Personalized medicine

Molecular profiling for advancing precision medicine in rheumatology in 2024

Coziana Ciurtin[1] and Elizabeth C Jury[1,†]

[1]Centre for Aging Rheumatology and Regenerative Medicine, University College London, UK

[†]Email: e.jury@ucl.ac.uk

Standfirst

Personalised molecular and immunological profiling for improved diagnosis, prognosis and treatment decisions remains an unmet need for patients with systemic autoimmune rheumatic diseases (ARDs) that are characterised by heterogeneous and overlapping clinical features and unpredictable disease course or response to treatment. Here we highlight three prominent publications in 2024, that have advanced the field.

Key advances

- Use of multiomic data from population-wide biobanks has the potential to accelerate the drive towards personalised medicine approaches for ARDs[1].
- Systemic autoimmune rheumatic diseases (ARDs) have distinct full spectrum genetic associations identifying shared serological and clinical phenotypes[2].
- Blood-based biomarkers can predict drug response across immune-mediated diseases[3].

Recent flow cytometry-based immune-phenotyping studies have substantiated the concept that ARDs can be grouped according to their immune profile regardless of disease classification[4,5]. One study identified five across-disease clusters in a cohort of >400 patients with 15 different autoimmune or autoinflammatory diseases[4], whereas the analysis of 1088 patients representing 11 ARDs identified multiple across-disease

groups with distinct immune and clinical phenotypes[5]. However, the development of robust biomarkers for personalised approaches is still lacking and is complicated by the complexity of high dimensional datasets and differences in cell-specific molecular profiles. This further requires validation both within and across established disease phenotypes, while accounting for age, sex and race or ethnicity[6].

Here we highlight three prominent publications in 2024, that have advanced the field by combining in-depth analysis of molecular or immune profiles coupled with novel computational analysis techniques to both stratify patients across traditional disease boundaries and provide an opportunity to discover novel biomarkers for improved prediction, diagnosis and prognosis of complex diseases. The use of large-scale biobanks such as the UK Biobank, which currently features data from 500,000 individuals, enables such research by making available omic data (including proteomics, metabolomics, genetic sequencing, imaging) with continuously updated electronic health records and standard blood test results[7].

The first 2024 study from Garg and colleagues highlights the power of the UK Biobank to drive predictive analysis using a powerful machine learning platform (MILTON) that enabled case-control studies across five ancestries[1]. This model identified disease-specific omic signatures from already diagnosed patients (using ICD10 codes) and predicted potential 'new' cases within the control cohort with relatively high predictive power for many disease phenotypes (AUC >0.7 for 1,596 ICD10 disease codes). This approach was able to predict new disease before onset and was able to outperform some known polygenic risk scores. The authors also performed a phenome-wide association using matched plasma proteomics data, which further improved prediction for some diseases (Δ AUC$\geq$0.1 for 52 diseases). These results were also validated using the FinnGen biobank. Not all disease phenotypes could be predicted using this model; this could be due to limitations in the ICD10 coding system or the need for more biomarker datasets. However, this approach might inform data collection in future biobanks, provide mechanistic insight into many diseases pathologies and have implications for future strategies for prevention and early detection of disease. Usefully, the open access MILTON platform enables researchers to explore information about ARDs using ICD10 codes.

Our second highlighted study investigated common molecular mechanisms in patients with systemic lupus erythematosus (SLE), primary Sjögren's, and myositis[2]. These conditions are characterised by overlapping clinical and serological features, with diagnostic and therapeutic implications, as well as high comorbid associations pertaining to individuals and families affected by more than one autoimmune condition, all suggestive of shared risk factors. The major histocompatibility complex (MHC) genes have been identified by genome wide association (GWAS), targeted genotyping studies

and meta-analyses, as the strongest single genetic cause of many systemic ARDs[8]. However, a converse approach, focused on investigating the impact of the full spectrum of genetic variation, including genes with lower allele frequency, is required to comprehensively evaluate genetic heritability. This strategy is important to counterbalance the limitations of GWAS analyses leading to the identification of common gene variants broadly spread across the genome with limited pathogenic or therapeutic implications, as advocated by the 'omnigenic model'[9]. Bianchi and colleagues[2] implemented this model using next-generation targeted DNA sequencing of regulatory and coding regions in a large cross-sectional cohort study including 2,292 individuals with SLE, primary Sjögren's and myositis, as well as 1,252 matched healthy individuals recruited from Scandinavian countries, specifically aiming to identify genetic profiles of individuals with shared clinical and laboratory features with implication for personalised medicine strategies[2].

The single variant analysis confirmed MHC, interferon pathway and reactive oxygen species metabolism, as the key genetic contributors to systemic ARDs. The gene based aggregate testing confirmed both MHC (*C2*, *HLA-C*, *MSH5*, *TNXB*) and non-MHC (*IRF5* and *YDJC*) associations, as well as reported associations with new genetic variants of *MAP3K6*, *SLC5A6* and *CGREF1*, which are all related to the type-I interferon pathway activation.

In addition to facilitating the discovery of genetic variants, this study identified common genetic associations of individuals with shared serological features across disease phenotypes. Rheumatoid factor, SSA-Ro52, SSA-Ro60, or SSB-La autoantibody positivity was strongly associated with MHC region variants, whereas dsDNA-specific autoantibody positivity was significantly associated with non-MHC genes, suggesting a role for environmental triggers on susceptible genetic backgrounds. Antinuclear antibody (ANA) positivity had weak associations with both MHC and non-MHC genes, supporting previous observations related to the molecular heterogeneity of ANA positive individuals. Furthermore, certain clinical features were linked with specific genetic profiles; arthritis was associated with genetic variants of protein kinase-C zeta and skin involvement was associated with dual specificity phosphatase, a finding confirmed at the functional level and externally validated in eczema and psoriasis studies.

Finally, in a third study, Gerassy-Vainberg and colleagues[3] used across-disease analysis looking for biomarkers of response to treatment. Immune features were profiled over time and in association with treatment responses to the tumour necrosis factor (TNF)-inhibitor infliximab within patients with inflammatory bowel disease (IBD).

High-dimensional data (whole blood RNA-sequencing, CYTOF and Luminex data) from patients with IBD responding or non-responding to infliximab were assessed at baseline, week-2 and week-14 post treatment. Variation in individual patient response

to drug over time was evaluated using a “Disruption Network" model, which indicated changes in immune and molecular regulation. By assessing the effect of every nonresponding patient to a pre-defined reference response network, the level of deviance from treatment response was calculated. This enabled disruptions in cell-specific functional modules/networks to be identified, that would otherwise not be detected using conventional analyses. Using this approach, cytoskeleton organization and VEGF-receptor signalling pathways in monocytes were most regulated in response to treatment at week-2. Furthermore, baseline monocytic expression of genes of the RAC1–PAK1 axis was predictive of infliximab response in patients with IBD and was validated in publicly available datasets from patients with rheumatoid arthritis. This approach supports earlier findings assessing secondary non-response to TNFα-inhibitors in a study using complex computational analysis of non-response due to development of anti-drug antibodies across disease phenotypes[10].

In conclusion, the three highlighted studies provide new evidence for the increased relevance of molecular diagnosis and its potential role in guiding patient diagnosis, stratification and targeted therapeutic approaches across established disease classification criteria in rheumatology. However, several challenges remain including validation of identified signatures in large cohorts, integrating molecular omic-signatures to provide a global picture of underlying pathogenesis in disease subsets and establishing robust and cost-effective tests for identified biomarkers that can be translated for routine use. An important goal is to improve the effectiveness of clinical trials in rheumatology by inclusion of patients based on their molecular signatures.

