Nature Reviews Rheumatology; Manuscript number: NRR-15-215V1

Title: Enhancing Translational Research in Paediatric Rheumatology Through Standardization

Authors: Rae S. M. Yeung[1,2,3], Salvatore Albani[4], Brian M. Feldman[1,3,5], Elizabeth Mellins[6], Berent Prakken[7], Lucy R. Wedderburn[8]

Affiliations:

1. Department of Paediatrics and Institute of Medical Science, Faculty of Medicine, University of Toronto, ON, Canada.
2. Department of Immunology, University of Toronto, Toronto, ON, Canada.
3. Division of Rheumatology, The Hospital for Sick Children, Toronto, ON, Canada.
4. Duke-NUS Graduate Medical School, Singapore
5. Institute of Health Policy Management and Evaluation, The Dalla Lana School of Public Health, University of Toronto, ON, Canada.
6. Department of Pediatrics, Program in Immunology, Stanford University, Stanford, CA, USA.
7. Department of Immunology, UMCU, Utrecht, The Netherlands.
8. Institute of Child Health University College London, Great Ormond Street Hospital London, UK.

To whom correspondence should be addressed:
Rae S M Yeung, MD, PhD, FRCPC
The Hospital for Sick Children
555 University Avenue
Toronto, Ontario, Canada M5G 1X8
Telephone: +1-416-813-8964, Fax: +1-416-813-4989
Email: rae.yeung@sickkids.ca

Abstract:

Translational studies in rheumatology have seen many successes in the last decade with dramatic improvements in outcomes with the introduction of new biologically-based therapeutic agents for management of inflammatory diseases and the discovery of new monogenic disorders of inflammation. Yet pharmacotherapy is still largely based on a trial-and-error approach with sequential use of therapies with different modes of action to find the most effective agent for an individual patient. Advances in molecular medicine combined with the drive towards precision care provide a significant opportunity to accelerate translation of biological understanding to the bedside. However, relative to the clinical component, where rheumatology is at the forefront of standardized data collection and measures of disease activity, standardized biologic sample collection and assay performance lag behind. Uniform approaches are required for robust collaborative research into pathobiology, especially for diseases where patient numbers at single institutions are small. Standardization is also critical to increase reproducibility between centers, which is a requisite step towards clinical implementation based on translational science. In the following discussion, we emphasize the need for standardization and best practices, 2) highlight current work and new directions in biospecimen science; and 3) review lessons learned from international networks.

The introduction of rheumatic disease therapies that selectively target cytokines or immune cell subsets has led to dramatic improvements in clinical outcomes for many paediatric and adult patients.[1,2] However, to achieve the full potential of these and other immune interventions will require systematic characterization and classification of patients, incorporating both clinical and biologic measures. Heterogeneity among patients in rheumatic diseases is reflected in the fact that few therapies are effective in all patients carrying the same diagnosis, and this variable efficacy persists even within clinically defined subgroups of patients with a given disease.[3] This heterogeneity is also reflected in the variations in results from molecular evaluations of patients with the same diagnosis, resulting in the equivalent of many diseases, each with smaller numbers, such that elucidating pathophysiology or identifying diagnostic or prognostic biomarkers is likely to require collaboration. However the lack of reproducibility of many biomarker studies suggests that in addition to the true biological heterogeneity, additional procedural issues likely add substantial variability. Differences in biologic sample handling is one of the most common sources of site-dependent results observed in current biomarker studies. Collaborative research involving multiple centers and groups requires harmonization of procedures to enable proper comparisons. The need for standardized operating procedures (SOP) for collection and handling of samples and data is a critical first step to ensure high-quality translational research. Sharing this information with researchers requires a flexible and secure data-sharing infra-structure supporting the integration of biologic data with precise clinical measures.

Relative to the clinical component, where rheumatology is at the forefront of clinical data collection with use of standardized measures of disease activity and damage, there remains significant disparity in biologic sample collection and standardization of biologic measures of disease activity.

Internationally accepted measures of clinical disease activity and damage are widely used in both adult and pediatric rheumatic diseases and many have been incorporated as standard-of-care in clinical assessments.[4] In contrast, biologic measures have not been incorporated at the bedside, despite evidence demonstrating that DNA, gene and protein expression and cellular immunophenotyping profiles help stratify patients and define homogeneous patient groups and point to molecular determinants of susceptibility and outcome in many rheumatologic diseases including childhood arthritis and rheumatic diseases.[3,5,6] Additional challenges arise when research is focused on rare diseases. Proper solutions require large-scale methodologic and organizational efforts aimed at bringing groups together that are able to collect, manage and share datasets. Uniform standards remove some of the barriers to sharing. Addressing standardization throughout the life-cycle of clinical and biologic data will accelerate the incorporation of translational studies into clinical decision-making (Figure 1).

