RHEUMATOLOGY AND IMMUNOLOGY RESEARCH

Editorial • DOI: 10.2478/rir-2020-0008 • 1(1) • 2020 • 1–3

DE GRUYTER

Biomarkers for treatment response in rheumatoid arthritis: Where are they?

Sizheng Steven Zhao[1,2], Robert J. Moots[1,3,*]

[1]*Department of Rheumatology, Liverpool University Hospitals, Liverpool, UK*
[2]*Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, UK*
[3]*Faculty of Health, Social Care and Medicine, Edge Hill University, Liverpool, UK*

Received July 7, 2020 accepted July 30, 2020

The development of many new and effective drugs, and the parallel approach of a dynamic "treat to target" strategy, have transformed outcomes for patients with rheumatoid arthritis (RA). Yet, despite these incredible advances, too many patients fail to achieve an adequate response. The current management approach is, for almost all patients, disappointingly uniform: try a drug and, if/when it fails to provide adequate response, try another. Drugs are chosen by protocol, by "gestalt", but all too often by trial and error. The "personalized medicine" revolution promised decades ago, and renewed with the announcement of each new biomarker, has yet to make a meaningful impact on routine practice. Patients and clinicians are increasingly aware of the need to induce and maintain remission, prevent irreversible damage, minimize adverse events, and reduce cost to the individual and society. To achieve this, the optimal drug must be delivered to the appropriate patient at the earliest time. Are we any closer to finding useful predictors of treatment response in RA?

Biomarkers for treatment response can be derived from genes (e.g., polymorphism, epigenetics), serum (antibodies, cytokines), tissue (cellular blood components, synovium), or clinical phenotype (disability, imaging). Individually or combined, it is hoped that biomarkers can predict who will or who will not respond. Following are some examples. A well performed and validated study of inadequate response to methotrexate (MTX) found high DAS28 (we will return to this later), smoking, and abstinence from alcohol to be good clinical predictors; together, their area under the receiver operating curve was 0.75.[1] This dropped to 0.67 when externally validated—not so impressive when 0.5 indicates no better than chance. These results are typical of the generally weak predictive ability of clinical factors, including sex, age, and body mass index (BMI).[2] Indeed, one study suggested that the combination of several "established" clinical predictors account for <17% of the variability in response.[3]

From phenotype to the other end of the spectrum, genotype: these studies are often limited by small sample sizes and/or lack of replication/reproducibility. Of the better-known candidates (e.g., SLC19A1 or ATIC for methotrexate; CD84 or PDE3A–SLCO1C1 for tumor necrosis factor [TNF] inhibitors), none were sufficiently predictive to inform clinical decision making.[3,4] Research into transcriptional or epigenetic factors (that differ from cell to cell) followed with much promise. However, they pose additional challenges of identifying the optimal tissue to study and carry much greater cost.[3] Focusing on the primary site of pathology, biomarker studies using biopsies of the RA synovium have reported conflicting results,[5] partly because of their small sample sizes. Practicalities that limit study size also limit integration into routine practice, also were a useful biomarker to be found.

The only biomarker class to make its way into widespread clinical consciousness are serum proteins. Seropositivity for rheumatoid factor and/or anti-citrullinated protein antibodies is

Address for correspondence:
*Prof. Robert Moots, Department of Rheumatology, Liverpool University Hospitals NHS, Foundation Trust, Lower Lane, L9 7AL, Liverpool, UK.
E-mail: R.J.Moots@liverpool.ac.uk

Editorial • DOI: 10.2478/rir-2020-0008 • 1(1) • 2020 • 1–3

associated with rituximab response, driving some consensus statements to suggest using other biologic disease-modifying antirheumatic drugs (bDMARDs) in seronegative patients.[6] Elevated baseline C-reactive protein (CRP) has been reported to predict response to TNF inhibitors (TNFis) and interleukin (IL)-6 inhibitors.[2] To what extent do current biomarkers capture more than baseline disease activity? In a comprehensive study of tocilizumab response, genotypic, transcriptional, and serum biomarkers of its therapeutic target, IL-6, failed to robustly predict response after accounting for baseline DAS28.[7]

Why are we no closer? Here are some important points for biomarker hunters to consider. Most importantly, studies must carefully define what is being predicted. Is the definition of response some fixed level of low disease activity (e.g., DAS28 remission), fixed degree of response (reduction by >1.2 units of DAS28), something borrowed from randomized controlled trials (RCTs) (e.g., American College of Rheumatology 70% response, ACR70), or created with real-world practice in mind (where the baseline disease activity matters)? Any patient group with high disease activity at baseline will simultaneously be (1) more able to achieve a fixed or proportional reduction because there is more capacity for improvement and (2) less able to achieve a threshold for low disease activity because a greater absolute improvement is required. For example, high baseline DAS28 has been reported as a biomarker that simultaneously increase odds of ACR70 response and decreases odds of DAS remission.[8] Using outcomes from RCTs, where baseline disease activity is balanced, is inappropriate when it is not. Adjusting for baseline disease activity is not a robust solution when using binary outcomes.[9]

The majority of prediction studies focus on finding the optimal drug, but comparatively little is done on predicting non-response; ironic, when recommending against rituximab in seronegative patients is the nearest to personalized medicine we have got. There may be benefits in predicting non-response, since statistical properties are superior for common and more homogenous outcomes (i.e., there are fewer categorizations of non-response than response). However, dichotomizing to response versus non-response or using percentage-change further sacrifices statistical power that is already limited in many biomarker studies.[10] Using continuous absolute change in an outcome measure, with consideration for baseline differences between biomarker groups, is more valid and statistically efficient. We may even need to question the concept of response altogether, no matter how it is measured, since popular trial designs make inferences about efficacy for populations, not individuals.[10]

Studies should consider for whom the potential biomarker will be used and for what purpose. Research studies often recruit from specialist academic centers with preference for classification criteria, while biomarkers are more likely to be applied to heterogenous “physician diagnosed” populations in routine practice. Even a good biomarker will fail to be predictive in a non-represented population. It is also unlikely that genetic or serum biomarkers will predict response to bDMARDs of different class equally well; even within a class, such as TNFi, response may not be equally predictable due to differences in pharmacology.

Novel biomarker studies often report the predictive ability without comparison to the existing predictors. The expensive biomarker may simply be a marginally better representation of, e.g., crude markers of high disease activity or prior steroid use that are easily and inexpensively captured. Novel biomarker results are often not reproducible in other populations, partly due to inadequate analysis and/or validation. Outdated or inappropriate statistical methods, e.g., univariate screening or stepwise variable selection, remain prevalent.[11,12] Newer “machine learning” methods often overfit, making them impossible to replicate externally.[13] Many biomarkers will struggle to be implemented in clinical practice, due to cost or acceptability (e.g., synovial biopsies). An ideal biomarker candidate would be one that is already collected as part of routine practice, e.g., the ratio of neutrophils to lymphocytes.[14]

Despite the extensive and expensive search for biomarkers in the past decades, very little has changed to personalize RA treatment. This research area uniquely demands insight and knowledge across the spectrum of science, from genotype to phenotype and laboratory to epidemiological methods. Biomarkers are desperately needed, but repeating the same failed approach, over and over again, will unlikely get us any closer. However, far from recommending defeat, we challenge readers to take up this gauntlet with the conclusion from a recent piece on personalized medicine in *Nature*: “Realizing that the scope for personalized medicine might be smaller than we have assumed over the past 20 years will help us to concentrate our resources more carefully. Ironically, this could also help us to achieve our goals.”[10]

Conflict of Interest

None declared.

Editorial • DOI: 10.2478/rir-2020-0008 • 1(1) • 2020 • 1–3

References

[1] Teitsma XM, Jacobs JWG, Welsing PMJ, *et al*. Inadequate Response to Treat-to-Target Methotrexate Therapy in Patients With New-Onset Rheumatoid Arthritis: Development and Validation of Clinical Predictors. Ann Rheum Dis. 2018;77(9): 1261–1267.

[2] Wijbrandts CA, Tak PP. Prediction of Response to Targeted Treatment in Rheumatoid Arthritis. Mayo Clin Proc. 2017;92(7):1129–1143.

[3] Plant D, Wilson AG, Barton A. Genetic and Epigenetic Predictors of Responsiveness to Treatment in RA. Nat Rev Rheumatol. 2014;10(6):329–337.

[4] Cherlin S, Plant D, Taylor JC, *et al*. Prediction of Treatment Response in Rheumatoid Arthritis Patients Using Genome-wide SNP Data. Genet Epidemiol. 2018;42(8):754–771.

[5] Humby FC, Al Balushi F, Lliso G, *et al*. Can Synovial Pathobiology Integrate with Current Clinical and Imaging Prediction Models to Achieve Personalized Health Care in Rheumatoid Arthritis? Front Med [Internet]. 2017 May 3 [cited 2020 Jun 28];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5413506/

[6] Buch MH, Smolen JS, Betteridge N, *et al*. Updated Consensus Statement on the Use of Rituximab in Patients With Rheumatoid Arthritis. Ann Rheum Dis. 2011;70(6):909–920.

[7] Wang J, Platt A, Upmanyu R, *et al*. IL-6 Pathway-Driven Investigation of Response to IL-6 Receptor Inhibition in Rheumatoid Arthritis. BMJ Open. 2013;3(8):e003199.

[8] Kristensen LE, Kapetanovic MC, Gülfe A, *et al*. Predictors of Response to Anti-TNF Therapy According to ACR and EULAR Criteria in Patients With Established RA: Results From the South Swedish Arthritis Treatment Group Register. Rheumatol Oxf Engl. 2008;47(4):495–499.

[9] Zhao SS, Goodson NJ, Robertson S, *et al*. Smoking in Spondyloarthritis: Unravelling the Complexities. Rheumatol Oxf Engl. 2020;59(7):1472–1481.

[10] Senn S. Statistical Pitfalls of Personalized Medicine. Nature. 2018;563(7733):619–621.

[11] Harrell F. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis [Internet]. New York: Springer-Verlag; 2001 [cited 2020 Jun 12]. (Springer Series in Statistics.) Available from: https://www.springer.com/gp/book/9781441929181.

[12] VanderWeele TJ. Principles of Confounder Selection. Eur J Epidemiol. 2019;34(3):211–219.

[13] The Elements of Statistical Learning—Data Mining, Inference, and Prediction, Second Edition | Trevor Hastie | Springer [Internet]. 2018 [cited 2018 Nov 12]. Available from: https://www.springer.com/us/book/9780387848570

[14] Boulos D, Proudman SM, Metcalf RG, *et al*. The Neutrophil-Lymphocyte Ratio in Early Rheumatoid Arthritis and Its Ability to Predict Subsequent Failure of Triple Therapy. Semin Arthritis Rheum. 2019;49(3):373–376.