OPINION

Enhancing translational research in paediatric rheumatology through standardization

Rae S. M. Yeung, Salvatore Albani, Brian M. Feldman, Elizabeth Mellins, Berent Prakken and Lucy R. Wedderburn

Abstract | The past decade has seen many successes in translational rheumatology, from dramatic improvements in outcomes brought about by novel biologic therapies, to the discovery of new monogenic inflammatory disorders. Advances in molecular medicine, combined with progress towards precision care, provide an excellent opportunity to accelerate the translation of biological understanding to the bedside. However, although the field of rheumatology is a leader in the standardization of data collection and measures of disease activity, it lags behind in standardization of biological sample collection and assay performance. Uniform approaches are necessary for robust collaborative research, particularly in rare diseases. Standardization is also critical to increase reproducibility between centres, a prerequisite for clinical implementation of translational research. This Perspectives article emphasizes the need for standardization and implementation of best practices, presented in the context of lessons learned from international biorepository networks.

The introduction of therapies that selectively target cytokines or immune cell subsets has led to a dramatic improvement in clinical outcomes for many paediatric and adult patients with rheumatic diseases[1,2]. However, achieving the full potential of these and other interventions will require systematic characterization and classification of patients, incorporating both clinical and biological data. Few therapies are effective in all patients diagnosed as having the same rheumatic disease, and this variable efficacy persists even within groups of patients with the same clinically defined disease subtype[3]. This heterogeneity is also reflected in molecular studies of patients with the same diagnosis, and can ultimately lead to subdivision of the diagnosis into many distinct disease forms, each involving only a few patients. Elucidating the pathophysiology of these disease subtypes and identifying relevant diagnostic or prognostic biomarkers is, therefore, likely to require collaboration. However, the lack of reproducibility of the results from many biomarker studies suggests that procedural inconsistencies add to the effects of genuine biological heterogeneity between patients. The handling of biological samples is one of the most common sources of site-dependent differences observed in biomarker studies. Collaborative research involving multiple centres and groups requires critical procedures to be harmonized to facilitate accurate comparison of data. The use of standard operating procedures (SOPs) for the collection and handling of samples and data is a critical first step in ensuring high-quality translational research. Disseminating such information among researchers requires a flexible and secure data-sharing infrastructure that supports the integration of biological data with precise measurement of clinical variables.

Although the use of standardized measures of disease activity and damage has long been prominent in rheumatology, standardization of both biological sample collection and assays remains suboptimal. Internationally accepted indices of clinical disease activity and damage are widely used in both adult and paediatric rheumatic disease and many of these tools have been adopted as standards of care in clinical assessments[4]. By contrast, biologic measurements have not been incorporated into routine clinical practice, despite evidence that DNA, gene expression, protein expression and cellular immunophenotyping profiles can help to stratify and define homogeneous groups of patients[3,5,6,7]. The results of such molecular profiling also point towards biological pathways that might be involved in susceptibility to disease and patient outcomes in many paediatric and adult rheumatologic diseases[3,6,8].

Additional challenges arise when research is focused on rare diseases. Uniform standards remove some of the barriers to data sharing, but appropriate solutions require large-scale organizational efforts to homogenize methodology aimed at bringing together groups of researchers who then share the collection and management of datasets. Successfully addressing the issues around standardization of clinical and biological data will accelerate the incorporation of translational studies into clinical decision-making (FIG. 1).

Standardization and guidelines

The identification of robust biomarkers of disease can greatly improve diagnosis and treatment decision-making. For example, anti-cyclic citrullinated peptide (CCP) antibodies can define a subgroup of patients with inflammatory arthritis, with important implications for disease course and clinical management[9]. Variability in biological sample collection, however, can affect the results of biological assays. Specimen collection is often perceived as the easy part of a study protocol, yet the validation of biomarkers relies on controlling for errors that could be introduced at each step in the collection process (FIG. 1). To increase efficiency and achieve appropriate sample sizes in translational studies, a concerted effort must be made to establish common principles for use in biobanks. Potential variability due to biospecimen