Standardization and best practices guidelines

Identification of robust biomarkers can dramatically improve diagnostic and treatment decisions. For example, the presence of anti-CCP antibodies have defined a subgroup of inflammatory arthritis.[8] Identification of biomarkers can be crucial in diagnostic and treatment decisions, yet variability can alter results of biological assays. Often specimen collection is perceived as the easy part of a protocol, yet validation of biomarkers relies on the control of error introduced during every step in the collection process (Figure 1). It is clear that in order to increase efficiencies and achieve appropriate sample sizes in translational studies, a concerted effort must be made to create a standard of practice for establishing common biobanking principles. In order to ensure the quality of downstream molecular analyses, potential variability due to biospecimen collection, processing and handling must

be identified and controlled by adhering to uniform SOPs. Important lessons have been learned from both successes and failures in multicenter translational studies in pediatric rheumatology – highlighting the need for stringent procedures, processing times and quality control. [9,10]

Among the different types of experimental variability, preanalytical variations, including pre- and post- sample acquisition, are often not appreciated and thus not controlled.[11,12] Preanalytical processes are defined as those procedures taking place between specimen collection and experimental assay/analysis. Even small differences in SOPs between groups or within groups could yield uninterpretable results due to variations at multiple steps in the collection and handling process.[13] The National Cancer Institute's Cancer Genome Atlas project was launched on the assumption that its researchers could obtain quality specimens provided by dozens of established biobanks with strong track records of success. Unfortunately, results indicated that very few of the initial samples were usable due to variability in collection, processing and storage in legacy biorepositories.[13] Variables that may impact analytic outcomes include: 1) the type of blood collection tube and additive, 2) sample processing protocols, times and temperatures, 3) hemolysis, 4) transport conditions, 5) storage parameters (temperatures and freeze-thaw cycles). [14-16] These types of differences between samples have a significant impact on the stability of analytes causing deterioration in data quality. The challenge to harmonization of SOPs is to move forward with SOPs developed from the current state of the science and based on empirical evidence and not on ritual.

Valuable lessons have been learned from international efforts addressing standardization of preanalytical biospecimen processing.[17] The International Society for Biological and Environmental Repositories (ISBER) has produced harmonized high-quality best practices, which they publish and

update regularly [current version 2012].[18] ISBER systematically reviews the biospecimen science literature, furnishing evidence and links to publications. Recommended protocols are available for collection and handling of a comprehensive list of biologic samples including cellular and noncellular fractions from human peripheral blood, but they clearly specify that SOPs must consider the analytical endpoints and available resources. Similarly, the Early Detection Research Network (EDRN) endorsed by the International Agency for Research on Cancer has also proposed similar SOPs for noncellular fractions from peripheral blood including plasma and serum in multi-institutional consortium environments.[19] The HUPO Plasma Proteome project noted that plasma is the noncellular fraction of choice for interrogation of the proteome, with meticulous documentation of preanalytical steps to aid in decision-making for suitable downstream applications.[16]

The recommendations from all these international organizations, however, focus exclusively on biospecimen collection from adult subjects. There are relatively few published biospecimen collection guidelines dealing with variables unique to children. Comparison to healthy controls needs to take into consideration age and stage of development, and biospecimen collection in children is more complex than simply decreasing the volume of specimens that are collected in adults. Even routine venipuncture can introduce a degree of preanalytical variability with the introduction of shear stress and the risk of cell and platelet activation, particularly with smaller gauge needles that are often used in young patients.[20, 21] Research ethics boards have their own policies around sampling in children with guidelines established around maximum allowable blood draw volumes based on body weight and approximate blood volumes.[22] Increasing availability of multiplex assays designed with a small sample size in mind have been fortuitous for paediatric translational research. For example, multiplex immunoassays, such as Luminex, can detect up to 100 different proteins using as little as 50 ul sample volume[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 where sample volumes are limited. These and other advances in biomolecular technology have greatly increased the power and precision of analytical tools used in immunologic research and have accelerated the drive toward personalized medicine. Human specimens that are analyzed using these emerging technology platforms are a critical resource for translational research in rheumatology because they are a rich source of biologic data from which molecular taxonomies can be derived and targets for therapy identified.[25]

