedicated to rheumatic diseases are presented in Table 2.

Paediatric rheumatic diseases, though rare, necessitate targeted research for advancing understanding and therapeutic development. Biospecimen collection from paediatric cohorts with rheumatic diseases presents unique ethical and logistical challenges, balancing participation rights with child protection mandates. Recent recommendations by Public and Professional Policy Committee of European Society of Human Genetics and European normative and legal framework surrounding paediatric clinical trials emphasize subsidiarity, the paediatric rule, minor protection, and burden minimization [68, 69]. Integrating minors in the consent process is crucial, ensuring their informed participation alongside legal guardian consent [70]. Future directions emphasize multi-center international paediatric biobanks, standardized practices, robust ethical frameworks, and innovative data management tools to facilitate critical research while protecting young participants [69, 70].

Among rare rheumatic diseases, IIM represent a field where international biobanking efforts are still emerging and would benefit from greater coordination and standardization. Over the past decades, numerous biobanks and registries have contributed to significant advancements in

understanding and treating IIM, across both neurologic and rheumatologic disciplines [71–76]. However, many biobanks continue to operate independently, maintaining separate biospecimen collections and adhering to distinct protocols [77, 78]. These fragmented approaches highlight the ongoing need for standardized practices to maximize the impact of research in this area. Recognizing this challenges, the Myositis International Health Research Collaborative Alliance (MIHRA) has been established to foster global cooperation in myositis research.

Opportunities and Future Perspective – the Need for Collaboration in Rare Rheumatic Diseases

The future of biobanking for rare rheumatic diseases depends on several critical advancements. The development and adoption of evidence-based guidelines for cold-chain management and sample processing protocols tailored to rare diseases are urgently needed. The current landscape remains fragmented due to differing national directives and technical methods, leading to significant heterogeneity among biobanks [99, 100].

Integrating standardized pre-analytical protocols into biobank management is also essential, especially for new facilities and high-priority specimens in established repositories. This approach is much needed as biorepositories face growing demands for quality assurance. Implementing SPREC, BRISQ or MIABIS in this context may harmonize the existing systems and align with global efforts to improve quality management [19]. Moving forward, collaboration across biobanking communities will be crucial ideally through virtual pilot projects that use shared online platforms and databases to simulate, and refine preanalytical barcode systems.

These initiatives could facilitate data sharing and remote collaborations, helping diverse biorepositories adopt consistent and interoperable practices without needing physical exchanges or on-site coordination.

Additionally, fostering international collaboration is crucial. Given the low prevalence and significant heterogeneity of rare diseases, large and diverse datasets are essential for meaningful research. Streamlined and efficient governance structures, including harmonized legal frameworks and expedited ethical review processes, are necessary to facilitate this global cooperation [101]. The Single Hub and Access point for pediatric Rheumatology in Europe (SHARE) initiative's development of the first recommendations for pediatric rheumatology collaborative research, including biobanking, exemplifies the progress that can be made through such collaboration [102].

Table 2 Established biobanks for rare rheumatic diseases

Disease	Biobank/Registries	Bio-specimens	Region	Characteristics
SLE	The Systemic Lupus Erythematosus International Collaborating Clinics (SLICC) [9, 79]	DNA, plasma, serum	North America and Europe, although includes a few centers from Asia and Oceania	<ul style="list-style-type: none"> • Includes 1,835 + SLE patients, with 1,378 DNA and 9,600 + serum/plasma samples; over 500 patients followed for 10 + years • Supports research on metabolic syndrome, vitamin D deficiency, and cardiovascular risk in lupus • Enables studies on genetic susceptibility and clinical-serologic correlations in SLE • Registry contributed to SLE classification criteria, damage index, disease flare index • National registry with 800 + lupus patients across 13 hospitals • Supported 50 + research projects and 8 multi-site studies • Key findings: <ul style="list-style-type: none"> - Ethnic differences in disease severity - Vitamin D deficiency linked to worse lupus - Steroid use associated with organ damage - Validated LLDAS as a disease control target • Facilitated development of steroid-sparing therapies • Open for collaborative research access
	Australian Lupus Registry and Biobank [80, 81]	Whole blood, serum, plasma and urine	Australia, although also developed Asia-Pacific Lupus collaboration	
	Lupus BioBank des OberRheins (Lupus BioBank of the upper Rhein “LBBR”) [82, 83]	Unknown	German and French cohort	<ul style="list-style-type: none"> • Established a globally unique cohort of SLE patients, enhancing research into rare autoimmune diseases. • Unified 17 institutions across France, Germany, and Switzerland, fostering transnational cooperation in autoimmune disease research.

Table 2 (continued)

Disease	Biobank/Registries	Bio-specimens	Region	Characteristics
Systemic Sclerosis	The EULAR Scleroderma Trials and Research Group (EUSTAR) [62, 84]	Whole blood, serum, tissues	Europe	<ul style="list-style-type: none"> • Extensive Patient Registry: Has developed one of the largest international SSc registries, encompassing over 25,000 patients across more than 150 centers worldwide. This comprehensive database facilitates large-scale studies on disease progression, treatment outcomes, and patient demographics. • Provides detailed instructions on biospecimen collection
	Collaborative National Quality and Efficiency Registry (CONQUER) [85, 86]	DNA, serum, plasma, whole blood, RNA	U.S	<ul style="list-style-type: none"> • Launched in 2018 as the first nationwide longitudinal scleroderma registry in the U.S • Builds a collaborative network of U.S. centers to ensure diverse and high-quality data • Enrolled 600 + patients from multiple scleroderma centers by 2021. Defined registry design and utility for improving clinical care and research in systemic sclerosis • Linked early GI involvement with higher healthcare use, highlighting the clinical impact of gastrointestinal symptoms in early disease • Backed by industry partners, including Boehringer Ingelheim and Actelion
	Scleroderma Biobank, Canada (Canadian Scleroderma research group) [87, 88]	Serum, DNA, tissue	Canada	<ul style="list-style-type: none"> • 1,750 + patients followed for up to 13 years • Over 1,500 variables, emphasizing early disease • Organ-Specific Studies: Includes lung, GI, skin, renal, and cardiac involvement • Lung Cancer Link: Interstitial lung disease associated with increased lung cancer risk • Environmental Risk: Potential link between organic solvent exposure and systemic sclerosis