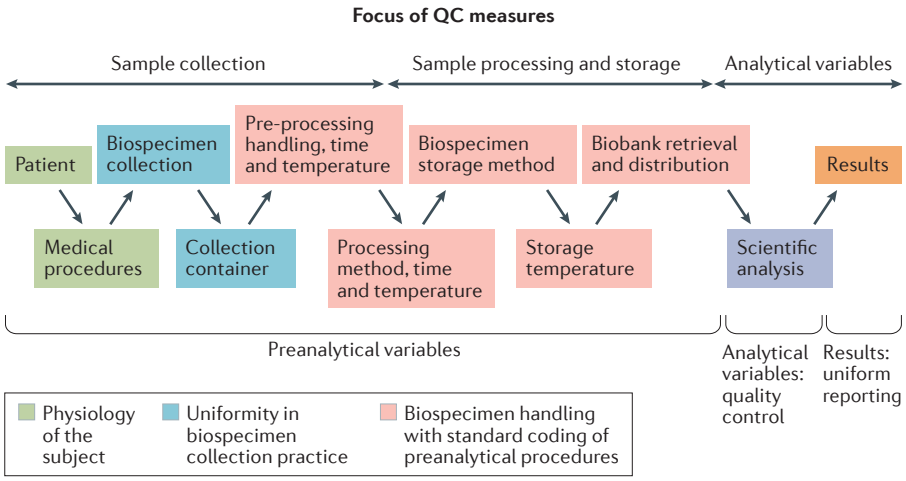

Figure 1 | **Biospecimen handling pipeline.** Establishing standard operating procedures for collection and handling of samples and data is a critical first step in ensuring high-quality translational research. Preanalytical variables can affect various steps of the pipeline, from specimen collection to downstream analysis. Key preanalytical variables include those relating to the patient's individual physiology (green), the uniformity of biospecimen collection (blue) and biospecimen handling procedures (pink). Quality control assays can highlight sample adequacy, confirm optimal sample processing and/or storage and predict the validity of downstream analysis.

collection, processing and handling should be identified and minimized by adhering to uniform SOPs in order to ensure the quality of downstream molecular analyses. Important lessons have been learned from both successes and failures in multicentre translational studies in paediatric rheumatology, highlighting the need for stringent procedures, processing times and quality control measures[7,10].

Although experimental variability can occur at various stages, inconsistencies in preanalytical procedures (those taking place between specimen collection and experimental assay or analysis) are often underappreciated and, therefore, not stringently controlled[11,12]. Even small differences in SOPs within or between research groups could yield results that are difficult to interpret owing to variations at multiple steps in the sample collection and handling process[13]. For instance, the US National Cancer Institute (NCI) Cancer Genome Atlas project was launched on the assumption that researchers could obtain high-quality specimens from dozens of established biobanks with successful track records. Unfortunately, very few of the biological samples initially tested were usable, owing to differences in collection, processing and storage in legacy biorepositories (existing historical biorepositories, usually from single investigators)[13]. The type of blood collection tube and additives used, differences in sample processing (such as analysis protocols, the time from collection to final analysis and temperatures), haemolysis, transport conditions, and storage parameters (such as temperatures and freeze–thaw cycles) can all impair the stability of analytes and cause a deterioration in data quality[14–16].

The challenge when harmonizing SOPs is to avoid reliance on historical protocols used out of habit, and instead to move towards state-of-the-art methodology based on empirical evidence. Valuable lessons have been learned from international efforts to standardize preanalytical biospecimen processing[17]. The International Society for Biological and Environmental Repositories (ISBER) publishes high-quality best practice guidelines for researchers that are updated every few years[18]. ISBER systematically reviews the scientific literature on biospecimen collection and provides researchers with links to relevant publications. Evidence-based recommended protocols are also available for the collection and handling of a comprehensive list of biological samples, including cellular and noncellular fractions from human peripheral blood, although ISBER clearly specifies that the choice of SOP must take into account the analytical end-points to be used, as well as available resources. Similarly, the Early Detection Research Network (EDRN), which is endorsed by the International Agency for Research on Cancer, has proposed SOPs for the collection and processing of noncellular fractions from peripheral blood (including plasma and serum) for use by multi-institutional consortia[19]. The Human Proteome Organization's Plasma Proteome Project has also meticulously documented the preanalytical steps required for preparation of plasma (the noncellular fraction of choice for interrogation of the proteome) to aid decision-making for suitable downstream applications[16].

Recommendations from the aforementioned international organizations focus exclusively on biospecimen collection from adults; relatively few published guidelines deal with variables unique to biospecimen collection from children. One major consideration is the need for recording the age and stage of development of the child, which is crucial for accurate comparison of biological samples with those from healthy children. Moreover, biospecimen collection in children is more complex than simply decreasing the volume of specimen (relative to sample volumes normally collected from adults). In small children, even routine venipuncture can introduce a degree of preanalytical variability owing to the introduction of shear stress and the associated risk of cell and platelet activation, particularly with the small-gauge needles often used for young patients[20,21]. Research ethics boards have their own policies around taking biological samples from children, with guidelines established for maximum allowable blood-draw volumes based on body weight and approximate total blood volumes[22]. The increasing availability of multiplex assays designed for use with small sample volumes has been fortuitous for paediatric translational research. For example, multiplex immunoassays, such as Luminex, can detect up to 100 different proteins using sample volumes as small as 50 µl (REF. 23), and high-resolution multiparameter immunophenotyping using mass cytometry quantified by time of flight (CyToF) can generate up to 35-parameter flow cytometry data from 1 ml of blood[24], making these types of platforms very useful in situations where sample volumes are limited. Advances in biomolecular technology have greatly increased the power and precision of analytical tools used in immunological research, and are accelerating the drive toward personalized medicine. Human biological samples analysed using these emerging technologies are a critical resource for translational research in rheumatology because they are a rich source of biological data from which molecular taxonomies can be derived and therapeutic targets identified[25].