Biospecimen science

Biospecimen science aims to determine the cellular and molecular alterations attributed to preanalytical processes.[26] The main goal is to control variability and limit how it might affect downstream analytic results. It is possible to significantly reduce preanalytical variability by developing quality assurance and quality control measures specific for each type of sample or analyte, and to partner that with an informatics infrastructure capable of collecting the data needed to rigorously annotate the biospecimen collection and storage processes. A number of groups are active in biospecimen science (Table 1). The National Cancer Institute has spent almost a decade developing and revising *Best Practice Guidelines* for biospecimen resources which describes guiding principles that define the-state-of-the-science for biospecimen resource practices, promote biospecimen and data quality, and support adherence to ethical and legal requirements. Results from biospecimen research initiatives will inform future guidelines as the community moves toward the development of evidence-based SOPs that are both biospecimen-type and analysis-platform specific. The concept associated with 'quality' in a biospecimen cannot be uniquely defined, as these are critically

dependent on the downstream application or assay. Thus processing conditions that are optimal for one assay are not the same for another, necessitating careful documentation of preanalytical information and harmonization of the methods to document these conditions. The BRISQ (Biospecimen Reporting for Improved Study Quality) guidelines[27] were developed in collaboration with the NCI Biospecimen Research Network to document where biospecimens came from and how they were treated. The guidelines list critical data elements that include general information for consistent documentation of categories of biospecimens and factors that might influence the integrity, quality, and/or molecular composition of biospecimens. Standardizing and codifying these elements for reporting and communication in scientific publications[26] will serve to complement existing reporting recommendations. Tools to facilitate these processes are being developed, such as the Sample Preanalytical Code (SPREC), a specimen barcode with details about preanalytical sample processing.[26] Recognizing and documenting these critical elements will further support evidence-based biobanking, foster collaborations between biobanks, add rigor to scientific reporting, and also empower all stakeholders involved in translational research by recognizing the importance of every step in biospecimen management.

One critical element to document is the type of collection tube used for biosample collection. Blood collection tubes (BCTs) have multiple components, all of which can affect the quality of the biospecimen and/or the performance of downstream assays.[28] Different BCTs affect performance of multiple downstream assays including cytokine measurements by immunoassay to functional assays with PBMCs. Supplementary Table 1 summarizes the pre-analytical variability associated with BCTs and the advantages and special considerations associated with different BCTs. Supplementary Table 2 describes biomarkers of interest in rheumatology and the pre-analytical variables directly affecting

them.

Although international efforts to establish best practices for handling biospecimens have greatly contributed to convergence in principles for technical SOPs, some issues remain unsolved. One of the critical problems in preanalytic validation studies is the absence of key biomarkers that will predict sample integrity. Quality control (QC) assays are different depending on the sample type and downstream application. Consensus is needed on both quality control and uniformity of procedures. To illustrate, ribosomal RNA and RNA integrity number (RIN) are standard approaches for assessing RNA; however, neither method is sensitive nor specific enough to assess potential error in downstream gene expression analysis.[29] Development and testing of various QC measures for biospecimens is in its infancy. ISBER endorsed a review of the literature,[30] summarizing existing research on QC measures (preprocessing delay, freeze thawing, storage conditions) and assessing potential evidence-based QC assays summarizing potential QC biomarkers. CD40L and VEGF were identified as potentially meaningful analytes for assessing serum exposure to variations in temperatures for assessing serum freeze-thawing.[30] However a caveat was that soluble CD40L levels are already artificially elevated in serum samples (which are not platelet free), and the increased sCD40L concentrations are a result of ex-vivo platelet activation during sample preparation and not due to in-vivo factors, therefore interpretation of studies of sCD40L already require caution.[31] Ascorbic acid has also been used as a QC marker for blood pre-centrifugation delay and serum storage conditions owing to its intrinsic instability.[32,33] Pitfalls are readily apparent in even these examples of anlaytes proposed as potential QC tools.

To date, there is little information around acquisition and processing requirements for some downstream applications, particularly where the technology is relatively new (i.e. epigenomics, metabolomics, CyToF) or associated with specific functional assays. Tissue repositories need to be compliant with current regulations and aligned with current hypotheses and approaches but also flexible enough to allow for future testing. As an example, UK Biobank, following half a million participants, developed sample handling and storage protocols[34] based on existing literature review and extensive validation studies with one of the main objectives to provide a resource that has applicability to a wide range of future scientific questions and technologies.[35]

Lessons from biomarker discovery

Biomarkers are important tools that have significant potential for guiding both clinical management of and therapeutic development for rheumatic diseases. Although translational studies from DNA to RNA, protein and cellular phenotyping have contributed to identifying molecules as potential determinants of susceptibility and outcome in many rheumatologic diseases, to date, few have proven useful for predicting response to treatment. Pharmacotherapy is still largely based on a trial-and-error approach with sequential use of therapies with different modes of action to find the most effective agent for an individual patient.[36] However, a few analytes have shown promise at the bedside. These include S100 proteins[37] and serum amyloid A (SAA),[38] promising surrogate biomarkers that serve as non-specific measures of inflammation that may be able to detect subclinical disease activity.[39,40, 41] One common advantage to S100 proteins and SAA is their ex-vivo stability. Blood samples can be collected, processed, stored using varying methods and the preanalytical variability has little effect on their measurement. This is in stark contrast to TNFα and IL-1β, which play key roles in immunopathogenesis of rheumatic diseases but are unstable molecules and subject to

significant preanalytical variability. Although these mechanistic biomarkers point to molecular pathways directly involved in disease pathogenesis,[42] technical limitations and preanalytical variability have prevented meaningful guidance for clinical decisions.