Table 2 (continued)

Disease	Biobank/Registries	Bio-specimens	Region	Characteristics
IIM	MyoCite biorepository [64]	Blood (serum, DNA, plasma), and urine	India	<ul style="list-style-type: none"> • Collected data on 350 + patients (adult and juvenile) with longitudinal data and bio samples • Collected clinical outcomes using predefined glossary and core set four standardized outcomes • Published protocol of biosample adapted for resource constrained setting • Published findings on diverse aspects: autoantibodies, biomarkers, myositis complications, relapse patterns, and prescription practices-insights specific to the Indian demographic • Described metabolomic profiles across multiple specimens (serum, urine, muscle) • Multicentre, prospective • Collects detailed clinical data on IIM patients, facilitating large-scale studies and improving understanding of disease patterns and outcomes. • Research Impact: Identified key predictors of treatment response, such as early therapy and autoantibody status • Integrated Biobanking: Collects blood and tissue samples alongside clinical data, supporting biomarker discovery and personalized medicine efforts • Multicenter, retrospective • Pediatric and adult cases are included
	Swedish Myositis Network (SweMyonet) [89]	Muscle biopsy, serum and DNA	Sweden	
	Pan American League of Associations for Rheumatology (PANLAR) Myositis Registry [90, 91]	Serum, Muscle biopsy	America (both North and South America countries)	
	The Juvenile Dermatomyositis National (UK and Ireland) Cohort Biomarker Study and Repository for Idiopathic Inflammatory Myopathies [92]	PBMCs, serum, genomic DNA and biopsy material	United Kingdom and Ireland	<ul style="list-style-type: none"> • Established in 2000, it's one of the largest European pediatric IIM registries, with over 285 children enrolled • Collects longitudinal clinical data and biospecimens to support research and improve care • Enabled identification of biomarkers for disease activity and prognosis • Contributed to standardized assessment tools and treatment protocols

Table 2 (continued)

Disease	Biobank/Registries	Bio-specimens	Region	Characteristics
Rheumatic diseases	United Kingdom Biobank [93–95]	Blood, urine, saliva	United Kingdom	<ul style="list-style-type: none"> • One of the world's largest health databases with data from 500,000 participants • Supports global research on cancer, heart disease, and autoimmune diseases • Offers broad researcher access via a cloud-based platform • Identified metabolomic and genetic profiles that improve prediction of RA risk • Showed that RA patients with multiple comorbidities face higher risks of mortality and cardiovascular events, stressing the need for holistic care
	Danish Rheumatologic Biobank (DRB) [12, 96]	Blood, synovial fluid, tissue and urine	Denmark	<ul style="list-style-type: none"> • National Danish registry established in 2000 for RA treatment monitoring • Captures > 90% of adult RA patients on biologics in Denmark • Dual-purpose system serving as both clinical documentation tool and research registry • Monitors disease activity, medication safety, and treatment outcomes • Informs national RA treatment guidelines and quality standards • Advances personalized medicine through comprehensive longitudinal data
	Childhood Arthritis and Rheumatology Research Alliance (CARRA) [97, 98] ^w	Plasma, serum, cells, DNA, RNA	North America	<ul style="list-style-type: none"> • North American pediatric rheumatology registry founded in 2002 with 10,000 + enrolled children • Collects longitudinal clinical and patient-reported data on juvenile arthritis, lupus, and other pediatric rheumatic diseases • Maintains biorepository with samples from juvenile idiopathic arthritis, SLE, dermatomyositis, scleroderma • Supports diverse research: observational studies, pharmacosurveillance, comparative effectiveness, translational science • Key findings include racial disparities in pediatric lupus outcomes and subtype-specific responses in juvenile lupus nephritis • Provides evidence for personalized treatment approaches in pediatric rheumatology

Public funding from national and transnational entities (e.g., EU, WHO) should prioritize centralized rare disease registries, emphasizing data curation and sharing. Platforms like the Rare Disease Cures Accelerator-Data and Analytics Platform (RDCA-DAP) could offer funding towards data-enabled registries, encouraging standardized biobank linkage [103]. This approach would dismantle proprietary data silos, meet stakeholder needs, and de-risk early-stage trials, thereby incentivizing industry investment in rare disease research [104].

Academic institutions should reconsider promotion and recognition criteria by shifting away from rewarding data ownership and instead recognizing data-sharing practices and collective achievements. This shift would encourage broader participation and promote more efficient resource utilization across the rare disease research field.

Investing in biobanking infrastructure is another critical priority. Developing advanced IT systems for data management, secure storage facilities, and comprehensive training programs for biobank personnel will ensure the long-term sustainability and efficiency of biobanking operations [105]. Innovative technologies, such as permissioned blockchain frameworks, can enhance transparency and traceability in biobanking processes [106]. Additionally, tools like text mining can facilitate easier data extraction from electronic medical records, improving data collection efficiency [107].

While innovative technologies can enhance data extraction efficiency, they may be difficult to be implemented in resource-restricted environments. In these settings, it is crucial to prioritize foundational infrastructure, such as basic cold chain solutions and essential data management systems, to ensure at least minimal compliance with biobanking standards. Focusing on biospecimens that are easier to handle, like DNA and formalin-fixed paraffin embedded tissues, may be more feasible. In contrast, specimens such as RNA, proteins, or fresh tissues, which are used for metabolomics and proteomics, require strict handling and cold chain control, making them impractical where processing protocols are hard to maintain [108].

A move towards integrative, open, and community-driven registries is also necessary to overcome the limitations of existing systems. Such registries should align and harmonize conflicting standards, make data and code publicly available, and engage community members in governance, thereby promoting sustainability and longevity [109].

Tailoring biobank operations to the specific types of samples and assays needed in each context is essential for maintaining high-quality research outcomes. Furthermore, securing funding—whether through grants, international collaborations, or partnerships—can provide the necessary

resources to build this foundational infrastructure and gradually scale up operations.

Engaging stakeholders, including patients and patient organizations, in governance processes will further build trust and ensure that biobanking initiatives reflect diverse perspectives. Moreover, promoting patient recruitment, especially in specific groups such as adolescents, requires creative recruitment and scheduling strategies tailored to the needs of these populations [110].

Educational initiatives aimed at young translational scientists are vital for accelerating advancements in the field. Encouraging researchers to think beyond their institutions and embrace collaborative approaches early in their careers can foster more innovative and inclusive research efforts [111, 112].

Organizations such as NCATS (National Center for Advancing Translational Sciences), RDCRN (Rare Disease Clinical Research Network) and MIHRA have established teams working in this space [113, 114]. Led by an interdisciplinary panel, MIHRA group is developing an evidence-based protocol for biospecimen processing and storage. This incorporates qualitative interviews to identify barriers in the field based on existing evidence of feasibility and logistics, while emphasizing biospecimen validity.

Conclusions

Biobanking has revolutionized rare disease research by enabling breakthrough discoveries and fostering national and international collaborations. As the field progresses towards personalized medicine, biobanks will become increasingly indispensable in advancing diagnostics, therapeutics, and biomarker discovery. Despite their transformative impact, several unmet challenges persist. Ethical, legal and social concerns regarding data privacy patient confidentiality, and equitable access must be continually addressed through robust governance frameworks and ethical oversight. Additionally, public awareness, building trust among the donors, transparency, standardization and harmonization, sustainability and funding and integration with digital health technologies will be essential for ensuring the long-term efficiency and interoperability of biorepositories. These challenges can be met by engaging multi-stakeholders, enhanced public participation, cross-disciplinary collaboration, supportive policymaking and better integration with digital health tools. By strengthening these areas, biobanking will continue to drive innovation and accelerate rare disease research.

Box. Key Initiatives in Biospecimen and Data Standardization

Initiative	Full Name	Main Purpose	Application Area	Relevance
SPREC [32]	Standard PRE-analytical Code	Provides a standardized coding system for pre-analytical variables during specimen collection, processing, and storage	Biobanking	Enhances reproducibility by documenting factors affecting biospecimen integrity; adaptable across clinical and non-clinical settings
BRISQ [31]	Biospecimen Reporting for Improved Study Quality	Offers a tiered reporting system to improve consistency and transparency in human biospecimen research	Reporting Standards	Improves reliability and comparability of study results; developed through multi-organization consensus including NCI and ISBER
MIABIS [33]	Minimum Information About Biobank Data Sharing	Standardizes core metadata required for biobank interoperability and data sharing	Biobank Metadata	Facilitates collaboration by harmonizing biorepository data structures; supports resource discoverability
FAIR [49]	Findable, Accessible, Interoperable, Reusable	Establishes principles to improve data sharing, reuse, and machine-readability in research	Data Stewardship	Widely adopted for health data; enables integration and reuse across systems; supported by initiatives like ELIXIR and GO FAIR

Initiative	Full Name	Main Purpose	Application Area	Relevance
CoBRA [46]	Citation of BioResources in Research Articles	Recommends standardized citation of biorepositories in scientific publications	Reporting/Citation	Enhances traceability, credit, and reuse of biorepositories; encourages transparency and resource visibility

Key References

- Snapes E, Astrin JJ, Bertheussen Krüger N, Grossman GH, Hendrickson E, Miller N, et al. Updating International Society for Biological and Environmental Repositories Best Practices, Fifth Edition: A New Process for Relevance in an Evolving Landscape. *Biopreservation and Biobanking*. 2023 Dec;21(6):537–46.

Addresses the evolving challenges in biobanking and repository management.

- Kirwan JA, Brennan L, Broadhurst D, Fiehn O, Cascante M, Dunn WB, et al. Preanalytical Processing and Biobanking Procedures of Biological Samples for Metabolomics Research: A White Paper, Community Perspective (for “Precision Medicine and Pharmacometabolomics Task Group”-The Metabolomics Society Initiative). *Clin Chem*. 2018 Aug;64(8):1158–82.

Addresses and emphasizes on the importance of standardized preanalytical and biobanking procedures in metabolomics research. The authors emphasize that variations in sample collection, processing, and storage can significantly impact the metabolomic profiles obtained, potentially leading to inconsistent or misleading results.

- Howard HC, Mascalzoni D, Mabile L, Houeland G, Rial-Sebbag E, Cambon-Thomsen A. How to responsibly acknowledge research work in the era of big data and biobanks: ethical aspects of the Bioresource Research Impact Factor (BRIF). *J Community Genet*. 2018 Apr;9(2):169–76.

Addresses an important topic: the challenges of recognizing and crediting the contributions of researchers and institutions that develop and maintain bioresources, such as biobanks.