Table 1 | **Best practices and guidelines from selected international networks**

Institution	Available resources	Recommendations
ISBER	• Best practice guidelines for repositories on collection, storage, retrieval and distribution of biological materials for research[18] • SPREC[26] • Evidence-based biospecimen quality control tools[30]	• Adhere to recommended best practice guidelines • Annotate standard operating protocols with SPRECs
NCI Office of Biorepositories and Biospecimen Research	• Best practice guidelines for biospecimen resources[59] • BRISQ[28]	Include BRISQ guidelines in translational research studies
EU-BioSHaRE	Ethical, legal and social implications of data and sample sharing, frameworks, tools and policies[60]	Adapt frameworks, tools and policies to strengthen institutional ethical, legal and social policies
UK Biobank	• Sample handling and storage protocols[34] • UK Biobank sample handling and storage validation studies[35] • Details of the design and implementation of a high-throughput biological sample processing facility[61]	Review protocols and validation studies for informed, evidenced-based decision making when designing translational research studies
ABNA	ABNA Biorepository Protocols (4th revision)[62]	Review protocols for informed decision making when designing translational research studies
UCAN	Best practice guidelines for the collection, processing, transfer and storage of biological samples, and access to biological data linked to clinical information in childhood arthritis[63]	Review protocols for informed decision making when designing paediatric translational research studies
BBMRI-ERIC	• MIABIS[64] • Guidelines to standardize the citation of bioresources in journal articles (CoBRA)[58]	Adhere to the proposed standardized methods for citing bioresources in journal articles
CTRNet	• Certification for biobanks[65] • Comprehensive set of standard operating procedures for a biorepository network[66]	Ensure biobanks are certified to meet the standard requirements for a biorepository network
Global Alliance for Genomics and Health, Regulatory and Ethics Working Group	Framework for responsible sharing of genomic and health-related data[67]	Maintain an awareness of the potential effects of sharing genomic and health-related data
SHARE	Work package for rare paediatric diseases, dedicated to ethical and legal barriers with multi-national collaborations[56]	Use the proposed strategies to overcome international ethical and legal barriers

ABNA, The Australasian Biospecimen Network; BBMRI-ERIC, Biobanking and Biomolecular Resources Research Infrastructure–pan-European Research Infrastructure; BRISQ, Biospecimen Reporting for Improved Study Quality; CTRNet, Canadian Tumour Repository Network; EU-BioSHaRE, Biobank Standardization and Harmonisation for Research Excellence in the European Union; ISBER, International Society for Biological and Environmental Repositories; NCI, US National Cancer Institute; MIABIS, minimum information about biobank data sharing; SHARE, Single Hub and Access Point for Paediatric Rheumatology in Europe; SPREC, Sample PREanalytical Code (standardized encoding of preanalytical data for biospecimens); UCAN, Understanding Childhood Arthritis Network.

Biospecimen handling

Biospecimen science aims to characterize the cellular and molecular alterations caused by preanalytical processing[26], to limit variability in these alterations and to prevent them from affecting downstream analytical results. Preanalytical variability can be reduced by implementing quality assurance and quality control measures that are specific for each type of sample, and by the creation of an informatics infrastructure that is capable of collecting the data needed to rigorously annotate the biospecimen collection and storage processes used. Results from biospecimen research initiatives will inform future guidelines as the community moves towards the development of evidence-based SOPs optimized for each biospecimen type and analysis platform (TABLE 1). The NCI, for instance, has spent almost a decade developing and revising its best practice guidelines for biospecimen resources. These guidelines describe the defining principles of state-of-the-art biospecimen resource management and promote adherence to ethical and legal requirements, thereby assuring the highest quality of collected biospecimens and data. The concept of biospecimen quality is difficult to define, as measures of quality for a biological sample are critically dependent on the intended downstream application or assay. Processing conditions that are optimal for one assay are suboptimal for another, necessitating careful documentation of preanalytical information and processing conditions. For example, the nature of the receptacle used for biospecimen collection is vital information to document. Blood collection tubes have multiple components, which can all affect the quality of blood samples and/or the performance of downstream assays[27], including cytokine measurements by immunoassay and functional assays of peripheral blood mononuclear cells (Supplementary information S1, S2 (tables)).

The Biospecimen Reporting for Improved Study Quality (BRISQ) guidelines[28] were developed (in collaboration with the NCI Biospecimen Research Network) to enable the capture of information on where biospecimens come from and how they are treated. These guidelines list critical data elements (such as information needed for consistent categorization of biospecimens) and factors that might influence the integrity, quality, or molecular composition of biological samples. Standardization of these elements for use in scientific publishing complements existing recommendations on data reporting. Tools to facilitate these processes, such as the Sample PREanalytical Code (SPREC)[26], a specimen barcode containing details of preanalytical sample processing, are being developed. Recognition and documentation of these critical data elements will further support the use of evidence-based protocols for biospecimen handling and storage, fostering collaboration

between biobanks and adding rigour to scientific reporting. Acknowledging the importance of every step in biospecimen management also empowers all stakeholders involved in translational research.