A multi-centre international study, which included 364 patients from 29 countries, showed that serum levels of S100A8/9 (also called MRP8/14 or calprotectin) provide predictive value in children with juvenile idiopathic arthritis (JIA) who reach clinical remission on methotrexate, allowing health care professionals and families to estimate risk of flare after MTX is stopped.[43] This biomarker also has some ability to predict flare of disease after stopping anti-TNF therapy.[44] Both S100 A8/9 and S100 A12 serum levels have been shown to be more accurate predictors of flare than the C reactive protein.[39] Serum S100A8/9 measured before starting treatment with MTX has also been shown to have value in predicting good response to this first line DMARD in JIA,[45] and decrease in S100A12 was associated with response to IVIG therapy in children with Kawasaki Disease, a multisystem vasculitis affecting young children.[46] The search for predictive biomarkers with better sensitivity and specificity to guide treatment is ongoing: some early results are promising.[47] An exciting development in systemic JIA, is a study showing that a panel of urine peptide biomarkers can discriminate between active and quiescent disease, if validated and shown to have predictive value this approach would have considerable benefit since it could be adapted to non-invasive testing.[48] This echoes encouraging early results for urinary biomarkers associated with renal disease activity in systemic lupus erythematosis including neutrophil gelatinase-associated lipocalin and monocyte chemoattractant protein-1. [49] In a more rare disease, juvenile dermatomyositis (JDM) recent progress has been facilitated by international efforts to define clinical utility of myositis specific autoantibodies (MSA).[50] Thus for example the MSA anti-MDA5 has been shown to be associated with an increased

risk of lung involvement and ulceration, yet mild muscle involvement, in both juvenile and adult DM.[51,52] Identifying anti-MDA-5 positive patients can enable careful screening by CT scan to detect lung disease and allow early treatment of this devastating complication. Again this area of progress has been rapid in part due to the fact that the biological stability of the biomarker protein (antibody) in serum enables sharing and biomarker comparisons from a wide range of centres.

Translational Research Networks

Rheumatology has long embraced collaboration through many established national and international networks. More recently, established research networks have joined together to tackle issues specific to translational research. For example, UCAN (Understanding Childhood Arthritis) is an international federation of research networks collectively representing over 50 countries and is committed to improving efficiencies by sharing complementary and collaborative research programs towards understanding the biology of clinical heterogeneity in childhood arthritis. UCAN was formed with the primary goal of harmonizing existing practices and creating a universally accepted standard for collection, processing, transfer, storage and access to biologic data linked to clinical information with a common minimal dataset for childhood arthritis. Using advances in evidence-based biospecimen science and international resources, UCAN is establishing best practice guidelines for preanalytical handling of biospecimens for use in translational studies in childhood arthritis and rheumatic diseases - a necessary prerequisite to incorporating biologic companion studies to clinical trials. Optimizing access to high-quality biospecimens and their associated data is the first step towards integrating the rapidly expanding knowledge of fundamental disease mechanisms with data-intensive biology from omic analysis with clinical phenotype towards precision medicine.

Standardized best practices in clinical measures and biospecimens need to be curated by an information technology infra-structure that minimizes technical barriers to sharing while enabling control over important data access issues. Computerized tools and infrastructure to support the acquisition, management, communication and analysis of research data are more and more critical, as translational and clinical research become increasingly data intensive.[53] A number of efforts are in progress to create information technology infrastructure for registry-based data-sharing. Examples from NIH sponsored research computing efforts with widespread uptake include REDCap (Research electronic data capture) [54] and i2B2 (Integrating Biology and the Bedside),[55] both freely available software solutions. When sharing between registries and biobanks is necessary, the hurdles are challenging with differences in terminology, data collections and database structure. Data standards facilitate data sharing, but flexibility and nimble responses to practical day-to-day tasks dictate end-user uptake. This tension between requirements for standard methods of naming and manipulating data and the practical functionality of the software solution is ubiquitous, leading to different solutions used by different researchers. Software solutions have also been developed to tackle these challenges – BiobankConnect is one example of a system that rapidly connects data elements for pooled analysis across different biobanks and software solutions by semi-automatically matching desired data elements with available ones using indexing terms and algorithms.[56] Standardized, collaborative digital infrastructure and advances in information technology tools to accelerate cross-platform searches are key enhancements to the data gathering process in the translational medicine pipeline (Figure 2).

