Personalized medicine in rheumatoid arthritis

Eektimmerman, F.

General discussion

Rheumatoid arthritis (RA) is an inflammatory disease of the joints that affects circa 1% of the Western population.[1] It is a progressive disease that results in joint damage and disability unless the inflammation is slowed or stopped by appropriate drug treatment. The treatment principle is to suppress the inflammation at an early stage of the disease with aggressive drug treatment to reach specified and sequentially measured goals, such as remission or low disease activity (so-called treat-to-target approach).[2] This approach is not only in RA a common concept but also in the field of other chronic diseases, including diabetes,[3] hypertension,[4] and hyperlipidemia.[5] Thus, the treatment of RA has a clear target and could combination therapies when monotherapy fails to achieve the goal.

Despite using such clear clinical endpoints in the treat-to-target approach, still, not all RA patients obtain an adequate effect to reduce the disease activity. Conventional drugs used in RA – the "disease-modifying antirheumatic drugs" (DMARDs) – reach their optimal effect after two to three months of treatment. Subsequently, it remains a hurdle for rheumatologists to identify the responders and non-responders beforehand, while patients can be treated with a DMARD that had an insufficient clinical response for several months. A variety of reasons could be conceived for the efficacy of DMARDs, but we hypothesize that pharmacogenetics (PGx) plays a substantial part in it. In essence, the patient's PGx could be determined prior to prescribing DMARDs, making that the drug treatment could be adjusted earlier instead of attempting multiple months of treatment without potential success.

Precision medicine – also known as personalized treatment – with the use of PGx is an evolving field in which the treatment is tailored to the individual patient. Precision medicine is already widely recognized and is for instance embedded in the daily practice of oncology and cardiology. One example of PGx testing in daily clinical practice is on CYP2C9/VKORC1 for the change of the maintenance dose of warfarin, a vitamin K antagonist that is used to inhibit the formation of coagulation factors II, VII, IX, and X, and protein C and S.[6] Normally, due to the narrow therapeutic index and the wide variability in individual dosing, frequent monitoring of the international normalized ratio over weeks is needed to determine the right dose. Guidelines from the Dutch Pharmacogenetics Working Group (DPWG) and the Clinical Pharmacogenetics Implementation Consortium (CPIC) can aid physicians to set patients earlier on the right dose, with determining PGx variants in multiple, such as *CYP2C9*, *CYP4F2*, and *VKORC1* combined with non-pharmacogenetic variants.[7–10]

Although a large number of studies investigates PGx in RA, still, not a single genetic marker has been implemented in daily clinical practice. Thus, the field of RA is still in the proof of principle phase, whereas numerous challenges must first be addressed before the implementation of PGx. Contradictory or non-convincing study results hamper the implementation, among others caused by lack of power (small sample size) or the use

of multi-ethnicities or different evaluation times. This general discussion highlights the obstacles to the implementation and the future directions in the field of PGx testing in RA.

Methodological and statistical aspects

Most PGx studies in RA were retrospective analyses, that were limited to patients with available outcomes data and sufficient material (blood, saliva) for genotyping. As a result, retrospective studies have some disadvantages over prospective studies. First, data from retrospective studies are frequently old, and therefore the study can lack important data which cannot be supplemented or has potential confounding factors. Second, it can be difficult to identify an appropriate exposed comparison group (for example controls in a case/control study) within the same study, and as a consequence, result in a small sample size. Third, differential losses to follow-up retrospective studies can introduce selection bias, thereby it may ensure that the included patients are not representative of the studied population. In the field of RA, in Chapter 2 we concluded that most PGx studies associated with methotrexate (MTX) efficacy had a small study group (<100 patients) with a small effect size (OR between 1.0 and 1.5) or lack the proper multivariate analysis that prevents association by confounders.