Although international efforts to establish best practices for biospecimen handling have greatly contributed to convergence in the principles underlying technical SOPs, some issues remain unresolved. The absence of predictive biomarkers indicating sample integrity is a critical problem in preanalytic validation studies. The development and testing of biospecimen quality control measures is in its infancy. Quality control assays differ depending on the type of sample and downstream application, and consensus is needed to ensure both consistent standards of quality and procedural uniformity. To illustrate, quantification of ribosomal RNA and calculation of the RNA integrity number are standard approaches for assessing RNA quality; however, neither method is sensitive or specific enough to predict potential errors in downstream gene expression analysis[29]. An ISBER-endorsed literature review, which summarized research on quality control measures such as preprocessing delay, freeze–thawing and storage conditions, and assessed various evidence-based assays for potential quality control biomarkers, identified soluble CD40 ligand and vascular endothelial growth factor as analytes that are likely to be useful for assessing serum exposure to temperature fluctuations and freeze–thawing[30]. However, the results of assays for soluble CD40 ligand require careful interpretation, since levels of this marker are always elevated in platelet-containing serum samples, owing to *ex vivo* platelet activation during sample preparation[31]. Low vitamin C levels have also been used as a quality control marker reflecting suboptimal serum storage conditions and pre-centrifugation delays in blood sample processing, owing to the poor intrinsic stability of this vitamin[32,33].

For some downstream applications, published information about biospecimen acquisition and processing requirements is scarce, particularly where the technology is relatively new (such as epigenomics, metabolomics and CyToF) or associated with specific functional assays (such as those examining the components of the immune system). To account for this uncertainty, SOPs used by tissue repositories need to be not only aligned with current hypotheses, approaches and regulations, but also flexible enough to enable future applications. As an example, the sample handling and storage protocols used by the UK Biobank[34], which stores specimens from 500,000 participants, are based on a review of the literature and extensive validation studies. One of the main objectives of this repository is to provide a resource that is applicable to a wide range of future scientific questions and technologies[35].

Lessons from biomarker discovery

Biomarkers are important tools with the potential to guide both the management of patients in clinical practice and the development of new therapies for rheumatic diseases. Although many biomarkers point towards molecular pathways that are directly involved in the pathogenic mechanisms of rheumatic diseases[36], technical limitations and preanalytical variability have curtailed their usefulness. Specifically, although translational studies at the DNA, RNA, protein and cellular levels have identified potential markers of susceptibility or clinical outcome in many rheumatologic diseases, few of these markers have proven to be useful for predicting individual outcomes or responses to treatment. As a result, pharmacotherapy is still largely based on a trial and error approach, in which therapies with different modes of action are administered sequentially to find the most effective agent for an individual patient[37].

A few analytes, however, such as S100 proteins[38] and serum amyloid A1 protein (SAA)[39], have shown promise in clinical studies. These biomarkers serve as nonspecific surrogate markers of inflammation, and some might be sensitive enough to detect subclinical disease activity[40,41]. One advantage common to both the S100 proteins and SAA is their *ex vivo* stability: blood samples can be collected, processed and stored using a variety of methods, with little effect on the reliability of downstream assays for their measurement. This *ex vivo* stability is in stark contrast to TNF and IL-1β, which have key roles in the immunopathogenesis of many rheumatic diseases but are highly unstable (a source of substantial preanalytical variability).

An international multicentre study of 364 patients from 29 countries found that serum levels of the S100A8 and S100A9 heterodimer (calprotectin, also known as S100A8–S100A9 or MRP8–MRP14) can be used to predict flare after cessation of methotrexate treatment in children with juvenile idiopathic arthritis (JIA) who reach clinical remission[42]. This biomarker also has some utility for predicting flare in children with JIA after cessation of anti-TNF therapy[43], and, along with protein S100A12, is a more accurate predictor of flare than C-reactive protein levels in this disease setting[40]. High serum levels of S100A8–S100A9 measured before the initiation of first-line treatment with methotrexate predict a good response to this DMARD in children with JIA[44], and a rapid decrease in serum levels of S100A12 is associated with a response to intravenous immunoglobulin therapy in children with Kawasaki disease, a multisystem vasculitis affecting young children[45].

The search continues for protein biomarkers with improved sensitivity and specificity that can be used to predict response to treatment in patients with JIA[46]. An exciting study in patients with systemic JIA showed that a panel of urinary peptide biomarkers could discriminate between active and quiescent disease[47]; if validated and shown to have predictive value, this noninvasive test could have considerable benefit. This finding echoes the encouraging early results for urinary biomarkers associated with renal disease activity in patients with systemic lupus erythematosus, such as neutrophil gelatinase-associated lipocalin and CC-motif chemokine 2 (CCL2, also known as monocyte chemoattractant protein 1)[48]. The search for relevant biomarkers in juvenile dermatomyositis has been facilitated by international efforts to define the clinical utility of myositis-specific autoantibodies[49]. A positive test for antibodies against melanoma differentiation-associated protein 5 (MDA5), for example, is associated with an increased risk of lung involvement and ulceration despite only mild muscular involvement, in both juvenile and adult dermatomyositis[50,51]. Identifying anti-MDA5-positive patients with dermatomyositis can prompt further CT screening to detect lung disease, enabling early treatment of this devastating complication. Again, progress in this area has been rapid, as the stability of the biomarker (in this case an antibody) in serum has enabled the sharing of samples and comparative data on biomarker performance between research institutes.