Collaboration and sharing of biologic samples and resources goes beyond providing access to data. Good sharing requires quality control for not only the sample but also documentation, metadata and

a clear framework for access. This is costly and requires dedicated resources to address barriers, but is not always recognized nor valued in our current academic settings. Legal, ethical and funding challenges present as fundamental barriers that precede collaboration and sharing. One of the goals of a European Union funded initiative – the SHARE (Single Hub and Access point for paediatric Rheumatology in Europe) project – aims to specifically identify and resolve the barriers preventing effective multi-national collaboration. In the USA, one approach to streamline ethics review is Reliant Review. A Reliant Review is a process that allows investigators to collaborate across institutions in a more efficient way by allowing a single institution to provide ethics review for a research project that involves multiple sites. The institutional review board (IRB) that performs the ethics review is designated the 'IRB of Record' and collaborating organizations agree to accept this review and are designated as the 'Relying IRBs', making the process more efficient and streamlined.

Lack of funding and recognition is another obstacle to effective sharing of biospecimens. This is amplified in the study of rare diseases, where the work-load for individual centers may be heavy with regard to protocols, translated informed consent forms and local ethical approvals, despite the fact that sometimes only one or two patients will be recruited from each center. The European Commission has recently recognized that research data are as important as publications and has incorporated this principle in current funding schemes.[57] Additionally, the pan-European initiative, The Biobanking and Biomolecular resources Research Infrastructure (BBMRI-ERIC) is exploring ways to encourage the biobanking research community through harmonized citation and recognition processes.[58] Working together with scientific journal editors, guidelines for the standardized citation of bioresources in journal articles (CoBRA) have been proposed.[58]

Moving beyond a clinical and descriptive understanding of rheumatic diseases will require partnership with biology. We have reviewed some key ingredients for success in overcoming the hurdles in translational research, which include gathering, integrating and sharing data, with standardization as the fundamental underlying principle. Sharing of biologic and health-related data for biomedical research is of utmost importance in securing continued advancement in our understanding of human health and disease. The challenges associated with international collaborative translational research require a commitment to developing a sharing framework that will facilitate responsible and excellent research with evidence-based standards and guidelines.[25] Now is the time. In order to maximise the opportunities for advances in molecular medicine, and move towards evidence-based precision decisions for treatment, it is vital to accelerate these harmonization and standardization efforts.

Figures Legends:

Figure 1: Biospecimen Science - Standardized operating procedures for collection and handling of samples and data is a critical first step to ensure high quality translational research. Preanalytical variables impact various steps from specimen collection to downstream analysis. These include 1) The physiology of the subject; 2) Uniformity in biospecimen collection; 3) Biospecimen handling procedures, illustrated in yellow, blue and red, respectively. Quality control measures may be diagnostic of upstream collection, processing and/or storage processes, or a predictive assay for the validity of the downstream analysis. (Adapted from reference 30)

Figure 2. Translational Research Pipeline – Resources and infra-structure support the start the of the translational medicine pipeline with fundamental studies and applied research that define molecular and cellular mechanisms and their relationship to disease. Data is generated, analyzed and integrated creating a knowledge network with the evidence to impact clinical decisions. (Adapted from reference 25)

Table 1: Summary of select international biospecimen collection, storage, handling and sharing best practices and guidelines

Institution	Document(s) and select references	Recommendations
International Society for Biological and Environmental Biorepositories	• Best practice guidelines for repositories: collection, storage, retrieval and distribution of biological materials for research[18] • Standard preanalytical coding for biospecimens defining the sample PREanalytical code (SPREC)[26] • Identification of evidence-based biospecimen quality control tools[30]	• Adhere to recommended best practice guidelines • Annotate standard operating protocols with SPREC codes
National Cancer Institute (NCI), Office of Biorepositories and Biospecimen Research (OBBR), United States	• Best practices for biospecimen resources[59] • Biospecimen reporting for improved study quality (BRISQ)[27]	• Include BRISQ guidelines in translational research studies
Biobank Standardization and Harmonisation for Research Excellence in the European Union (EU-BioSHaRE)	• Ethical, legal and social implications of data and sample sharing, framework, tools and policies[60]	• Adapt framework, tools and policies to strengthen institutional ethical, legal and social policies
UK Biobank, Ltd.	• Sample handling and storage protocols[34] • The UK Biobank sample handling and storage validation studies[35] • Design and implementation of a high-throughput biological sample processing facility[61]	• Review protocols and validation studies for informed, evidenced-based decision making around translational research study design
The Australasian Biospecimen Network	• Australasian Biospecimen Network Biorepository Protocols – 4th Revision[62]	• Review protocols for informed decision making around translational research study design
Understanding Childhood Arthritis Network (UCAN) International Federation	• Best practice guidelines for the collection, processing, transfer, storage and access to biologic data linking to clinical information in childhood arthritis[63]	• Review protocols for informed decision making around paediatric translational research study design
Biobanking and Biomolecular resources Research Infrastructure (BBMRI-ERIC), pan-European research infrastructure	• A minimum data set for sharing biobank samples, information, and data: MIABIS[64] • BBMRI subgroup – Bioresource Research Impact Factor: Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA)[58]	• Adhere to the proposed standardized methods for citing bioresources in journal articles
Canadian Tumour	• Certification for biobanks[65]	• Ensure biobanks are

Repository Network (CTRNet)	• Comprehensive set of standard operating procedures for a biorepository network[66]	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 impact on sharing genomic and health-related data
Single Hub and Access point for paediatric Rheumatology in Europe (SHARE) Project	• Paediatric rare diseases project • Work package dedicated to ethical and legal barriers with multi-national collaborations	• Use the proposed strategies to overcome international ethical and legal barriers

Supplemenarey Table 1: Pre-analytical variability with blood collection tubes (BCT)

Tube	Advantages	Potential Limitations
EDTA	• Hematology testing[28] • Proteomic analysis[16] • DNA studies	• Cytogenetic assays • Immunoassays • Binds required reagents and co-factors[68] • Affects antibody binding[68]
Heparin	• PBMC Immediate isolation for functional assays • Lithium heparin is recommended over sodium heparin for downstream T cell assays [69]	• Ag-Ab reactions (slows interaction) [70] • Cryoprotein measurement (precipitates cryoprotein) [28] • Albumin measurements[71] • Creatine kinase, γ-glutamyl transferase[72]
Citrate	• Coagulation testing[73] • Citrate-stabilized blood results in better quality RNA and DNA, yields more lymphocytes for culture	• Aspartate aminotransferase • Alkaline phosphatase[73]
Serum Separator Tubes (SST)	• Serum testing including S100 proteins and antibodies	• Hydrophobic drugs including phenytoin, phenobarbital, carbamazepine, etc (absorbed by gel) [74] • Myoglobin and CK-MB[75] • Testosterone
P100 (additive: protease inhibitor cocktail)	• Well suited for proteomic studies, including analysis of unstable proteins (ie TNFα)[76] • Provides an option for delayed processing after blood collection	• Gel stopper malfunction could interfere with sample • Potential for hemolysis • Cost

Supplementary Table 2: Pre-analytical variables affecting biomarker measurements

Biomarkers	Pre-analytical variable and impact[21]
CYTOKINES: IL-1α, IL-1B, IL-2, IL-4, IL-5, IL-6, IL-13, IL-15, IL-17, IL-18, TNFα, INFγ, CXCL-8[23]	• Cytokine stability is affected by additives in the blood collection tube (serum, sodium heparin plasma, EDTA plasma, sodium citrate plasma[23], P100 plasma[76])
PLATELET DERIVED PROTEINS: CD40L, VEGF, MMPs[31,77]	• Platelet derived proteins are falsely elevated in serum due to platelet release during activation and/or coagulation process
MATRIKINES: • MMP-1, MMP-2, MMP-13 • MP-2, MMP-9	• Matrikines are generally elevated in plasma (regardless of blood collection tube) due to isolation dependent activation of neutrophils and mononuclear cells[78,79]
OTHER BIOMARKERS: • Troponin • Glucose	• Other biomarkers are decreased in plasma from heparinized tubes due to binding[8] • Other biomarkers are decreased in plasma (regardless of blood collection tube) due to an anticoagulant-induced fluid shift from erythrocytes to plasma[9]

References

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, doi:10.1038/nrrheum.2015.79 (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, doi:10.1038/nrrheum.2015.126 (2015).

3 Eng, S. W. *et al.* The biologic basis of clinical heterogeneity in juvenile idiopathic arthritis. *Arthritis Rheumatol* **66**, 3463-3475, doi:10.1002/art.38875 (2014).

4 Boers, M. *et al.* Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* **67**, 745-753, doi:10.1016/j.jclinepi.2013.11.013 (2014).

5 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, doi:10.1093/rheumatology/kep125 (2009).

6 de Kleer, I. M. *et al.* CD4+CD25bright 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).

7 Kastner, D. L., Aksentijevich, I. & Goldbach-Mansky, R. Autoinflammatory disease reloaded: a clinical perspective. *Cell* **140**, 784-790, doi:10.1016/j.cell.2010.03.002 (2010).

8 Schellekens, G. A. *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum* **43**, 155-163, doi:10.1002/1529-0131(200001)43:1<155::AID-ANR20>3.0.CO;2-3 (2000).

9 Barnes, M. G. *et al.* Subtype-specific peripheral blood gene expression profiles in recent-onset juvenile idiopathic arthritis. *Arthritis Rheum* **60**, 2102-2112, doi:10.1002/art.24601 (2009).

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, doi:10.1002/art.24534 (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, doi:10.1200/JCO.2009.25.6065 (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, doi:10.1016/j.jpba.2009.10.014 (2010).

13 Scott, C. T., Caulfield, T., Borgelt, E. & Illes, J. Personal medicine--the new banking crisis. *Nat Biotechnol* **30**, 141-147, doi:10.1038/nbt.2116 (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, doi:10.1373/clinchem.2007.091470 (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, doi:10.1043/2010-0702-RAIR.1 (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, doi:10.1002/pmic.200401245 (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, doi:10.1158/1055-9965.EPI-11-0021 (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, doi:10.1089/bio.2012.1025 (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, doi:10.1021/pr800545q (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, doi:10.1515/CCLM.2006.172 (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, doi:10.1542/peds.2012-1181 (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, doi:10.1089/neu.2011.1861 (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, doi:10.1186/1471-2172-10-52 (2009).
24 Benoist, C. & Hacohen, N. Immunology. Flow cytometry, amped up. *Science* **332**, 677-678, doi:10.1126/science.1206351 (2011).
25 in *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease The National Academies Collection: Reports funded by National Institutes of Health* (2011).
26 Betsou, F. *et al.* Standard preanalytical coding for biospecimens: defining the sample PREanalytical code. *Cancer Epidemiol Biomarkers Prev* **19**, 1004-1011, doi:10.1158/1055-9965.EPI-09-1268 (2010).
27 Moore, H. M. *et al.* Biospecimen reporting for improved study quality (BRISQ). *J Proteome Res* **10**, 3429-3438, doi:10.1021/pr200021n (2011).
28 Bowen, R. A. & Remaley, A. T. Interferences from blood collection tube components on clinical chemistry assays. *Biochem Med (Zagreb)* **24**, 31-44, doi:10.11613/BM.2014.006 (2014).
29 Opitz, L. *et al.* Impact of RNA degradation on gene expression profiling. *BMC Med Genomics* **3**, 36, doi:10.1186/1755-8794-3-36 (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, doi:10.1016/j.jmoldx.2012.06.008 (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, doi:10.1373/clinchem.2007.085332 (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, doi:10.1038/sj.ejcn.1602655 (2007).
34 Elliott, P., Peakman, T. C. & Biobank, U. K. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int J Epidemiol* **37**, 234-244, doi:10.1093/ije/dym276 (2008).

35 Peakman, T. C. & Elliott, P. The UK Biobank sample handling and storage validation studies. *Int J Epidemiol* **37 Suppl 1**, i2-6, doi:10.1093/ije/dyn019 (2008).
36 Duurland, C. L. & Wedderburn, L. R. Current developments in the use of biomarkers for juvenile idiopathic arthritis. *Curr Rheumatol Rep* **16**, 406, doi:10.1007/s11926-013-0406-3 (2014).
37 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, doi:10.1016/j.clim.2012.11.008 (2013).
38 Lachmann, H. J. *et al.* Natural history and outcome in systemic AA amyloidosis. *N Engl J Med* **356**, 2361-2371, doi:10.1056/NEJMoa070265 (2007).
39 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, doi:10.1136/annrheumdis-2012-201329 (2012).
40 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, doi:10.1136/annrheumdis-2011-200598 (2012).
41 Miyamae, T. *et al.* Serum protein profile in systemic-onset juvenile idiopathic arthritis differentiates response versus nonresponse to therapy. *Arthritis Res Ther* **7**, R746-755, doi:10.1186/ar1723 (2005).
42 Robinson, W. H. & Mao, R. Biomarkers to guide clinical therapeutics in rheumatology? *Curr Opin Rheumatol*, doi:10.1097/BOR.0000000000000250 (2015).
43 Foell, D. *et al.* Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* **303**, 1266-1273, doi:10.1001/jama.2010.375 (2010).
44 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, doi:10.1186/s13075-015-0723-1 (2015).
45 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, doi:10.1093/rheumatology/ket152 (2013).
46 Foell, D. *et al.* S100A12 (EN-RAGE) in monitoring Kawasaki disease. *Lancet* **361**, 1270-1272 (2003).
47 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 arthrtis . *Pediatric Rheumatology* **12 (Suppl 1)**, Y1 (2014).
48 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, doi:10.1007/s12014-010-9058-8 (2010).
49 Abulaban, K. M. & Brunner, H. I. Biomarkers for childhood-onset systemic lupus erythematosus. *Curr Rheumatol Rep* **17**, 471, doi:10.1007/s11926-014-0471-2 (2015).
50 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, doi:10.1097/BOR.0000000000000224 (2015).
51 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, doi:10.1002/acr.21728 (2012).
52 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, doi:10.1186/ar4600 (2014).
53 Masys, D. R., Harris, P. A., Fearn, P. A. & Kohane, I. S. Designing a public square for research computing. *Sci Transl Med* **4**, 149fs132, doi:10.1126/scitranslmed.3004032 (2012).
54 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, doi:10.1016/j.jbi.2008.08.010 (2009).
55 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, doi:10.1136/amiajnl-2011-000492 (2012).
56 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, doi:10.1136/amiajnl-2013-002577 (2015).
57 2020, E. C. H. The EU framework program for research and innovation. http://ec.europa.eu/programmes/horizon2020/ (2014).
58 Bravo, E. *et al.* Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA). *BMC Med* **13**, 33, doi:10.1186/s12916-015-0266-y (2015).
59 National Cancer Institute. Best Practices for Biospecimen Resources. http://biospecimens.cancer.gov/bestpractices/2011-NCIBestPractices.pdf (2011).
60 Budin-Ljosne, I. *et al.* ELSI challenges and strategies of national biobank infrastructures. *Norwegian Journal of epidemiology* **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**, i46-50, doi:10.1093/ije/dyn031 (2008).
62 Australasian Biospecimen Network. Repository Protocols, 4th Revision. http://abrn.net/what-we-do/protocols/ (2012).
63 (UCAN), U. C. A. N. Best practice guidelines for the collection, processing, transfer, storage and access to biologic data linking to clinical information in childhood arthritis. http://www.ucan-u.org (2015).
64 Norlin, L. *et al.* A Minimum Data Set for Sharing Biobank Samples, Information, and Data: MIABIS. *Biopreserv Biobank* **10**, 343-348, doi:10.1089/bio.2012.0003 (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, doi:10.1089/bio.2012.0026 (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, doi:10.1089/bio.2013.0061 (2013).
67 Knoppers, B. M. Framework for responsible sharing of genomic and health-related data. *The HUGO Journal* **8**, 3-9 (2014).
68 Tate, J. & Ward, G. Interferences in immunoassay. *Clin Biochem Rev* **25**, 105-120 (2004).
69 Mallone, R. *et al.* Isolation and preservation of peripheral blood mononuclear cells for analysis of islet antigen-reactive T cell responses: position statement of the T-Cell Workshop Committee of the Immunology of Diabetes Society. *Clin Exp Immunol* **163**, 33-49, doi:10.1111/j.1365-2249.2010.04272.x (2011).