Reference:

1. Garg, M., Karpinski, M., Matelska, D., Middleton, L., Burren, O.S., Hu, F., Wheeler, E., Smith, K.R., Fabre, M.A., Mitchell, J., et al. (2024). Disease prediction with multi-omics and biomarkers empowers case–control genetic discoveries in the UK Biobank. Nature Genetics *56*, 1821-1831. 10.1038/s41588-024-01898-1.
2. Bianchi, M., Kozyrev, S.V., Notarnicola, A., Sandling, J.K., Pettersson, M., Leonard, D., Sjöwall, C., Gunnarsson, I., Rantapää-Dahlqvist, S., Bengtsson, A.A., et al. (2024). Unraveling the Genetics of Shared Clinical and Serological Manifestations in Patients With Systemic Inflammatory Autoimmune Diseases. Arthritis Rheumatol. 10.1002/art.42988.
3. Gerassy-Vainberg, S., Starosvetsky, E., Gaujoux, R., Blatt, A., Maimon, N., Gorelik, Y., Pressman, S., Alpert, A., Bar-Yoseph, H., Dubovik, T., et al. (2024). A personalized network framework reveals predictive axis of anti-TNF response across diseases. Cell Reports Medicine *5*, 101300. https://doi.org/10.1016/j.xcrm.2023.101300.
4. Tchitchek, N., Binvignat, M., Roux, A., Pitoiset, F., Dubois, J., Marguerit, G., Saadoun, D., Cacoub, P., Sellam, J., Berenbaum, F., et al. (2024). Deep immunophenotyping reveals

that autoimmune and autoinflammatory disorders are spread along two immunological axes capturing disease inflammation levels and types. Ann Rheum Dis *83*, 638-650. 10.1136/ard-2023-225179.
5. Tanaka, H., Okada, Y., Nakayamada, S., Miyazaki, Y., Sonehara, K., Namba, S., Honda, S., Shirai, Y., Yamamoto, K., Kubo, S., et al. (2024). Extracting immunological and clinical heterogeneity across autoimmune rheumatic diseases by cohort-wide immunophenotyping. Annals of the Rheumatic Diseases *83*, 242-252. 10.1136/ard-2023-224537.
6. Kalliolias, G.D., and Papavassiliou, A.G. (2024). Advancing precision rheumatology through tissue and blood profiling. Nat Rev Rheumatol *20*, 391-392. 10.1038/s41584-024-01115-7.
7. Allen, N.E., Lacey, B., Lawlor, D.A., Pell, J.P., Gallacher, J., Smeeth, L., Elliott, P., Matthews, P.M., Lyons, R.A., Whetton, A.D., et al. (2024). Prospective study design and data analysis in UK Biobank. Sci Transl Med *16*, eadf4428. 10.1126/scitranslmed.adf4428.
8. Harroud, A., and Hafler, D.A. (2023). Common genetic factors among autoimmune diseases. Science *380*, 485-490. 10.1126/science.adg2992.
9. Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell *169*, 1177-1186. 10.1016/j.cell.2017.05.038.
10. Hässler, S., Bachelet, D., Duhaze, J., Szely, N., Gleizes, A., Hacein-Bey Abina, S., Aktas, O., Auer, M., Avouac, J., Birchler, M., et al. (2020). Clinicogenomic factors of biotherapy immunogenicity in autoimmune disease: A prospective multicohort study of the ABIRISK consortium. PLoS Med *17*, e1003348. 10.1371/journal.pmed.1003348.

Competing interests

The authors declare no competing interests.

Figure 1: Approaches for personalised molecular and immunological profiling in autoimmune rheumatic diseases. (A) Using the longitudinal health record and omic data from participants in the UK Biobank and a new ensemble machine-learning framework (machine learning with phenotype associations, MILTON) a range of biomarkers to predict 3,213 diseases was developed and validated in the FinnGen biobank. (B) Targeted DNA sequencing of coding and regulatory regions in a large well-characterized ARD cohort of patients with SLE, primary Sjogren's, myositis, and healthy individuals combined with knowledge of autoantibody profiles and clinical features identified subgroups with unique genetic profiles. Common molecular mechanisms could underlie overlapping clinical manifestations and inform clinical heterogeneity. (C) Longitudinal high dimensional whole blood data from patients with inflammatory bowel disease (IBD) undergoing treatment with infliximab combined with a computational approach to examine individual response to therapy (Disruption Networks) was able to

identify cell-centred individual-level molecular networks, predictive of treatment response. This molecular pattern of response was validated in publicly available data from patients with rheumatoid arthritis (RA), suggesting common mechanisms of non-response to infliximab across disease classifications. ICD10, international classification of disease 10; WGS, whole genome sequencing; SLE, systemic lupus erythematosus.