Translational research networks

Rheumatology has a long tradition of national and international collaboration through established research networks, many of which have now joined together to tackle issues specific to translational research. For example, the Understanding Childhood Arthritis Network (UCAN) is an international federation of research

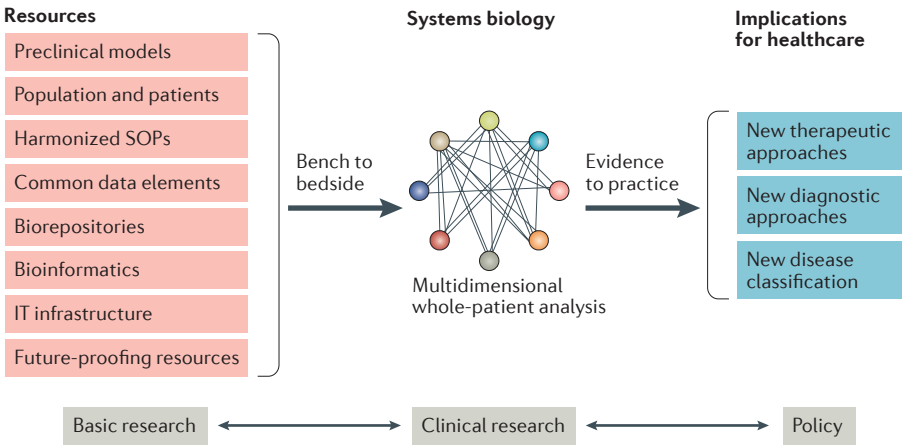

Figure 2 | **Translational research pipeline.** Basic and translational studies aim to define molecular and cellular pathogenetic mechanisms. Collaborative frameworks and infrastructure support the start of the translational medicine pipeline. Data is generated, analysed and integrated, creating evidence that will influence clinical decisions.

networks collectively representing over 50 countries. The primary goal of UCAN is harmonization of existing practices and research infrastructure, including the creation of universally accepted standards for the collection, processing, transfer and storage of biological samples from patients with childhood arthritis. UCAN also provides access to biological data linked to clinical information utilizing a common minimal dataset, and is committed to improving the efficiency of complementary and collaborative research programmes through data sharing, with the aim of improving understanding of the biologic factors underpinning the clinical heterogeneity of childhood arthritis. UCAN is also using advances in evidence-based biospecimen science and international resources to establish best practice guidelines for preanalytical handling of biospecimens for use in translational studies in childhood arthritis and rheumatic diseases — a necessary prerequisite to incorporating biomarker studies into clinical trials. The first step towards integrating knowledge of fundamental disease mechanisms with the large datasets generated by 'omics' technologies and clinical phenotyping, with a view to their use in precision medicine, is optimizing access to high-quality biospecimens and their associated data.

Standardized best practices in clinical measurements or assessments and biospecimen handling need to be paired with an information technology infrastructure that minimizes technical barriers to data sharing while enabling control over important access issues. *In silico* tools that support the acquisition, management, communication and analysis of research data are becoming more and more critical as translational and clinical research become increasingly data intensive[52]. A number of freely available software solutions, such as Research Electronic Data Capture (REDCap)[53] and Integrating Biology and the Bedside (i2B2)[54], have been developed to create an information technology infrastructure for registry-based data sharing. However, substantial challenges remain to be overcome when sharing data between registries and biobanks, including differences in terminology, data collection and database structure. Uniform standards facilitate data sharing, but end-user uptake is dictated by their flexibility and ease of application to day-to-day tasks. This pervasive tension between the need for standard methods of naming and manipulating data and the practical functionality of software leads to different solutions being used by different research groups. Several software products have been developed specifically to tackle these challenges. BiobankConnect is one such product; this system rapidly connects data elements, enabling pooled analysis across different biobanks and software solutions by semi-automatically matching desired data elements with those available using indexing terms and algorithms[55]. Advances in information technology tools that accelerate cross-platform searches and the standardization of collaborative digital infrastructure are key enhancements to the data-gathering portion of the translational medicine pipeline (FIG. 2).

The sharing of biological samples and resources goes beyond providing collaborators with a clear framework for access to data. Good sharing is costly because it requires control not only of sample quality, but also of the quality of documentation and metadata. Dedicated resources are necessary to address legal, ethical and financial barriers to data sharing, but this need is not always recognized or valued in the current academic environment. In Europe, one of the goals of the Single Hub and Access Point for Paediatric Rheumatology in Europe (SHARE) project, an initiative funded by the European Union, is to specifically identify and remove the barriers preventing effective multinational collaboration[56]. Likewise, the Reliant Review process in the USA aims to streamline the ethics review process by allowing a single institution, the central institutional review board (CIRB) of record, to provide the ethics review for a research project that spans multiple sites, increasing the efficiency of inter-institutional collaboration[57].

Other obstacles to effective sharing of biospecimens are the lack of funding and problems relating to recognition of work. These difficulties are amplified in the study of rare diseases, where the high workload relating to protocols, translation of informed consent forms and local ethical approval for an individual centre can be disproportionate to the number of patients recruited. The European Commission has stated that research data are as important as publications, and has incorporated this principle into its Horizon 2020 funding schemes. Additionally, the Biobanking and Biomolecular Resources Research Infrastructure–pan-European Research Infrastructure (BBMRI-ERIC) is exploring ways to encourage the biobanking research community by working with scientific journal editors to propose guidelines for the standardized citation of bioresources in journal articles (CoBRA)[58].