70 Zaninotto, M., Mion, M., Altinier, S., Forni, M. & Plebani, M. Quality specifications for biochemical markers of myocardial injury. *Clin Chim Acta* **346**, 65-72, doi:10.1016/j.cccn.2004.02.035 (2004).
71 Meng, Q. H. & Krahn, J. Lithium heparinised blood-collection tubes give falsely low albumin results with an automated bromcresol green method in haemodialysis patients. *Clin Chem Lab Med* **46**, 396-400, doi:10.1515/CCLM.2008.079 (2008).
72 Lippi, G., Avanzini, P., Cosmai, M., Aloe, R. & Ernst, D. Incomplete filling of lithium heparin tubes affects the activity of creatine kinase and gamma-glutamyltransferase. *Br J Biomed Sci* **69**, 67-70 (2012).
73 Narayanan, S. The preanalytic phase. An important component of laboratory medicine. *Am J Clin Pathol* **113**, 429-452, doi:10.1309/C0NM-Q7R0-LL2E-B3UY (2000).
74 Bergqvist, Y., Eckerbom, S. & Funding, L. Effect of use of gel-barrier sampling tubes on determination of some antiepileptic drugs in serum. *Clin Chem* **30**, 465-466 (1984).
75 Daves, M., Trevisan, D. & Cemin, R. Different collection tubes in cardiac biomarkers detection. *J Clin Lab Anal* **22**, 391-394, doi:10.1002/jcla.20277 (2008).
76 Matheson, L. A., Duong, T. T., Rosenberg, A. M. & Yeung, R. S. Assessment of sample collection and storage methods for multicenter immunologic research in children. *J Immunol Methods* **339**, 82-89, doi:10.1016/j.jim.2008.08.003 (2008).
77 Webb, N. J., Myers, C. R., Watson, C. J., Bottomley, M. J. & Brenchley, P. E. Activated human neutrophils express vascular endothelial growth factor (VEGF). *Cytokine* **10**, 254-257, doi:10.1006/cyto.1997.0297 (1998).
78 Ayache, S., Panelli, M., Marincola, F. M. & Stroncek, D. F. Effects of storage time and exogenous protease inhibitors on plasma protein levels. *Am J Clin Pathol* **126**, 174-184, doi:10.1309/3WM7-XJ7R-D8BC-LNKX (2006).
79 Boyanton, B. L., Jr. & Blick, K. E. Stability studies of twenty-four analytes in human plasma and serum. *Clin Chem* **48**, 2242-2247 (2002).