The second issue in pharmacogenetic testing is multiple testing, which refers to simultaneously investigate more than one hypothesis on the same study group of study subjects. The risk of multiple testing is that a significant difference is more often found based on coincidence when doing multiple tests. A correction for multiple testing maintains a stricter level of significance, but there is no firm rule whether you should correct it or not. In our thesis, we correct mostly for multiple testing by the simplest and most conservative approach using the Bonferroni correction. In Chapter 6 we performed a pathway analysis, whereas the SNPs within gene regions overlap each other. Consequently, Bonferroni correction is too stringent, because it is based on independent tests, and therefore a global p-value must be calculated using p-min, tail strength, and sequence kernel association test statistics. In most PGx rheumatology studies, no correction for multiple testing was applied.

Yet it is not justified to claim that studies that do not correct for multiple testing or consist of small sample sizes are wrong or even useless in science. Exploratory studies are essential, have a low threshold, and show a basis for new (pharmacogenetic) findings with more knowledge about the mechanism of action (efficacy or toxicity) of drugs. However, associations with good causality must be validated to confirm the result. After proper validation, the studies must be prospectively tested for both confirming the pharmacogenetic association as feasible in the clinical practice.

Hurdles of the translation into clinical practice

PGx has several hurdles to overcome for translation into clinical practice in the field of RA. The main problem is the paucity of available data that PGx testing demonstrates clinical improvement. Furthermore, a common assumption is that the findings of clinical improvement must be replicated in independent cohorts. Problems could be that there is no equivalent replication cohort or mostly the findings could not be replicated. For example, therapeutic drug monitoring (TDM) has been researched for a long time before it was embedded in clinical daily practice to adjust the drug dose according to the explored therapeutic window. Just like TDM, research must be done elaborately, and eventually – when there is sufficient evidence – PGx will be implemented in daily practice. One example that PGx is used in daily clinical practice is for the chemotherapeutic agent 5-fluorouracil (5-FU). Hereby, multiple studies showed that there is a relationship between different allelic variants in the DPYD gene (the gene that encodes DPD) and a deficiency in DPD activity. A deficiency of DPD activity leads to an increased risk for 5-FU toxicity, and therefore pre-emptive PGx testing for DPYD variants is performed.

After validation, there must be carrying capacity among clinicians to implement it into clinical practice. Different organizations associated with PGx – such as CPIC, Dutch Pharmacogenetics Working Group (DPWG), Pharmacogenomics Research Network, and Ubiquitous Pharmacogenomics – contribute to the implementation of PGx in clinical practice by establishing pharmacogenomic information, developing implementation tools, and also release public guidelines to implement PGx in clinical practice.

Phenotype definition and evaluation time

In RA, numerous criteria are used to determine the efficacy of drug therapy. For instance, we described in **Chapter 2** that the endpoints DAS(28), remission (DAS<1.6 or DAS28<2.6), low disease activity (DAS<2.4 or DAS28<3.2), EULAR response criteria, and ACR20, 50, 70 response criteria are often used to determine the efficacy of MTX. Granted that the majority of recent studies increasingly inclined towards the EULAR response criteria, still a few studies used different endpoints. For this reason, it is difficult to combine or compare the results directly. Attention is not only needed for the efficacy endpoint, but also the time of evaluation of those endpoints. In the case of MTX, most studies used the evaluation time points after three or six months of therapy, which properly reflects the effect of low dose MTX in RA.

Different studies investigate the effect of pharmacogenetics on the side effects or efficacy of MTX. A systematic review and meta-analysis showed that *RFC-1* 80G>A (*SLC19A*) rs1051266 is associated with the toxicity of MTX.[11] Our systematic review (Chapter 2) showed that this

SNP (rs1051266) among five other SNPs was associated with MTX efficacy, but still needs further validation.

For future research, it is preferable to take a uniform approach; with consistent criteria and time of evaluation. For example, as an efficacy endpoint the EULAR response criteria is suitable since it also corrects for DAS at baseline. Three or six months after the start of the treatment are appropriate choices of evaluation, whereas DMARDs are effective. For pharmacogenetics related to efficacy, simultaneously testing multiple SNPs seems more obvious than testing single SNPs, because DMARDs act on different pathways and combining SNPs can probably impact the response. For example, the associated six SNPs in Chapter 2 could be used to test if they together form a better prediction and associations on the efficacy of MTX.