Conclusions

Moving beyond a clinical and descriptive understanding of rheumatic disease requires partnership with research. A key component of success in overcoming the hurdles in translational research is to incorporate standardization as a fundamental underlying principle when gathering, integrating and sharing data. Sharing of biological and health-related data for biomedical research is of the utmost importance in securing further advances in our understanding of human health and disease. The challenges associated

with international collaborative and translational research require a commitment by researchers to the development of a sharing framework that will facilitate responsible and excellent research with evidence-based standards and guidelines[25]. Now is the time to maximize the opportunities for advances in molecular medicine and move towards evidence-based precision care. To reach this goal, it is vital to accelerate international harmonization and standardization efforts.

Rae S. M. Yeung is at the Department of Paediatrics, Division of Rheumatology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; and at the Department of Immunology and the Institute of Medical Science, University of Toronto Faculty of Medicine, Medical Sciences Building, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada.

Salvatore Albani is at Duke-National University of Singapore Graduate Medical School, 8 College Road, 169857, Singapore.

Brian M. Feldman is at the Department of Paediatrics and Institute of Medical Science, University of Toronto Faculty of Medicine, Medical Sciences Building, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada; the Division of Rheumatology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; and at the Institute of Health Policy Management and Evaluation, The Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario M5T 3M7, Canada.

Elizabeth Mellins is at the Department of Pediatrics and the Stanford Program in Immunology, Stanford University, 300 Pasteur Drive, Stanford, California 94305, USA.

Berent Prakken is at the Department of Immunology, University Medical Centre, Heidelberglaan 100, 3584 CX Utrecht, Netherlands.

Lucy R. Wedderburn is at the Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK.