- Rush A, Byrne JA, Watson PH. Applying Findable, Accessible, Interoperable, and Reusable Principles to Biospecimens and Biobanks. *Biopreserv Biobank*. 2024 Dec;22(6):550–6.

Emphasizes that applying FAIR principles to biospecimens and biobanks can improve data sharing, enhance research productivity, and ensure efficient use of resources in health research.

- Beyer C, Distler JHW, Allanore Y, Aringer M, Avouac J, Czirják L, et al. EUSTAR biobanking: recommendations for the collection, storage and distribution of biospecimens in scleroderma research. *Ann Rheum Dis*. 2011 Jul;70(7):1178–82.

Serves as a foundational resource for researchers involved in Systemic sclerosis studies, providing comprehensive guidelines to standardize biobanking practices and enhance the quality of systemic sclerosis research.

- Rider LG, Dankó K, Miller FW. Myositis registries and biorepositories: powerful tools to advance clinical, epidemiologic and pathogenic research. *Curr Opin Rheumatol*. 2014 Nov;26(6):724–41.

Highlights the significant role of clinical registries and biorepositories in enhancing our understanding of idiopathic inflammatory myopathies. Provides a comprehensive overview of the existing registries and biobanks in idiopathic inflammatory myopathies.

Author Contributions Conceptualization: LG Data curation: MRP, BY Formal analysis: MRP Funding acquisition: NA Investigation: LG, MRP, JD, BY Methodology: LG Project administration: LG, CL Software: NA Resources: NA Supervision: LG, CL Validation: LG Visualization: LG, JD Writing of the original draft: MRP, LG, and MS Review and editing of the manuscript: All authors.

Funding No funding was received.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Disclosures LG and CL co-lead the MIHRA biospecimen core. The views and opinions expressed belong to the authors and do not necessarily reflect those of the organizations with which they are affiliated. **LAS:** Abbvie, Argenx, aTyr Pharmaceuticals, Boehringer Ingelheim, CSL Behring, EMD Serono, Horizon Pharmaceuticals, Kadmon Pharmaceuticals, Kinevant Pharmaceuticals, Mallinckrodt Pharmaceuticals, Pfizer Inc, Priovant Pharmaceuticals, U.S. Department of Defense.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. The Lancet Diabetes Endocrinology null. Spotlight on rare diseases. *Lancet Diabetes Endocrinol*. 2019;7(2):75.
2. Rheumatology TL. Long-awaited action on rare diseases. *The Lancet Rheumatology*. 2022;4(4):e229.
3. Graham CE, Molster C, Baynam GS, Bushby K, Hansson M, Kole A, et al. Current trends in biobanking for rare diseases: a review. *Journal of Biorepository Science for Applied Medicine*. 2014;21(2):49–61.
4. Gasparyan AY, Aivazyan L, Blackmore H, Kitas GD. Writing a narrative biomedical review: considerations for authors, peer reviewers, and editors. *Rheumatol Int*. 2011;31(11):1409–17.
5. Coppola L, Cianflone A, Grimaldi AM, Incoronato M, Bevilacqua P, Messina F, et al. Biobanking in health care: evolution and future directions. *J Transl Med*. 2019;22(17):172.
6. Conroy MC, Lacey B, Bešević J, Omiyale W, Feng Q, Effingham M, et al. UK Biobank: a globally important resource for cancer research. *Br J Cancer*. 2023;128(4):519–27.
7. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility *J Intern Med*. 1993;233(1):45–51.
8. Hee JY, Huang S, Leong KP, Chun L, Zhang YO, Gongye R, et al. Pregnancy loss and the risk of rheumatoid arthritis in Chinese women: findings from the China Kadoorie biobank. *BMC Public Health*. 2022;22(1):1768.
9. Biobank :SLICC [Internet]. [cited 2025 Feb 16]. Available from: <https://sliccgroup.org/research/biobank/>
10. Shickle D, Griffin M, El-Arifi K. Inter- and intra-biobank networks: classification of biobanks. *Pathobiology*. 2010;77(4):181–90.
11. Watson PH. Biobank classification: communicating biorepository diversity. *Biopreserv Biobank*. 2014;12(3):163–4.

12. Kringelbach T, Glintborg B, Johansen J, Høgdall E, Hetland ML. AB1053 The Nationwide Danish Rheumatologic Biobank – Paving The Road To Personalized Treatment of Patients with Inflammatory Rheumatic Diseases. *Ann Rheum Dis*. 2016;75(Suppl 2):1262–1262.
13. Vaught J, Lockhart N. The Evolution of Biobanking Best Practices. *Clin Chim Acta*. 2012;413(19–20):1569–75.
14. Quinlan PR, Groves M, Jordan LB, Stobart H, Purdie CA, Thompson AM. The Informatics Challenges Facing Biobanks: A Perspective from a United Kingdom Biobanking Network. *Biopreserv Biobank*. 2015;13(5):363–70.
15. best practices for repositories collection. storage, retrieval, and distribution of biological materials for research international society for biological and environmental repositories. *Biopreserv Biobank*. 2012;10(2):79–161.
16. Dagher G. Quality matters: International standards for biobanking. *Cell Prolif*. 2022;55(8):e13282.
17. Rao A, Vaught J, Guan P, Weil C, Moore H. Abstract 5947: The NCI Best Practices for Biospecimen Resources : 2016 revised recommendations. *Can Res*. 2017;77:5947–5947.
18. Riegman PHJ, Morente MM, Betsou F, de Blasio P, Geary P. Biobanking for better healthcare. *Mol Oncol*. 2008;2(3):213–22.
19. Snapes E, Astrin JJ, Bertheussen Krüger N, Grossman GH, Hendrickson E, Miller N, et al. Updating International Society for Biological and Environmental Repositories Best Practices, Fifth Edition: A New Process for Relevance in an Evolving Landscape. *Biopreserv Biobank*. 2023;21(6):537–46.
20. Kirwan JA, Brennan L, Broadhurst D, Fiehn O, Cascante M, Dunn WB, et al. Preanalytical Processing and Biobanking Procedures of Biological Samples for Metabolomics Research: A White Paper, Community Perspective (for “Precision Medicine and Pharmacometabolomics Task Group”-The Metabolomics Society Initiative). *Clin Chem*. 2018;64(8):1158–82.
21. Kirwan JA, Kaddurah-Daouk R, Mitchell T, Pischon T, Schmidt MA, Velagapudi V. Biobanking for Metabolomics and Lipidomics in Precision Medicine. *Clin Chem*. 2019;65(7):827–32.
22. Kirk MJ, Hayward RM, Sproull M, Scott T, Smith S, Cooley-Zgela T, et al. Non-patient related variables affecting levels of vascular endothelial growth factor in urine biospecimens. *J Cell Mol Med*. 2008;12(4):1250–5.
23. Timms JF, Arslan-Low E, Gentry-Maharaj A, Luo Z, T’Jampens D, Podust VN, et al. Preanalytic influence of sample handling on SELDI-TOF serum protein profiles. *Clin Chem*. 2007;53(4):645–56.
24. Webster MJ. Tissue preparation and banking. In: Hemby SE, Bahn S, editors. *Progress in Brain Research* [Internet]. Elsevier; 2006 [cited 2025 Feb 17]. p. 3–14. (Functional Genomics and Proteomics in the Clinical Neurosciences; vol. 158). Available from: <https://www.sciencedirect.com/science/article/pii/S007961230658001X>
25. Banks RE, Stanley AJ, Cairns DA, Barrett JH, Clarke P, Thompson D, et al. Influences of Blood Sample Processing on Low-Molecular-Weight Proteome Identified by Surface-Enhanced Laser Desorption/Ionization Mass Spectrometry. *Clin Chem*. 2005;51(9):1637–49.
26. Drake SK, Bowen RAR, Remaley AT, Hortin GL. Potential Interferences from Blood Collection Tubes in Mass Spectrometric Analyses of Serum Polypeptides. *Clin Chem*. 2004;50(12):2398–401.
27. Breit S, Nees M, Schaefer U, Pfoersich M, Hagemeyer C, Muckenthaler M, et al. Impact of pre-analytical handling on bone marrow mRNA gene expression. *Br J Haematol*. 2004;126(2):231–43.
28. Srinivasan M, Sedmak D, Jewell S. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol*. 2002;161(6):1961–71.
29. Wotton L, Gali B, Carvalho K, Tarling T, Matzke L, Watson PH. Analysis of Trends in Biospecimen Complexity in Cancer Research Over Two Decades. *Biopreserv Biobank*. 2022;20(2):195–200.
30. Garcia M, Downs J, Russell A, Wang W. Impact of biobanks on research outcomes in rare diseases: a systematic review. *Orphanet J Rare Dis*. 2018;13(1):202.
31. Moore HM, Kelly AB, Jewell SD, McShane LM, Clark DP, Greenspan R, et al. Biospecimen reporting for improved study quality (BRISQ). *Cancer Cytopathol*. 2011;119(2):92–101.
32. Betsou F, Lehmann S, Ashton G, Barnes M, Benson EE, Coppola D, et al. Standard preanalytical coding for biospecimens: defining the sample PREanalytical code. *Cancer Epidemiol Biomarkers Prev*. 2010;19(4):1004–11.
33. Norlin L, Fransson MN, Eriksson M, Merino-Martinez R, Anderberg M, Kurtovic S, et al. A Minimum Data Set for Sharing Biobank Samples, Information, and Data: MIABIS. *Biopreserv Biobank*. 2012;10(4):343–8.
34. Annaratone L, De Palma G, Bonizzi G, Sapino A, Botti G, Berriño E, et al. Basic principles of biobanking: from biological samples to precision medicine for patients. *Virchows Arch*. 2021;479(2):233–46.
35. Rush A, NSW Biobank Study Group, Catchpoole DR, Watson PH, Byrne JA. An Approach to Evaluate the Costs and Outputs of Academic Biobanks. *Biopreserv Biobank*. 2024;22(5):463–74.
36. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885.
37. Secko DM, Preto N, Niemeyer S, Burgess MM. Informed consent in biobank research: a deliberative approach to the debate. *Soc Sci Med*. 2009;68(4):781–9.
38. Boulton M, Parker M. Informed consent in a changing environment. *Soc Sci Med*. 2007;65(11):2187–98.
39. Andreas Brekke O, Sirnes T. Population Biobanks: The Ethical Gravity of Informed Consent. *BioSocieties*. 2006;1(4):385–98.
40. Hansson MG, Dillner J, Bartram CR, Carlson JA, Helgesson G. Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol*. 2006;7(3):266–9.
41. Hartanti W, Wahdi AE, Prasetyawati T, Izhati QA, Fachiroh J. Developing informed consent for academic hospital-based biobank modeling: an experience from Indonesia. *Biopreserv Biobank*. 2024.
42. Coebergh JWW, van Veen EB, Vandenbroucke JP, van Diest P, Oosterhuis W. One-time general consent for research on biological samples: opt out system for patients is optimal and endorsed in many countries. *BMJ*. 2006;332(7542):665.
43. Phillips H, Deshpandey M, Staveski S. Consenting in the time of the COVID-19 pandemic. *J Pediatr Nurs*. 2024;74:129–31.
44. Dankar FK, Gergely M, Malin B, Badji R, Dankar SK, Shuaib K. Dynamic-informed consent: A potential solution for ethical dilemmas in population sequencing initiatives. *Comput Struct Biotechnol J*. 2020;2(18):913–21.
45. Cambon-Thomsen A, Thorisson GA, Mabile L, BRIF Workshop Group. The role of a bioresource research impact factor as an incentive to share human bioresources. *Nat Genet*. 2011;43(6):503–4.
46. Napolitani F, Calzolari A, Cambon-Thomsen A, Mabile L, Rossi AM, De Castro P, et al. Biobankers: Treat the Poison of Invisibility with CoBRA, a Systematic Way of Citing Bioresources in Journal Articles. *Biopreserv Biobank*. 2016;14(4):350–2.
47. Howard HC, Mascalzoni D, Mabile L, Houeland G, Rial-Sebbag E, Cambon-Thomsen A. How to responsibly acknowledge research work in the era of big data and biobanks: ethical aspects of the Bioresource Research Impact Factor (BRIF). *J Community Genet*. 2018;9(2):169–76.

48. Rubinstein YR, Robinson PN, Gahl WA, Avillach P, Baynam G, Cederroth H, et al. The case for open science: rare diseases. *JAMIA Open*. 2020;3(3):472–86.
49. Rush A, Byrne JA, Watson PH. Applying Findable, Accessible, Interoperable, and Reusable Principles to Biospecimens and Biobanks. *Biopreserv Biobank*. 2024;22(6):550–6.
50. Shi J, Guo Y, He N, Xia W, Liu H, Li H. Data Governance and Distribution of Biobank: A Case from a Chinese Cancer Hospital. *Biopreservation and Biobanking*. 2024. <https://doi.org/10.1089/bio.2024.0081>
51. Holub P, Kohlmayer F, Prasser F, Mayrhofer MT, Schlünder I, Martin GM, et al. Enhancing Reuse of Data and Biological Material in Medical Research: From FAIR to FAIR-Health. *Biopreserv Biobank*. 2018;16(2):97–105.
52. Boeckhout M, Zielhuis GA, Bredenoord AL. The FAIR guiding principles for data stewardship: fair enough? *Eur J Hum Genet*. 2018;26(7):931–6.
53. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc*. 2014;21(4):578–82.
54. Waitman LR, Bailey LC, Becich MJ, Chung-Bridges K, Dusetzina SB, Espino JU, et al. Avenues for Strengthening PCORnet's Capacity to Advance Patient-Centered Economic Outcomes in Patient-Centered Outcomes Research (PCOR). *Med Care*. 2023;61(12 Suppl 2):S153–60.
55. GO FAIR [Internet]. [cited 2025 Feb 17]. Rare Diseases. Available from: <https://www.go-fair.org/implementation-networks/overview/rare-diseases/>
56. Reihls R, Proynova R, Maqsood S, Ataian M, Lablans M, Quinlan PR, et al. BBMRI-ERIC Negotiator: Implementing Efficient Access to Biobanks. *Biopreserv Biobank*. 2021;19(5):414–21.
57. Samuel G, Lucassen A. Access to Biobanks: Responsibilities Within a Research Ecosystem. *Biopreserv Biobank*. 2023;21(3):275–81.
58. Yadav BK, Ng W, Vu H, Fachiroh J, Tsuruyama T, Zhou L, et al. Improving Public Trust in Biobanking: Roundtable Discussions from the 2021 ISBER Annual Meeting. *Biopreserv Biobank*. 2023;21(3):308–12.
59. OECD Guidelines on Human Biobanks and Genetic Research Databases [Internet]. 2009. Available from: <http://www.rettdatabasenetwork.org/Guidelines%20databases.pdf>
60. Abdaljalalel M, Singer EJ, Yong WH. Sustainability in biobanking. *Methods Mol Biol*. 1897;2019:1–6.
61. Bledsoe MJ, Watson PH, Grizzle WE. Metrics for the Evaluation of Bioresources: Tell Us the Whole Story. *Biopreserv Biobank*. 2023;21(3):223–4.
62. Beyer C, Distler JHW, Allanore Y, Aringer M, Avouac J, Czirják L, et al. EUSTAR biobanking: recommendations for the collection, storage and distribution of biospecimens in scleroderma research. *Ann Rheum Dis*. 2011;70(7):1178–82.
63. Isenberg DA, Ramsey-Goldman R, Gladman D, Hanly JG. The Systemic Lupus International Collaborating Clinics (SLICC) group - it was 20 years ago today. *Lupus*. 2011;20(13):1426–32.
64. Naveen R, Anuja AK, Rai MK, Agarwal V, Gupta L. Development of the Myocyte Biobank: Cost-Efficient Model of Public Sector Investigator-Driven Biobank for Idiopathic Inflammatory Myositis. *Indian J Rheumatol*. 2020;15(Suppl 2):S194.
65. Huang HJ, Liu C, Sun XW, Wei RQ, Liu LW, Chen HY, et al. The rheumatoid arthritis gut microbial biobank reveals core microbial species that associate and effect on host inflammation and autoimmune responses. *iMeta*. 2024;3(5):e242.
66. de la Vega M, da Silveira de Carvalho HM, Ventura Ríos L, Goycochea Robles MV, Casado GC. The importance of rheumatology biologic registries in Latin America. *Rheumatol Int*. 2013;33(4):827–35.
67. Beukelman T, Kimura Y, Ilowite NT, Mieszkalski K, Natter MD, Burrell G, et al. The new Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry: design, rationale, and characteristics of patients enrolled in the first 12 months. *Pediatr Rheumatol Online J*. 2017;17(15):30.
68. Hens K, Van El CE, Borry P, Cambon-Thomsen A, Cornel MC, Forzano F, et al. Developing a policy for paediatric biobanks: principles for good practice. *Eur J Hum Genet*. 2013;21(1):2–7.
69. Pinxten W, Dierickx K, Nys H. Ethical principles and legal requirements for pediatric research in the EU: an analysis of the European normative and legal framework surrounding pediatric clinical trials. *Eur J Pediatr*. 2009;168(10):1225–34.
70. Giesbertz NA, Bredenoord AL, van Delden JJ. Consent procedures in pediatric biobanks. *Eur J Hum Genet*. 2015;23(9):1129–34.
71. Rider LG, Dankó K, Miller FW. Myositis registries and biorepositories: powerful tools to advance clinical, epidemiologic and pathogenic research. *Curr Opin Rheumatol*. 2014;26(6):724–41.
72. Chinoy H, Adimulam S, Marriage F, New P, Vincze M, Zilahi E, et al. Interaction of HLA-DRB1*03 and smoking for the development of anti-Jo-1 antibodies in adult idiopathic inflammatory myopathies: a European-wide case study. *Ann Rheum Dis*. 2012;71(6):961–5.
73. Lundberg I, Svensson J. Registries in idiopathic inflammatory myopathies. *Curr Opin Rheumatol*. 2013;25(6):729–34.
74. Labirua-Iturburu A, Selva-O'Callaghan A, Vincze M, Dankó K, Vencovsky J, Fisher B, et al. Anti-PL-7 (anti-threonyl-tRNA synthetase) antisynthetase syndrome: clinical manifestations in a series of patients from a European multicenter study (EUMY-ONET) and review of the literature. *Medicine (Baltimore)*. 2012;91(4):206–11.
75. Miller FW, Cooper RG, Vencovský J, Rider LG, Danko K, Wedderburn LR, et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis Rheum*. 2013;65(12):3239–47.
76. Reza M, Cox D, Phillips L, Johnson D, Manoharan V, Grieves M, et al. MRC Centre Neuromuscular Biobank (Newcastle and London): Supporting and facilitating rare and neuromuscular disease research worldwide. *Neuromuscul Disord*. 2017;27(11):1054–64.
77. Huber AM, Robinson AB, Reed AM, Abramson L, Bout-Tabaku S, Carrasco R, et al. Consensus treatments for moderate juvenile dermatomyositis: beyond the first two months. Results of the second Childhood Arthritis and Rheumatology Research Alliance consensus conference. *Arthritis Care Res (Hoboken)*. 2012;64(4):546–53.
78. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM, et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)*. 2013;92(1):25–41.
79. Hanly JG, O'Keefe AG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. *Rheumatology (Oxford)*. 2016;55(2):252–62.
80. O'Neill S, Morand EF, Hoi A. The Australian Lupus Registry and Biobank: a timely initiative. *Med J Aust [Internet]*. 2017 Mar 20 [cited 2025 Feb 17];206(5). Available from: <https://www.mja.com.au/journal/2017/206/5/australian-lupus-registry-and-biobank-timely-initiative>
81. lupus-registry [Internet]. [cited 2025 Apr 21]. Australian Lupus Registry. Available from: <https://www.lupusregistry.com>

82. Arnaud L, Gavand PE, Voll R, Schwarting A, Maurier F, Blaison G, et al. Predictors of fatigue and severe fatigue in a large international cohort of patients with systemic lupus erythematosus and a systematic review of the literature. *Rheumatology (Oxford)*. 2019;58(6):987–96.
83. Interreg [Internet]. [cited 2025 Apr 21]. LBBR: Lupus BioBank des OberRheins. Available from: <https://www.interreg-oberrhein.eu/projet/lbbr-lupus-biobank-des-oberrheins/>
84. EUSTAR – European Scleroderma Trials and Research Group [Internet]. [cited 2025 Apr 21]. Available from: <https://eustar.org/>
85. Shanmugam VK, Frech TM, Steen VD, Hummers LK, Shah AA, Bernstein EJ, et al. Collaborative National Quality and Efficacy Registry (CONQUER) for Scleroderma: outcomes from a multicenter US-based systemic sclerosis registry. *Clin Rheumatol*. 2020;39(1):93–102.
86. CONQUER Registry [Internet]. Scleroderma Research Foundation. [cited 2025 Apr 21]. Available from: <https://srfcure.org/research/conquer/>
87. Wu M, Baron M, Pedroza C, Salazar GA, Ying J, Charles J, et al. CCL2 in the Circulation Predicts Long-Term Progression of Interstitial Lung Disease in Patients With Early Systemic Sclerosis: Data From Two Independent Cohorts. *Arthritis Rheumatol*. 2017;69(9):1871–8.
88. CSRG Website [Internet]. [cited 2025 Apr 21]. CSRG Website. Available from: <https://www.canadiansclerodermaresearchgroup.org>
89. Galindo-Feria AS, Lodin K, Horuluoglu B, Sarrafzadeh-Zargar S, Wigren E, Gräslund S, et al. Anti-FHL1 autoantibodies in adult patients with myositis: a longitudinal follow-up analysis. *Rheumatology*. 2024;64(3):keae317.
90. O'Hanlon TP, Carrick DM, Arnett FC, Reveille JD, Carrington M, Gao X, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1 and -DQA1 allelic profiles and motifs define clinicopathologic groups in caucasians. *Medicine (Baltimore)*. 2005;84(6):338–49.
91. Inicio | panlar.org [Internet]. [cited 2025 Apr 21]. Available from: <https://panlar.org/en>
92. Martin N, Krol P, Smith S, Murray K, Pilkington CA, Davidson JE, et al. A national registry for juvenile dermatomyositis and other paediatric idiopathic inflammatory myopathies: 10 years' experience; the Juvenile Dermatomyositis National (UK and Ireland) Cohort Biomarker Study and Repository for Idiopathic Inflammatory Myopathies. *Rheumatology (Oxford)*. 2011;50(1):137–45.
93. Fang XY, Zhang J, Qian TT, Gao P, Wu Q, Fang Q, et al. Metabolomic profiles, polygenic risk scores and risk of rheumatoid arthritis: a population-based cohort study in the UK Biobank. *RMD Open*. 2023;9(4):e003560.
94. Sandoval-Plata G, Nakafero G, Chakravorty M, Morgan K, Abhishek A. Association between serum urate, gout and comorbidities: a case-control study using data from the UK Biobank. *Rheumatology*. 2021;60(7):3243–51.
95. UK Biobank - UK Biobank [Internet]. 2025 [cited 2025 Apr 21]. Available from: <https://www.ukbiobank.ac.uk>
96. RBGB – Regionernes Bio- og Genombank [Internet]. [cited 2025 Apr 21]. Available from: <https://www.rbgb.dk/>
97. Fuhlbrigge RC, Schanberg LE, Kimura Y. CARRA. *Rheumatic Disease Clinics of North America*. 2021;47(4):531–43.
98. Sheridan J. CARRA. 2022 [cited 2025 Feb 16]. CARRA Registry & Biorepository. Available from: <https://carragroup.org/research/carra-registry/>
99. Capocasa M, Anagnostou P, D'Abramo F, Matteucci G, Dominici V, Destro Bisol G, et al. Samples and data accessibility in research biobanks: an explorative survey. *PeerJ*. 2016;4:e1613.
100. Ransohoff DF, Gourlay ML. Sources of bias in specimens for research about molecular markers for cancer. *J Clin Oncol*. 2010;28(4):698–704.
101. Julkowska D, Austin CP, Cutillo CM, Gancberg D, Hager C, Halftermeyer J, et al. The importance of international collaboration for rare diseases research: a European perspective. *Gene Ther*. 2017;24(9):562–71.
102. Kuemmerle-Deschner JB, Hansmann S, Wulffraat NM, Vastert SJ, Hens K, Anton J, et al. Recommendations for collaborative paediatric research including biobanking in Europe: a Single Hub and Access point for paediatric Rheumatology in Europe (SHARE) initiative. *Ann Rheum Dis*. 2018;77(3):319–27.
103. Denton N, Mulberg AE, Molloy M, Charleston S, Fajgenbaum DC, Marsh ED, et al. Sharing is caring: a call for a new era of rare disease research and development. *Orphanet J Rare Dis*. 2022;17(1):389.
104. Moses H, Matheson DHM, Cairns-Smith S, George BP, Palisch C, Dorsey ER. The anatomy of medical research: US and international comparisons. *JAMA*. 2015;313(2):174–89.
105. Zelis HS, de Ridder S, van Gent MDJM, Mom CH, Wisman GBA, Roes EM, et al. The Information Technology (IT) Infrastructure of the Multicenter Archipelago of Ovarian Cancer Research Biobank: A Potential Blueprint for Other Biobanks. *Biopreserv Biobank*. 2024;22(6):568–77.
106. Ortiz-Lizcano MI, Arias-Antunez E, Hernández Bravo Á, Caminero MB, Rojo Guillen T, Nam Cha SH. Increasing the security and traceability of biological samples in biobanks by blockchain technology. *Comput Methods Programs Biomed*. 2023;231:107379.
107. Brunekreef TE, Otten HG, van den Bosch SC, Hoefer IE, van Laar JM, Limper M, et al. Text Mining of Electronic Health Records Can Accurately Identify and Characterize Patients With Systemic Lupus Erythematosus. *ACR Open Rheumatol*. 2021;3(2):65–71.
108. Soo CC, Mukomana F, Hazelhurst S, Ramsay M. Establishing an academic biobank in a resource-challenged environment. *S Afr Med J*. 2017;107(6):486.
109. Hoyt CT, Balk M, Callahan TJ, Domingo-Fernández D, Haendel MA, Hegde HB, et al. Unifying the identification of biomedical entities with the Bioregistry. *Sci Data*. 2022;9(1):714.
110. Sanford SD, Beaumont JL, Snyder MA, Reichel J, Salsman JM. Clinical research participation among adolescent and young adults at an NCI-designated Comprehensive Cancer Center and affiliated pediatric hospital. *Support Care Cancer*. 2017;25(5):1579–86.
111. Karataş M, Azbazar ME, Camkiranlar M, Tercan-Avci S, Atabey N. Biobank Education for Future Physicians: Training Medical Students Through Student Research Association Networks. *Biopreserv Biobank*. 2024;22(3):217–24.
112. Seidler D, Karlíková M, Topolčan O, Snitílá M, Niedermair T, Brochhausen C. Establishing Biobanking in Medical Curricula-The Education Program “Precision Medicine International” (edu-BRoTHER). *Biopreserv Biobank*. 2023;21(2):200–7.
113. Rare Diseases Clinical Research Network - National Organization for Rare Disorders [Internet]. 2022 [cited 2025 Feb 4]. Available from: <https://rarediseases.org/organizations/rare-diseases-clinical-research-network/>
114. Saketkoo LA, Paik JJ, Alexanderson H, Dimachkie MM, Ernste FC, Naddaf E, et al. Collaborative research in myositis-related disorders: MIHRA, a global shared community model. *Clin Exp Rheumatol*. 2024;42(2):207–12.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.