Functional SNPs

To better understand if pharmacogenetic variants are associated with the efficacy or adverse events it is essential to know if those variants are functional SNPs that alter the function of a gene. However, there are at least 3.1 million SNP in the human genome, and most of them are not defined as (non-)functional and pragmatically these are extrapolated to assign an effect to a gene. Most common polymorphisms (MAF>5%) are potential regulatory polymorphisms located in 1) noncoding regions, including promoter/upstream, downstream and intron regions, that may affect transcription; 2) in intron and untranslated regions transcribed as RNA that may affect transcription, RNA splicing, stability or translation; or 3) in intergenic regions of unknown function.

Even if a SNP is functional it can have minimal impact on the alteration of a protein and lead to clinically unimportant changes.[12] In our studies, we tested individual SNPs that may have a (minimal) effect or no effect on the gene, but it could also be possible that a set of SNPs that form a haplotype could have a functional association of the efficacy or toxicity of drugs. However, to detect haplotype associations, other genotypic methods with sequencing data are needed.

Prediction models: trend or necessity?

Prediction models for DMARD treatment are developed with the purpose to support drug decision-making for rheumatologists. In recent years the number of publications on statistical models and decision models increased, but yet, these models are not clinically applied. One of the obstacles is that validation is a necessity before a model can be applied. In the developing phase, good models have internal validation. Further to internal validation,

prediction models must also be externally validated in another cohort in the same type of patients, preferably by other investigators. Unfortunately, most multivariate prediction models fail external validation due to poor study design, missing data, or weak-mediocre key performance. Also, not all important issues are reported, and therefore the TRIPOD statements have been introduced. The TRIPOD statements consist of the minimal details to report when developing or validating a multivariate diagnostic or prognostic prediction model.[13]

As mentioned earlier, (in)efficacy of drug response is probably multi genetic, and therefore a combination of different pharmacogenetic biomarkers could play a role and needs further investigation. This is also embedded in our prediction model, which consists of four different genes. Even though the prediction model, tested in **Chapter 5**, compromises four SNPs in four genes, those genes were included a decade ago, while there were only a limited potential known SNPs. Nowadays, there are more investigated SNPs known and it seems that other SNPs have more potential to link them between MTX efficacy and PGx.

Future perspectives on genetic testing

Both the candidate gene studies and GWAS are subject to the same artifacts of spurious association findings. GWAS relies on the indirect association to locate a pharmacogenetic-causing variant but only identifies putative candidate genes that still need a functional assay to determine the functioning of the active substance rather than just its PGx part. The direct candidate gene analysis relies on *a priori* hypothesis to identify a pharmacogenetic-causing variant by direct sequencing.

A novel method, next-generation sequencing (NGS) could be the future that will unravel complex disease genetics, like RA. NGS performs sequencing of millions of small DNA fragments in parallel. These fragments are mapped together with the individual reads to the human genome. Each of the three billion bases in the human genome is sequenced multiple times, providing accurate data and more insight into unexpected DNA variation. The advantage of NGS is that it will capture a broader spectrum of mutations than Sanger sequencing, is unselective, and is used to interrogate full genomes or exomes to discover entirely novel mutations and disease-causing genes, and could detect mosaic mutations.[14]

However, sequencing has the property that it results in huge data and being that, could lead to more spurious findings than GWAS or candidate gene studies. A better method seems to select genes from significant associations in a GWAS and sequence those genes and filter potential associations. This not only leads to a narrow, and more objective result, but is also more affordable.

Prospects towards personalized treatment in rheumatoid arthritis

In the last decade, great progress has been made in the clinical management of RA to achieve low disease activity or remission (so-called treat-to-target principle).[2] Thanks to the treat-to-target approach patients are earlier onset on an effective DMARD and ultimately had less joint damage. Additionally, the introduction of the new drugs (TNF-and JAK-inhibitors) ensures that there is an ample choice in the treatment of RA and offers a solution when the conventional DMARDs had an insufficient clinical effect. Despite those developments, it remains the question of the field of rheumatology has still engrossment about genetic testing.

In this thesis, four SNPs were associated with the efficacy of adalimumab (**Chapter 6**) but need additional replication to validate those findings. In **Chapter 1** we found in the literature six potential SNPs associated with the efficacy of MTX, but five of them did not have any replication studies that could confirm the results. Also, the prediction model for MTX monotherapy (Chapter 4) seems useful, but still, nowadays most of the included non-genetic variants are taken into consideration for drug decision making and the four pharmacogenetic variants in the prediction model showed a small contribution to its total effect.

Up to now, no genetic variants have yet been robustly and consistently associated with response to DMARD use in RA. Also, the results of candidate gene studies, including ours, had led to conflicting results with margin effect sizes. Given the fact that MTX acts on various biological pathways, it is more likely that multiple genes are related to the efficacy and therefore a combination of multiple genes seems more logical. Genetic studies, that tested single SNPs, are not sufficient to unravel complex immunological diseases like RA or multi-target drug treatment as MTX. Therefore, future studies must focus more on a combination of multiple SNPs (haplotype), eventually in combination with other non-genetic factors. One method to take this into account is by using a polygenetic risk score (PRS). A PRS summarises the estimated effect of multiple genetic variants on an individual's phenotype, calculated as a weighted sum of trait-associated alleles. A practical example of the application of PRS is in the prediction of subtype-specified breast cancer, which was based on a large GWAS dataset.[15]

Future research should consider the potential effects of combining results from GWA studies with sequencing data, so the discovery of genetic variants will be accelerated and could ultimately lead to the implementation of pharmacogenetics in RA patients. Also, PRS could be used to improve the predictive value of the efficacy or toxicology of DMARDs and thus help to improve stratification for the screening of suitable DMARDs.

REFERENCES

1. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. Lancet 376, 1094–1108 (2010).
2. Solomon DH, Bitton A, Katz JN *et al*. Review: Treat to target in rheumatoid arthritis: Fact, fiction, or hypothesis? Arthritis Rheumatol. 66, 775–782 (2014).
3. Garber AJ. Treat-to-target trials: Uses, interpretation and review of concepts. Diabetes Obes. Metab. 16, 193–205 (2014).
4. Staessen JA, Wang J, Thijs L. Cardiovascular protection and blood pressure reduction: a meta-analysis. Lancet 358, 1305–1315 (2001).
5. Graham I, Atar D, Borch-Johnsen K *et al*. European guidelines on cardiovascular disease prevention in clinical practice: Executive summary - Fourth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (Constituted by Eur. Heart J. 28, 2375–2414 (2007).
6. Hirsh J, Dalen JE, Deykin D *et al*. Oral anticoagulants: mechanism of action, clinical effectiveness, and optimal therapeutic range. Chest 119, 8S–21S (2001).
7. Johnson JA, Gong L, Whirl-Carrillo M *et al*. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. Clin. Pharmacol. Ther. 90, 625–629 (2011).
8. Swen JJ, Huizinga TW, Gelderblom H *et al*. Translating pharmacogenomics: Challenges on the road to the clinic. PLoS Med. 4, 1317–1324 (2007).
9. Swen JJ, Nijenhuis M, de Boer A *et al*. Pharmacogenetics: From bench to byte an update of guidelines. Clin. Pharmacol. Ther. 89, 662–673 (2011).
10. Swen JJ, Wilting I, de Goede AL *et al*. Pharmacogenetics: From bench to byte. Clin. Pharmacol. Ther. 83, 781–787 (2008).
11. Qiu Q, Huang J, Shu X *et al*. Polymorphisms and Pharmacogenomics for the Clinical Efficacy of Methotrexate in Patients with Rheumatoid Arthritis: A Systematic Review and Meta-analysis. Sci. Rep. 7, 44015 (2017).
12. Albert PR. What is a functional genetic polymorphism? Defining classes of functionality. J. Psychiatry Neurosci. 36, 363–365 (2011).
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 350, g7594 (2015).
14. Behjati S, Tarpey PS. What is next generation sequencing? Arch. Dis. Child. Educ. Pract. Ed. 98, 236–238 (2013).
15. Mavaddat N, Michailidou K, Dennis J *et al*. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. Am. J. Hum. Genet. 104, 21–34 (2019).