Correspondence to R.S.M.Y.
rae.yeung@sickkids.ca

doi:10.1038/nrrheum.2016.156
Published online 22 Sep 2016

1. Holzinger, D., Kessel, C., Omenetti, A. & Gattorno, M. From bench to bedside and back again: translational research in autoinflammation. *Nat. Rev. Rheumatol.* **11**, 573–585 (2015).
2. Bottini, N. & Firestein, G. S. Decade in review-translational rheumatology: Ten years after: rheumatology research from bench to bedside. *Nat. Rev. Rheumatol.* **11**, 623–624 (2015).
3. Eng, S. W. *et al.* The biologic basis of clinical heterogeneity in juvenile idiopathic arthritis. *Arthritis Rheumatol.* **66**, 3463–3475 (2014).
4. Boers, M. *et al.* Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J. Clin. Epidemiol.* **67**, 745–753 (2014).
5. Hunter, P. J. *et al.* Biologic predictors of extension of oligoarticular juvenile idiopathic arthritis as determined from synovial fluid cellular composition and gene expression. *Arthritis Rheum.* **62**, 896–907 (2010).
6. van den Ham, H. J., de Jager, W., Bijlsma, J. W., Prakken, B. J. & de Boer, R. J. Differential cytokine profiles in juvenile idiopathic arthritis subtypes revealed by cluster analysis. *Rheumatology (Oxford)* **48**, 899–905 (2009).
7. Barnes, M. G. *et al.* Subtype-specific peripheral blood gene expression profiles in recent-onset juvenile idiopathic arthritis. *Arthritis Rheum.* **60**, 2102–2112 (2009).
8. de Kleer, I. M. *et al.* $CD4^+CD25^{bright}$ regulatory T cells actively regulate inflammation in the joints of patients with the remitting form of juvenile idiopathic arthritis. *J. Immunol.* **172**, 6435–6443 (2004).
9. Schellekens, G. A. *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum.* **43**, 155–163 (2000).
10. Griffin, T. A. *et al.* Gene expression signatures in polyarticular juvenile idiopathic arthritis demonstrate disease heterogeneity and offer a molecular classification of disease subsets. *Arthritis Rheum.* **60**, 2113–2123 (2009).
11. Ransohoff, D. F. & Gourlay, M. L. Sources of bias in specimens for research about molecular markers for cancer. *J. Clin. Oncol.* **28**, 698–704 (2010).
12. Calvano, C. D. *et al.* Optimization of analytical and pre-analytical conditions for MALDI-TOF-MS human urine protein profiles. *J. Pharm. Biomed. Anal.* **51**, 907–914 (2010).
13. Scott, C. T., Caulfield, T., Borgelt, E. & Illes, J. Personal medicine — the new banking crisis. *Nat. Biotechnol.* **30**, 141–147 (2012).
14. McLerran, D. *et al.* Analytical validation of serum proteomic profiling for diagnosis of prostate cancer: sources of sample bias. *Clin. Chem.* **54**, 44–52 (2008).
15. Engel, K. B. & Moore, H. M. Effects of preanalytical variables on the detection of proteins by immunohistochemistry in formalin-fixed, paraffin-embedded tissue. *Arch. Pathol. Lab. Med.* **135**, 537–543 (2011).
16. Rai, A. J. *et al.* HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics* **5**, 3262–3277 (2005).
17. Moore, H. M., Compton, C. C., Alper, J. & Vaught, J. B. International approaches to advancing biospecimen science. *Cancer Epidemiol. Biomarkers Prev.* **20**, 729–732 (2011).
18. Campbell, L. D. *et al.* Development of the ISBER best practices for repositories: collection, storage, retrieval and distribution of biological materials for research. *Biopreserv. Biobank.* **10**, 232–233 (2012).
19. Tuck, M. K. *et al.* Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group. *J. Proteome Res.* **8**, 113–117 (2009).
20. Lippi, G., Salvagno, G. L., Montagnana, M., Brocco, G. & Cesare Guidi, G. Influence of the needle bore size used for collecting venous blood samples on routine clinical chemistry testing. *Clin. Chem. Lab. Med.* **44**, 1009–1014 (2006).
21. Gillio-Meina, C., Cepinskas, G., Cecchini, E. L. & Fraser, D. D. Translational research in pediatrics II: blood collection, processing, shipping, and storage. *Pediatrics* **131**, 754–766 (2013).
22. Berger, R. P. *et al.* Common data elements for pediatric traumatic brain injury: recommendations from the biospecimens and biomarkers workgroup. *J. Neurotrauma* **29**, 672–677 (2012).
23. de Jager, W., Bourcier, K., Rijkers, G. T., Prakken, B. J. & Seyfert-Margolis, V. Prerequisites for cytokine measurements in clinical trials with multiplex immunoassays. *BMC Immunol.* **10**, 52 (2009).
24. Benoist, C. & Hacohen, N. Immunology. Flow cytometry, amped up. *Science* **332**, 677–678 (2011).
25. National Research Council (US) Committee on A Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease* (National Acadamies Press, 2011).
26. Betsou, F. *et al.* Standard preanalytical coding for biospecimens: defining the sample PREanalytical code. *Cancer Epidemiol. Biomarkers Prev.* **19**, 1004–1011 (2010).
27. Bowen, R. A. & Remaley, A. T. Interferences from blood collection tube components on clinical chemistry assays. *Biochem. Med. (Zagreb)* **24**, 31–44 (2014).
28. Moore, H. M. *et al.* Biospecimen reporting for improved study quality (BRISQ). *J. Proteome Res.* **10**, 3429–3438 (2011).
29. Opitz, L. *et al.* Impact of RNA degradation on gene expression profiling. *BMC Med. Genomics* **3**, 36 (2010).
30. Betsou, F. *et al.* Identification of evidence-based biospecimen quality-control tools: a report of the International Society for Biological and Environmental Repositories (ISBER) Biospecimen Science Working Group. *J. Mol. Diagn.* **15**, 3–16 (2013).
31. Ivandic, B. T., Spanuth, E., Haase, D., Lestin, H. G. & Katus, H. A. Increased plasma concentrations of soluble CD40 ligand in acute coronary syndrome depend on *in vitro* platelet activation. *Clin. Chem.* **53**, 1231–1234 (2007).
32. Chung, W. Y., Chung, J. K., Szeto, Y. T., Tomlinson, B. & Benzie, I. F. Plasma ascorbic acid: measurement, stability and clinical utility revisited. *Clin. Biochem.* **34**, 623–627 (2001).
33. Karlsen, A., Blomhoff, R. & Gundersen, T. E. Stability of whole blood and plasma ascorbic acid. *Eur. J. Clin. Nutr.* **61**, 1233–1236 (2007).
34. Elliott, P., Peakman, T. C. & Biobank, U. K. The, U. K. Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int. J. Epidemiol.* **37**, 234–244 (2008).
35. Peakman, T. C. & Elliott, P. The, U. K. Biobank sample handling and storage validation studies. *Int. J. Epidemiol.* **37** (Suppl. 1), i2–i6 (2008).
36. Robinson, W. H. & Mao, R. Biomarkers to guide clinical therapeutics in rheumatology? *Curr. Opin. Rheumatol.* **28**, 168–175 (2016).
37. Duurland, C. L. & Wedderburn, L. R. Current developments in the use of biomarkers for juvenile idiopathic arthritis. *Curr. Rheumatol. Rep.* **16**, 406 (2014).
38. Kessel, C., Holzinger, D. & Foell, D. Phagocyte-derived S100 proteins in autoinflammation: putative role in pathogenesis and usefulness as biomarkers. *Clin. Immunol.* **147**, 229–241 (2013).
39. Lachmann, H. J. *et al.* Natural history and outcome in systemic AA amyloidosis. *N. Engl. J. Med.* **356**, 2361–2371 (2007).
40. Gerss, J. *et al.* Phagocyte-specific S100 proteins and high-sensitivity C reactive protein as biomarkers for a risk-adapted treatment to maintain remission in juvenile idiopathic arthritis: a comparative study. *Ann. Rheum. Dis.* **71**, 1991–1997 (2012).
41. Holzinger, D. *et al.* The Toll-like receptor 4 agonist MRP8/14 protein complex is a sensitive indicator for disease activity and predicts relapses in systemic-onset juvenile idiopathic arthritis. *Ann. Rheum. Dis.* **71**, 974–980 (2012).
42. Foell, D. *et al.* Methotrexate withdrawal at 6 versus 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* **303**, 1266–1273 (2010).
43. Anink, J. *et al.* MRP8/14 serum levels as a predictor of response to starting and stopping anti-TNF treatment in juvenile idiopathic arthritis. *Arthritis Res. Ther.* **17**, 200 (2015).
44. Moncrieffe, H. *et al.* A subgroup of juvenile idiopathic arthritis patients who respond well to methotrexate are identified by the serum biomarker MRP8/14 protein. *Rheumatology (Oxford)* **52**, 1467–1476 (2013).
45. Foell, D. *et al.* S100A12 (EN-RAGE) in monitoring Kawasaki disease. *Lancet* **361**, 1270–1272 (2003).
46. Palman, J., Ursu, S., Heywood, W., Mills, K. & Wedderburn, L. Development of novel protein biomarkers for the prediction of response to treatment in juvenile idiopathic arthritis. *Pediatr. Rheumatol. Online J.* **12** (Suppl. 1), Y1 (2014).
47. Ling, X. B. *et al.* Urine peptidomic and targeted plasma protein analyses in the diagnosis and monitoring of systemic juvenile idiopathic arthritis. *Clin. Proteomics* **6**, 175–193 (2010).
48. Abulaban, K. M. & Brunner, H. I. Biomarkers for childhood-onset systemic lupus erythematosus. *Curr. Rheumatol. Rep.* **17**, 471 (2015).
49. Tansley, S. & Wedderburn, L. R. Comparing and contrasting clinical and serological features of juvenile and adult-onset myositis: implications for pathogenesis and outcomes. *Curr. Opin. Rheumatol.* **27**, 601–607 (2015).
50. Cao, H. *et al.* Clinical manifestations of dermatomyositis and clinically amyopathic dermatomyositis patients with positive expression of anti-melanoma differentiation-associated gene 5 antibody. *Arthritis Care Res. (Hoboken)* **64**, 1602–1610 (2012).
51. Tansley, S. L. *et al.* Anti-MDA5 autoantibodies in juvenile dermatomyositis identify a distinct clinical phenotype: a prospective cohort study. *Arthritis Res. Ther.* **16**, R138 (2014).
52. Masys, D. R., Harris, P. A., Fearn, P. A. & Kohane, I. S. Designing a public square for research computing. *Sci. Transl Med.* **4**, 149fs132 (2012).
53. Harris, P. A. *et al.* Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).

54. Kohane, I. S., Churchill, S. E. & Murphy, S. N. A translational engine at the national scale: informatics for integrating biology and the bedside. *J. Am. Med. Inform. Assoc.* **19**, 181–185 (2012).
55. Pang, C. *et al.* BiobankConnect: software to rapidly connect data elements for pooled analysis across biobanks using ontological and lexical indexing. *J. Am. Med. Inform. Assoc.* **22**, 65–75 (2015).
56. Wulffraat, N. M., Vastert, B. & SHARE consortium. Time to share. *Pediatr. Rheumatol. Online J.* **11**, 5 (2013).
57. Ervin, A. M., Taylor, H. A., Meinert, C. L. & Ehrhardt, S. Evidence gaps and ethical review of multicenter studies. *Science* **350**, 632–633 (2015).
58. Bravo, E. *et al.* Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA). *BMC Med.* **13**, 33 (2015).
59. National Cancer Institute. *NCI best practices for biospecimen resources biospecimens.cancer.gov* http://biospecimens.cancer.gov/bestpractices/2011-NCIBestPractices.pdf (2011).
60. Budin-Ljosne, I. *et al.* ELSI challenges and strategies of national biobank infrastructures. *Norweg. J. Epidemiol.* **21**, 155–160 (2012).
61. Downey, P. & Peakman, T. C. Design and implementation of a high-throughput biological sample processing facility using modern manufacturing principles. *Int. J. Epidemiol.* **37** (Suppl. 1), i46–i50 (2008).
62. Australasian Biospecimen Network. *ABN Protocol Guidelines* http://abrn.net/what-we-do/protocols/ (2012).
63. Understanding Childhood Arthritis Network-Utrecht. *UCAN-U Knowledge Base* http://www.ucan-u.org/kb/mainpage (2015).
64. Norlin, L. *et al.* A minimum data set for sharing biobank samples, information, and data: MIABIS. *Biopreserv. Biobank.* **10**, 343–348 (2012).
65. Matzke, E. A. *et al.* Certification for biobanks: the program developed by the Canadian Tumour Repository Network (CTRNet). *Biopreserv. Biobank.* **10**, 426–432 (2012).
66. Barnes, R. *et al.* Generating a comprehensive set of standard operating procedures for a biorepository network-The CTRNet experience. *Biopreserv. Biobank.* **11**, 387–396 (2013).
67. Knoppers, B. M. Framework for responsible sharing of genomic and health-related data. *HUGO J.* **8**, 3–9 (2014).

Author contributions

R.S.M.Y. researched the data and wrote the article. All authors provided a substantial contribution to discussions of the content and contributed to review and/or editing of the manuscript.

Competing interests statement

R.S.M.Y. declares that she has received grant support from Novartis. E.D.M. declares that she has received grant support from GlaxoSmithKline and Novartis. The other authors declare no competing interests.

FURTHER INFORMATION

European Commission. Horizon 2020 The EU Framework Programme for Research and Innovation: http://ec.europa.eu/programmes/horizon2020

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF