Willem J.J. Falkenburg*, Helen J. von Richthofen, Jana Koers, Cas Weykamp,
Marco W.J. Schreurs, Liesbeth E. Bakker-Jonges, Inez-Anne Haagen, Willem F. Lems,
Dörte Hamann, Dirkjan van Schaardenburg and Theo Rispens*

Clinically relevant discrepancies between different rheumatoid factor assays

Abstract

Background: Accurate measurements of rheumatoid factors (RFs), autoantibodies binding IgG, are important for diagnosing rheumatoid arthritis (RA) and for predicting disease course. Worldwide, various RF assays are being used that differ in technique and target antigens. We studied whether assay choice leads to clinically important discrepancies in RF status and level.

Methods: RF measurements using four commercial RF assays were compared in 32 RF$^+$ samples. Using enzyme-linked immunosorbent assays (ELISAs), the influence of the target antigen source – human IgG (hIgG) versus rabbit IgG (rIgG) – on measured RF levels was investigated in arthralgia patients and RA patients.

Results: Substantial discrepancies were found between RF levels measured in the four commercial assays. Six samples (19%) with RF levels below or slightly above the cutoff in the rIgG-based Phadia assay were RF$^+$ in three assays using hIgG as the target antigen, some with very high levels. Direct ELISA comparisons of RF reactivity against hIgG and rIgG estimated that among 173 ACPA$^+$ arthralgia patients, originally RF negative in rIgG-based assays, up to 10% were single positive against hIgG. Monoclonal RFs binding to hIgG and rIgG or hIgG only supported these findings. In a cohort of 69 early RA patients, virtually all RF responses reacted with both targets, although levels were still variable.

Conclusions: The use of RF assays that differ in technique and target antigen, together with the different specificities of RF responses, leads to discrepancies in RF status and levels. This has important consequences for patient care if RA diagnosis and disease progression assessments are based on RF test results.

Keywords: arthralgia; assays; rabbit IgG; rheumatoid arthritis; rheumatoid factor; target antigen.

Introduction

Rheumatoid factors (RFs) are autoantibodies that bind to epitopes within the constant region (Fc) of IgG. The first evidence for their existence was found in 1937 when sera from patients with rheumatoid arthritis (RA) were shown to cause agglutination of sheep red blood cells opsonized with rabbit IgG (rIgG) [1, 2]. Presently, both RF status and level are part of the classification criteria for RA [3], and while the more recently discovered anti-citrullinated protein antibodies (ACPAs) have a higher specificity for RA [4, 5], RFs are still considered to have value in predicting development of disease in patients at-risk for RA and predicting disease course in RA patients [6–9].

*Corresponding authors: Willem J.J. Falkenburg,** Amsterdam Rheumatology and Immunology Center, Reade, Doctor Jan van Breemenstraat 2, 1056 AB Amsterdam, The Netherlands, E-mail: w.falkenburg@sanquin.nl. http://orcid.org/0000-0002-6739-8581; and Department of Immunopathology, Sanquin Research, PO Box 9892, 1006 AN Amsterdam, The Netherlands; and **Theo Rispens,** Department of Immunopathology, Sanquin Research and Landsteiner Laboratory, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, E-mail: T.Rispens@sanquin.nl
Helen J. von Richthofen and Jana Koers: Department of Immunopathology, Sanquin Research, Amsterdam, The Netherlands
Cas Weykamp: Queen Beatrix Hospital, Winterswijk, The Netherlands
Marco W.J. Schreurs: Department of Immunology, Laboratory Medical Immunology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Liesbeth E. Bakker-Jonges: Medical Laboratories, Department of Immunology, Reinier de Graaf Groep, Delft, The Netherlands
Inez-Anne Haagen: Department of Hematology and Clinical Chemistry Laboratories, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands
Willem F. Lems: Amsterdam Rheumatology and Immunology Center, Reade, Amsterdam, The Netherlands; and Amsterdam Rheumatology and Immunology Center, VU University Medical Center, Amsterdam, The Netherlands
Dörte Hamann: Department of Immunopathology and Blood Coagulation, Sanquin Diagnostic Services, Amsterdam, The Netherlands
Dirkjan van Schaardenburg: Amsterdam Rheumatology and Immunology Center, Reade, Amsterdam, The Netherlands; and Amsterdam Rheumatology and Immunology Center, Academic Medical Center, Amsterdam, The Netherlands

Table 1: Characteristics of the RF assays.

Name and technique of RF assay	Manufacturer	Antigen	RF Istoype
EliA™ RF IgM: Fluoroenzymimmunoassay	Phadia	Aggregated rabbit IgG	IgM
RF II Tina-quant: Immunoturbidimetry	Roche	Human IgG-coated latex particles	Total
AUTOSTAT II: ELISA	HYCOR	Human IgG	IgM
Rate nephelometry	Beckman Coulter	Particle bound human IgG	Total
Sanquin IgM RF ELISA	Sanquin	Human IgG	IgM

Many different assays are used in the clinic to detect RF. Nephelometry and turbidimetry assays are based on the originally observed phenomenon of RF-induced agglutination of IgG-coupled particles [10]. Enzyme-linked immunosorbent assays (ELISAs), including manual or automatic assays such as the widely used EliA™ system (Phadia), use isotype-specific secondary antibodies to detect RF bound to coated IgG [11]. The use of these different assays for detection and quantification of RF introduces potential sources of variability at multiple levels.

First, although ELISA-based assays specifically detect one RF isotype (IgM-, IgA- or IgG-RF), the exact contribution of the different RF isotypes to the agglutination measured in nephelometry and turbidimetry assays is unclear. Most likely, IgM-RFs are primarily responsible for the agglutination, as their polyvalent penta/hexameric structure makes them superior to IgA- and IgG-RF in crosslinking IgG-Fcs [12]. It is even possible that IgG-RF and/or IgA-RF could have an inhibitory effect on agglutination by competing with IgM-RF for IgG-Fc binding sites.

Second, different assays use different sources of IgG as the target antigen to which the RFs bind. Some assays, including the Phadia EliA™ system, use rIgG, analogous to the experiments in which RF was first discovered. Other assays use human IgG (hIgG) or human Fc domains. Although rIgG contains a histidine residue at amino acid (aa) position 435 that is critical in the "Ga epitope", which is an important binding site for RFs [13, 14], homology between rabbit and human CH2-CH3 domains is only 74%.

A third factor likely to cause intertest variability is the polyclonal and polyspecific nature of the RF response. Although some RF responses seem to have restricted reactivity to one epitope on IgG-Fc, others are composed of various RF clones specific for different epitopes [15, 16].

Finally, the concentration of the target IgG-antigen used may influence measurement of RF levels. IgM-RFs are generally considered to be of low affinity compared to IgG antibodies against recall antigens [17, 18]. Their binding depends on making a polyvalent interaction with multiple target IgGs using multiple of their 10 (pentameric IgM) or 12 (hexameric IgM) antigen binding domains. At low concentrations of target IgG, with the IgG-Fcs too far apart to facilitate sufficiently multivalent interaction, only the higher affinity fraction of an RF response may bind. Two sera with the same amount of RF but different RF affinities may show the same level of RF in a test with a high target IgG density and yet significantly differ in a test with a low target density.

Because RF status and RF level can have clinical consequences for diagnosing RA, predicting disease course and treatment determination [19, 20], it is essential to determine the extent of intertest variability between the different RF assays used in the clinic and to understand the causes of this variability. Here we analyze data from the Dutch Foundation for Quality Assessment in Medical Laboratories (SKML) comparing RF measurements between four internationally used RF assays each differing in technique, isotypes measured and target-IgG-antigen used (Table 1). Furthermore, we examined in different clinical cohorts the occurrence of RF responses specific for hIgG that are potentially missed by RF assays using rIgG as the target antigen.

Materials and methods

RF measurements by commercial assays: analysis of variation

For analysis of variation in RF measurements between different commercially available assays, 32 RF+ samples were selected from leftover sera that had been sent to Sanquin Diagnostics Services for RF testing. This panel included a wide range of RF levels as determined by an in-house IgM-RF ELISA: 19–837 IU/mL. For the in-house IgM-RF ELISA, 96-well flat-bottom plates coated with 25 μg/mL hIgG were used. Serum samples were put at 37 °C for 30 min, diluted in phosphate-buffered saline (PBS) supplemented with 0.1% Tween-20 and incubated after washing the plates 4× 100 μL/well for 30 min (shaking) at room temperature. After washing, IgM-RF was detected by incubating the wells for 30 min with 100 μL horseradish peroxidase (HRP)-conjugated mouse monoclonal antihuman IgM (μ-chain-specific) antibodies diluted 1:1500 (0.5 mg/mL, MH15; Sanquin) and visualized with Uptima TMB ELISA peroxidase substrate (Interchim)

diluted 1:1 with distilled water. The reaction was stopped with 2 M H_2SO_4, and the optical density (OD) was read at 450 nm. IgM-RF levels were calculated against a calibration curve of a reference sample (Reference Laboratory for Rheumatologic Serology [RELARES]) with a defined IgM-RF level of 200 IU/mL [21].

The 32 selected samples were aliquoted and sent to laboratories in the Netherlands that use different assays for RF testing (Table 1). All samples were tested in triplo on the Phadia EliA™ system and with HYCOR ELISAs at the Erasmus MC University Medical Center in Rotterdam, with immunoturbidimetry (Roche RF II Tina-quant) at the Onze Lieve Vrouwe Gasthuis in Amsterdam and with rate nephelometry (Beckman Coulter) at the Queen Beatrix Hospital in Winterswijk.

Direct comparison of RF reactivity against hIgG and rIgG in ELISA

Nunc MaxiSorp 96-well flat-bottom plates (Thermo Scientific) were used for all ELISAs. Target antibodies – hIgG and rIgG – were diluted in PBS to 15 or 25 µg/mL, or 1 µg/mL for the experiments with monoclonal RFs, and coated overnight at 4 °C. Polyclonal hIgG was obtained from Intravenous immunoglobulin (IVIG, Nanogam, Sanquin). Polyclonal rIgG was purified from rabbit plasma using protein G affinity chromatography (HiTrap Prot G HP; GE Healthcare Life Sciences). After overnight coating, plates were washed 5× with 0.02% Tween 20-PBS. All subsequent washing steps were identical. One hundred microliters of serum samples, controls, monoclonal IgM-RFs or reference serum diluted in 0.1% Tween 20-PBS was added to the wells and incubated for 30 min, shaking, at room temperature. After washing, IgM-RF was detected by incubating the wells for 30 min with MH-15 and visualized with 3,3′,5,5′-tetramethylbenzidine (100 µg/mL) in 0.11 M acetate buffer, pH 5.5, containing 0.003% H_2O_2 (Merck). The reaction was stopped with 2 M H_2SO_4, and OD was read at 450 nm and 540 nm for background correction using a BioTek microtiter plate reader. IgM-RF levels were calculated against a titration curve of the RELARES reference sample. For the hIgG-based ELISAs, which are performed similarly to the Sanquin IgM-RF ELISA, the same cutoff of 19 IU/mL was used. Testing 54 healthy donor samples showed that using the same cutoff for the rIgG-based ELISAs results in the same percentage of positive samples (3.7%) as for the hIgG-based ELISAs.

Serum samples used in the ELISAs

Three sets of serum samples were used in the ELISAs. The first set consisted of 51 leftover samples from sera that had previously tested RF+ using the in-house IgM-RF ELISA at Sanquin Diagnostic Services. This set excluded but was comparable to the set of 32 samples used for the comparison between the commercial assays. The second set of serum samples consisted of 173 ACPA+ arthralgia patients who originally tested RFneg. in an in-house IgM-RF ELISA or the Phadia EliA, both of which use rIgG as the target antigen. These samples were selected from a previously described [6] cohort of ACPA+ and/or IgM-RF+ patients with (a history of) arthralgia recruited since 2004 at the Jan van Breemen Research Institute, Reade (Amsterdam, The Netherlands). The third set consisted of 69 serum samples from disease-modifying antirheumatic drug (DMARD)-naïve RA patients based in the Amsterdam region. These patients fulfilled the 1987 RA classification criteria of the American College for Rheumatism (ACR) [22]. For all RA samples, the original RF assay was an in-house IgM-RF ELISA with rabbit target-IgG. Both the hIgG-based original assay in the diagnostics cohort and the rIgG-based original assays in the patient cohorts introduce a selection bias. Samples with RF specific for rIgG but not hIgG would not have been included in the diagnostics cohort; ACPAneg. patients with RF specific for hIgG but not rIgG would not have been included in the arthralgia cohort and would have had to score more points on other classification criteria to be diagnosed with RA and included in the RA cohort.

Ethics approval: Arthralgia patients and RA patients signed informed consent forms for use of their serum samples. No informed consent was obtained for the sera from Sanquin Diagnostic Services because these were the leftover samples from sera obtained for routine diagnostic purposes. These materials were used anonymously without any connection to clinical or patient-specific data. The study complies with the World Medical Association Declaration of Helsinki regarding ethical conduct of research involving human subjects.

Monoclonal IgM-RF: Two monoclonal IgM rheumatoid factors (mRFs), RF61 [17, 23] and RF-AN [24], were produced as described previously [18].

Results

Comparing RF levels determined with four commercial assays

To analyze intertest variability of RF measurements, RF levels were determined in 32 serum samples with four commercially available techniques (Table 1) and plotted in xy graphs and Bland-Altman plots (Figure 1). All samples had previously been determined RF+ in an in-house IgM-RF ELISA at Sanquin Diagnostic Services, and the panel was selected to encompass a wide range of RF levels (19–837 IU/mL, median 120 IU/mL). It is immediately apparent that some samples are RFneg. in the Phadia assay but have high RF levels in the other three commercial assays and the Sanquin ELISA (Figure 1A, upper panels). The Table in Figure 1 shows that for six samples (18%), all four assays that use hIgG as the target antigen detected RF, also at high levels, whereas the rIgG-based Phadia detected levels below or just above its cutoff point. This suggests that some RF+ samples have an RF response that is specific for hIgG and cannot be detected in an assay that uses rIgG as the target antigen. The Bland-Altman plots in the lower panels show that apart from these Phadianeg. samples also many other samples had much higher levels in the other four assays. This is especially true for those with relatively low average levels, which was partly compensated by the lower cutoff value for positivity in the Phadia assay.
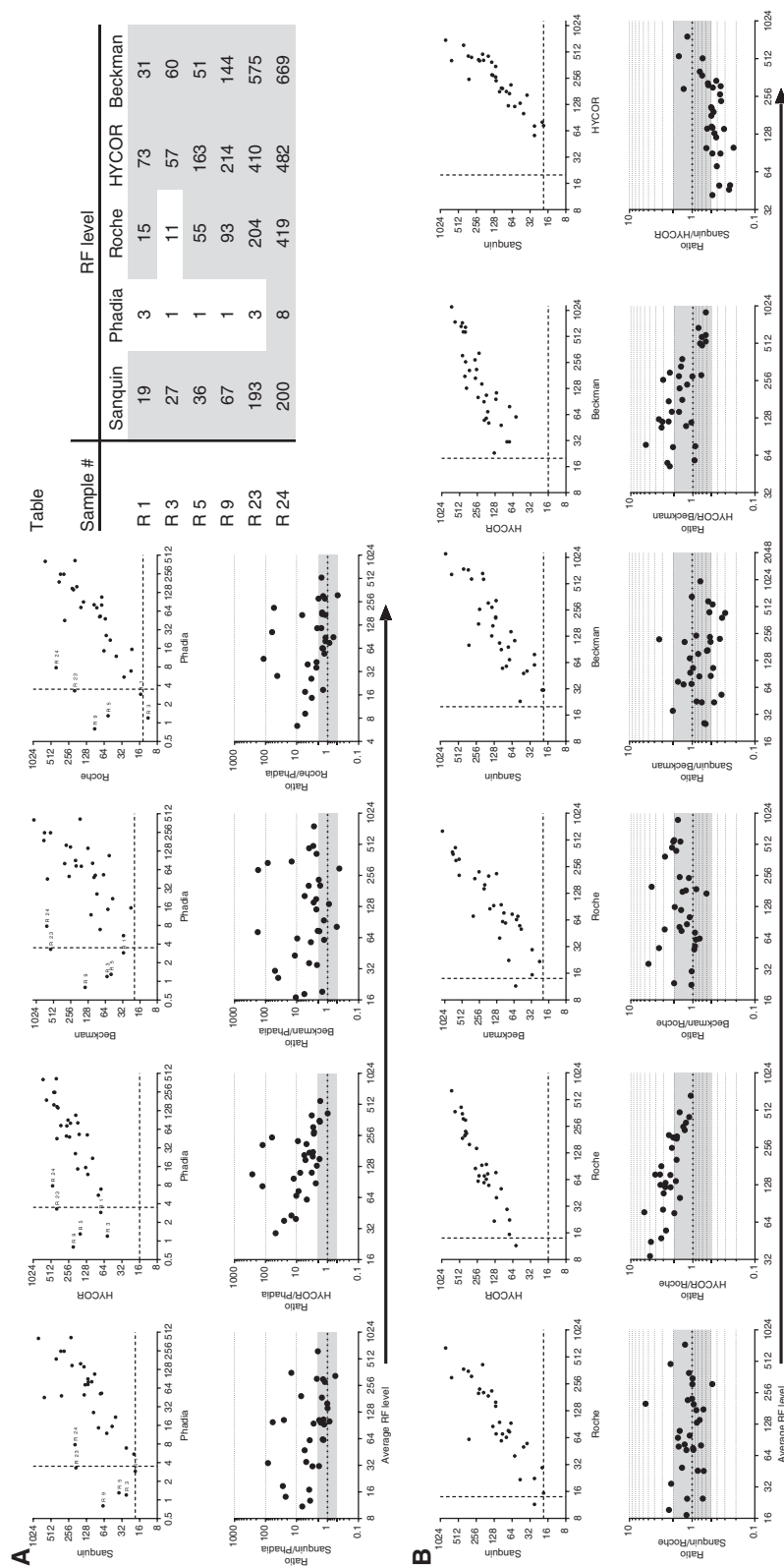
Table	RF level					
Sample #	Sanquin	Phadia	Roche	HYCOR	Beckman	
R 1	19	3	15	73	31	
R 3	27	1	11	57	60	
R 5	36	1	55	163	51	
R 9	67	1	93	214	144	
R 23	193	3	204	410	575	
R 24	200	8	419	482	669	

Figure 1: Comparisons of RF levels measured with four commercial RF assays.

(A–B upper panels) RF levels determined in 32 RF+ samples with four commercial RF assays and the original in-house Sanquin IgM-RF ELISA, plotted in x-y graphs. One sample with an extremely high level in the Roche assay was excluded from the graphs. Samples below the dotted lines are considered RF negative in the respective assays. (A–B lower panels) Bland-Altman plots of the same data as in the upper panels. Samples outside the gray area showed a more than twofold difference in RF level between two assays. Table: RF levels for six samples with a large discrepancy between Phadia and the other four assays. Levels with white background are below the cutoff value of the respective assay.

Although there appears to be a better correlation between RF levels measured with the different hIgG-based assays (Figure 1B), there was still substantial variation. No samples were found to be both positive in the assays detecting "total RF" (Beckman and Roche) and negative in both commercial assays that specifically detect IgM-RF (Phadia and HYCOR). This was expected, as the samples were selected from a panel of sera previously determined RF[+] in the Sanquin ELISA that measures IgM-RF. Because it is likely that not all tested samples also had IgA- and/ or IgG-RF, the fact that there was only one sample RF[neg.] in the Roche assay and no samples in the Beckman assay suggests that, in these turbidimetry and nephelometry assays, the IgM isotype is sufficient for causing the agglutination.

To test whether the discrepancies between the results from the different assays would also be found when pooled samples are used to compare assays, the 32 samples tested in Figure 1 were divided into quartiles based on their RF level, pooled and tested in the four commercial RF assays. Figure 2 shows that when comparing pooled samples, the discrepancies seen for the individual samples level out, seemingly improving agreement between the assays.

Comparison of RF reactivities to hIgG and rIgG

To further investigate the possibility that the RF response has a strong preferential binding to hIgG over rIgG in certain individuals, a direct comparison of RF reactivity against the two targets was performed for the additional 51 RF[+] samples sent to Sanquin Diagnostic Services for RF testing (the diagnostics cohort). The majority of these samples showed comparable reactivity against both targets in ELISA (Figure 3A). Some samples, especially among those with lower RF levels, bound much better to hIgG than rIgG. Conversely, there were no samples with substantially higher reactivity against rIgG. Because we do not have access to the clinical data of this cohort, we cannot determine which samples came from RA patients.

To determine if hIgG- or rIgG-specific RF responses also influence RF status and/or levels in better defined (at risk for-) RA cohorts, the same direct ELISA comparison of RF reactivity against hIgG and rIgG was performed in two patient cohorts. The first cohort consisted of patients with (a history of) arthralgia, i.e. joint pain but no clinical arthritis, tested positive for ACPA and/or IgM-RF originally determined in rIgG-based assays. From this cohort 173 ACPA[+] RF[neg.] patients were selected and re-tested for RF reactivity against hIgG and rIgG. Figure 3B shows that reactivity against hIgG was higher than what would be expected for this patient group, which had originally tested RF[neg.]. Although most of the samples with reactivity above the cutoff for positivity in the ELISA with hIgG targets (x-axis) also showed reactivity against rIgG (y-axis), a considerable proportion of the total cohort – 18/173 (10.4%) – 42% of the anti-hIgG RF[+] samples, showed substantially higher (>2×) reactivity against hIgG than against rIgG. This is not only true for samples with very low levels, where small absolute differences translate into large ratio differences in the Bland-Altman plots (Figure 3B, lower panel), but also for samples with anti-hIgG RF levels well above the cutoff. This suggests that these samples may have been incorrectly classified as RF[neg.] because their RF response is specific for hIgG. Indeed, if we use the 19 U/mL cutoff also for the rIgG-based ELISAs, 10% of samples (18/173) would be classified anti-hIgG positive anti-rIgG negative.

Next, we analyzed RF reactivities in 69 DMARD-naïve patients with active early RA, either RF[neg.] or RF[+] in their initial rIgG-based RF assay. In contrast to the diagnostics
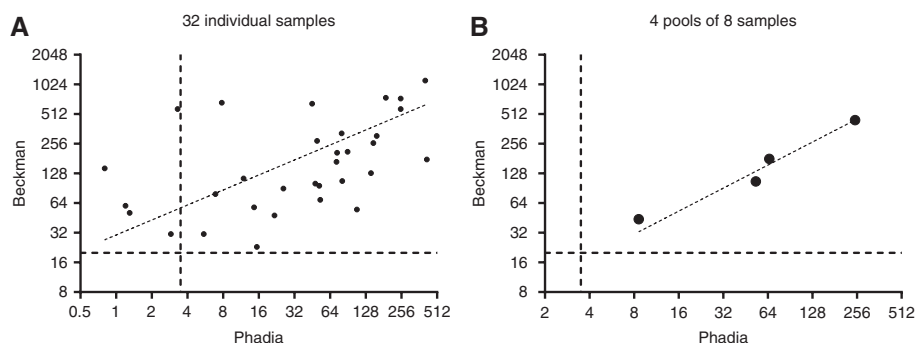

Figure 2: Discrepancy between assays for individual versus pooled samples.
The 32 samples tested in the four commercial RF assays – representative example in (A) – were divided into quartiles based on their RF level, pooled and also tested in the four RF assays – representative example in (B). For reference, a straight (dotted) line was fitted through the data points with nonlinear regression.
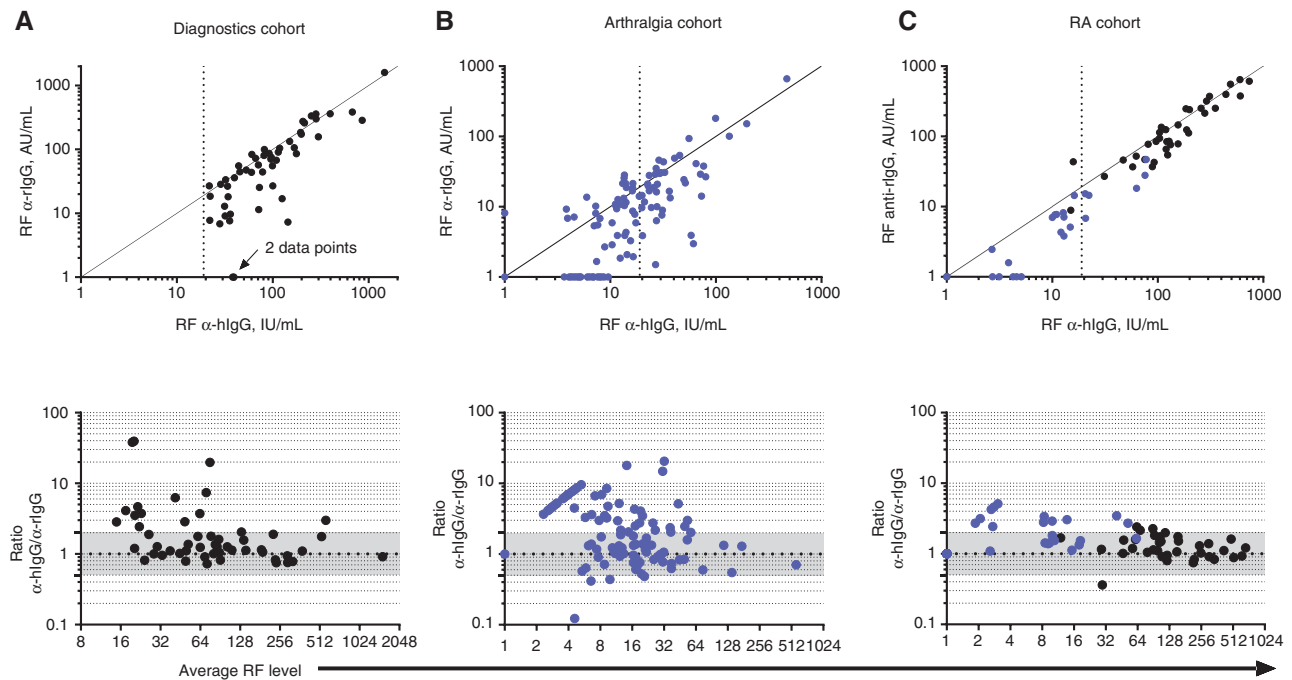
Figure 3: Direct comparison of IgM-RF reactivity against rabbit IgG and human IgG in three cohorts.
ELISAs with rabbit IgG or human IgG as coated target antigen were used to directly compare IgM-RF reactivity against these targets in three cohorts: (A) 51 samples sent to Sanquin Diagnostic Services for RF testing, all originally tested RF+ in a human IgG-based IgM-RF ELISA (diagnostics cohort). (B) 173 ACPA+ patients with (a history of) arthralgia, originally tested RFneg. in rabbit IgG-based IgM-RF assays. (C) 69 DMARD-naïve RA patients, originally RF+ or RFneg. in rabbit IgG-based assays, with active early RA diagnosed according to the 1987 ACR criteria (RA cohort). For every cohort samples tested positive in the original (cohort-specific) RF assays are represented by black dots, negative samples by blue dots. AU, arbitrary units; IU, international units; rIgG, rabbit IgG; hIgG, human IgG. Straight lines: x = y. Vertical dotted lines: cutoff of hIgG-based IgM-RF ELISA.

cohort and the arthralgia cohort, few RA patients showed a large discrepancy between RF reactivity against hIgG and rIgG (Figure 3C), only one or two originally classified RFneg. samples shows high anti-hIgG reactivity with low anti-rIgG reactivity and one RF+-assigned sample shows the opposite pattern.

To further substantiate possible hIgG specificity of RFs, we tested target-specific binding of two recombinantly produced monoclonal IgM-RFs originally isolated from two RA patients [23–26]. Monoclonal RF-AN was reported to bind an epitope at the CH2-CH3 interface of IgG, with an important role for His435 [24, 27], which is identical in hIgG and rIgG. Regarding the contact residues between RF-AN and hIgG1, there is just one amino acid difference with rIgG, within the 15 aa comprising the epitope. By contrast, RF61 binds hIgG at epitopes in the CH3
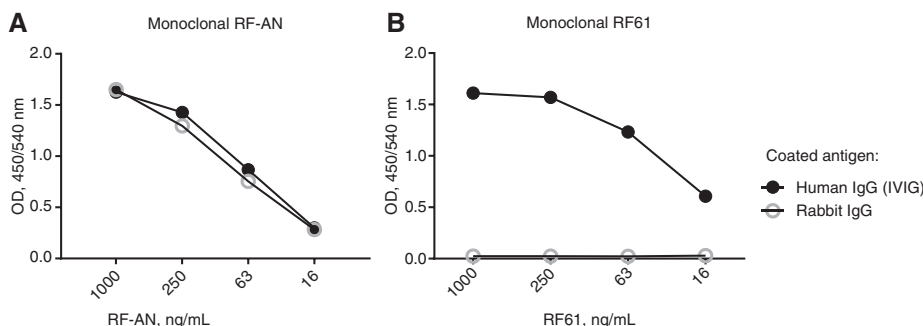

Figure 4: Binding profiles of two monoclonal IgM-RFs.
Binding to rIgG and hIgG was analyzed in ELISAs for two recombinantly produced monoclonal IgM-RFs originally isolated from RA patients, RF-AN (A) and RF61 (B) at different concentrations of mRF. OD, optical density.

domains, close to the C-terminus [17]. Arg355 is crucial for RF binding, and although conserved in rIgG, there are four residue differences at the interaction interface with the IgG-CH3 domains out of a total of 13 Fc contact residues. As illustrated in Figure 4, monoclonal RF-AN binds hIgG and rIgG equally well, whereas monoclonal RF61 binds to hIgG but not to rIgG. Although it is unknown what percentage of the RF response comprises these individual RF clones, the results show that RFs binding both targets or only hIgG can be present in RA patients.

Discussion

RF status and level are important factors in the most recent classification criteria for RA [3]. The presence of RF is a predictor for the development of RA in at-risk individuals, signals a more severe disease course in RA patients and can influence treatment decisions. Here we show that there can be large discrepancies in measured levels between commercial RF assays and even disagreement on whether individual samples are RF positive. The choice of the target antigen, human versus rabbit IgG, was identified as one important potential cause of this disagreement. In a cohort selected on IgM-RF reactivity against hIgG, multiple samples with low or undetectable reactivity against rIgG were identified. These would likely be incorrectly classified as RF[neg.] in rIgG-based assays. Because rIgG- and hIgG-based assays have not been compared head to head on clinical value, we cannot yet determine the clinical consequences of missing these hIgG-specific RF responses. It is clear however that diagnostic and therapeutic decisions based on RF results will be affected.

Problems with comparability between RF levels determined with nephelometry and turbidimetry were reported previously by Ameratunga et al. [28]. Van der Linden showed that RF levels measured in two high-level samples varied considerably between ELISA and nephelometry and turbidimetry and between different labs using the same method [29]. Comparing enzyme immunoassays, large discrepancies in qualitative as well as quantitative RF test results have been reported. A study by Bas et al. [30] found the same qualitative result in only 33% of samples when six IgM-RF ELISAs were compared. Agreement was 51% and 61% when results were stratified for the three rIgG- and the three hIgG-based assays. The authors concluded that quantitative results could not be compared across assays. This conclusion is supported by our data that show substantial discrepancies when comparing RF levels measured in the different commercial assays. There is disagreement on RF positivity in five samples

(17%). Similar to the experiments described in the present study, Bas et al. also performed a direct comparison for RF reactivity against rabbit and human target antigen, using rFc and hFc instead of the entire IgG molecule. Reported sensitivity was slightly higher for IgM-RF and significantly higher for IgA-RF when using hFc compared to rFc.

Because RFs are classified as autoantibodies, one would expect that testing for RF reactivity against hIgG has more value than against rIgG. An early study by Tuomi [31] found that the RF response in most RA patients bound both hFc and rFc in ELISA. However, some RA sera (4/93) that were RF[neg.] in agglutination tests using rIgG as well as in ELISAs with rFc did bind to hFc. Conversely, 3/37 RA sera with a positive agglutination test recognized only rFc, albeit at low levels. Notably, in non-RA subjects, Tuomi found a much larger proportion to be single[+] for rFc or hFc when retested against both targets. The data from the present study are in line with these earlier results. In the diagnostics cohort (Figure 3A), most likely a mix of early RA patients and non-RA subjects, a significant proportion of the samples contains an RF response that binds much better to hIgG than to rIgG. From the arthralgia cohort, we retested the subgroup that was ACPA[+] and RF[neg.] according to the original rIgG-based RF tests used for this cohort. We expected a better chance to find missed hIgG-specific RF responses in this group than in the originally ACPA[neg.] RF[neg.] arthralgia patients group because ACPA and RF often co-occur [32]. Some samples showed high reactivity against both hIgG and rIgG when retested. This may have been due to errors in the original RF assay or could have been caused by differences in the method of preparation of the rIgG, method of coating the rIgG (directly versus biotinylated rIgG on streptavidin-coated plates) or by differences in coating concentration. Strikingly, a substantial proportion of the arthralgia patients showed high anti-hIgG RF reactivity with low or almost no detectable anti-rIgG RF reactivity (Figure 3B). We estimate that up to 10% of these patients could have an anti-rIgG negative but anti-hIgG positive RF test result. It is important to note that the cutoff for anti-rIgG reactivity was not extensively validated for our assays and that these results cannot be directly translated into discrepancies between other rIgG- and hIgG-based assays.

Accurately classifying RF status is important in the arthralgia patient group because RF positivity is associated with a higher risk of progressing from arthralgia to RA in ACPA+ patients [6, 8]. The early RA cohort, consisting exclusively of confirmed RA cases included based on the presence of bad prognostic factors, showed reactivity between virtually all RF[+] samples and both targets (Figure 3C). Still, many samples showed higher levels

against hIgG, with six anti-hIgG+ samples showing RF levels $>2\times$ higher against hIgG than against rIgG. Based on the data obtained in the RA cohort that shows just one sample with much lower anti-hIgG than anti-rIgG RF reactivity, the risk of false positivity appears low when using rIgG as the target antigen. Instead, the data from the diagnostics cohort and the arthralgia cohort, showing RF responses with much higher anti-hIgG than anti-rIgG RF reactivity, indicate that there is a significant risk of false negativity especially among samples with relatively low RF levels. However, it may not be justified to speak of "false positivity" and "false negativity" until the clinical value of hIgG- and rIgG-specific RF responses has been determined. Using a hIgG-based assay with potentially higher sensitivity would be important particularly in the arthralgia phase, as discussed above.

The fact that, in all three cohorts, samples that show equal RF reactivity against hIgG can differ extensively in reactivity against rIgG suggests that RF responses can contain a mix of RF clones that recognize both targets and clones that solely recognize hIgG. This was further illustrated by monoclonal IgM-RF RF-AN reacting with both targets and RF61 exclusively reacting with hIgG (Figure 4). A number of previous studies [33–35] describe monoclonal RFs with restrictive reactivity against hIgG, whereas few studies suggest the presence of monoclonal RFs that bind to rIgG but not hIgG [36, 37]. When hIgG-specific RF clones such as RF61 constitute a large part of an RF response, measured levels will be much lower in rIgG-based assays. RF responses consisting exclusively of such RF clones will go undetected in rIgG-based assays. Misclassifying a patient with such an RF response as RFneg. can have real-life consequences in the clinic and it is therefore important that both clinical chemists and rheumatologist are aware of this issue.

The findings presented here also have important implications for harmonization of RF assays. The RF response is a very polyspecific response, and different individuals will have a different mix of RFs composing their RF response, with distinct contributions of RFs binding rIgG and/or hIgG and RFs binding with high or low affinity. Every assay detects these distinct characteristics of an RF response differently. The resulting discrepancies between assays are only detected when individual samples are compared between the assays, as has been done in the present study. When pooled samples are used to compare assays, these discrepancies level out, and are therefore overlooked, as illustrated in Figure 2. Here the 32 samples tested in Figure 1 were divided into quartiles based on their RF level, pooled and tested in the four RF assays. It is clear from Figure 2B that the pools give much better agreement between the assays. These data show that at the current state of art, harmonization is feasible for pooled sera, but not for individual samples. Articles referring directly or indirectly to RF should report the characteristics of the assay used for the measurements [38], and authors should be aware that both qualitative and quantitative results cannot be reliably compared across different assays.

In conclusion, there is substantial discrepancy in detected RF levels between the different RF assays used in the clinic. This is clinically relevant because the discrepancies are large enough to affect classification of RA using the ACR/EULAR classification criteria and can affect diagnosis of RA patients, prediction of their disease course and treatment decisions (if dependent on RF levels). Others have suggested ignoring RF level and looking solely at RF status [29], or leaving out the RF test in the work-up of a suspected RA patient [39]. We propose to focus on improving RF tests by further dissecting the RF response to identify the clinically most relevant RF reactivities. The next step could be validation of our findings in other cohorts from different countries, followed by international consensus meetings about standardization of RF tests. Eighty years after the first RF tests by Waaler, the rheumatology field should collaborate in the optimization of this still valuable test.

References

1. Waaler E. On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. Acta Pathol Microbiol Scand 1940;17:172–88.
2. Fraser KJ. The Waaler-Rose Test: anatomy of the eponym. Semin Arthritis Rheum 1988;18:61–71.

3. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.
4. Avouac J, Gossec L, Dougados M. Diagnostic and predictive value of anti-cyclic citrullinated protein antibodies in rheumatoid arthritis: a systematic literature review. Ann Rheum Dis 2006;65:845–51.
5. Lee DM, Schur PH. Clinical utility of the anti-CCP assay in patients with rheumatic diseases. Ann Rheum Dis 2003;62:870–4.
6. Bos WH, Wolbink GJ, Boers M, Tijhuis GJ, de Vries N, van der Horst-Bruinsma IE, et al. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. Ann Rheum Dis 2010;69:490–4.
7. Berglin E, Johansson T, Sundin U, Jidell E, Wadell G, Hallmans G, et al. Radiological outcome in rheumatoid arthritis is predicted by presence of antibodies against cyclic citrullinated peptide before and at disease onset, and by IgA-RF at disease onset. Ann Rheum Dis 2006;65:453–8.
8. Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
9. van Schaardenburg D, Hazes JM, de Boer A, Zwinderman AH, Meijers KA, Breedveld FC. Outcome of rheumatoid arthritis in relation to age and rheumatoid factor at diagnosis. J Rheumatol 1993;20:45–52.
10. Renaudineau Y, Jamin C, Saraux A, Youinou P. Rheumatoid factor on a daily basis. Autoimmunity 2005;38:11–6.
11. Brink M, Hansson M, Mathsson-Alm L, Wijayatunga P, Verheul MK, Trouw LA, et al. Rheumatoid factor isotypes in relation to antibodies against citrullinated peptides and carbamylated proteins before the onset of rheumatoid arthritis. Arthritis Res Ther 2016;18:43.
12. Sager D, Wernick RM, Davey MP. Assays for rheumatoid factor: a review of their utility and limitations in clinical practice. Lab Med 1992;23:15–8.
13. Bonagura VR, Artandi SE, Davidson A, Randen I, Agostino N, Thompson K, et al. Mapping studies reveal unique epitopes on IgG recognized by rheumatoid arthritis-derived monoclonal rheumatoid factors. J Immunol 1993;151:3840–52.
14. Allen JC, Kunkel HG. Hidden rheumatoid factors with specificity for native gamma globulins. Arthritis Rheum 1966;9:758–68.
15. Falkenburg WJ, van Schaardenburg D, Ooijevaar-de Heer P, Wolbink G, Rispens T. IgG subclass specificity discriminates restricted IgM rheumatoid factor responses from more mature anti-citrullinated protein antibody-associated or isotype-switched IgA responses. Arthritis Rheumatol 2015;67:3124–34.
16. Sasso EH, Barber CV, Nardella FA, Yount WJ, Mannik M. Antigenic specificities of human monoclonal and polyclonal IgM rheumatoid factors. The C gamma 2-C gamma 3 interface region contains the major determinants. J Immunol 1988;140:3098–107.
17. Duquerroy S, Stura EA, Bressanelli S, Fabiane SM, Vaney MC, Beale D, et al. Crystal structure of a human autoimmune complex between IgM rheumatoid factor RF61 and IgG1 Fc reveals a novel epitope and evidence for affinity maturation. J Mol Biol 2007;368:1321–31.

18. Falkenburg WJ, Kempers AC, Dekkers G, Ooijevaar-de Heer P, Bentlage AE, Vidarsson G, et al. Rheumatoid factors do not preferentially bind to ACPA-IgG or IgG with altered galactosylation. Rheumatology (Oxford) 2017;56:2025–30.
19. Buch MH, Smolen JS, Betteridge N, Breedveld FC, Burmester G, Dorner T, et al. Updated consensus statement on the use of rituximab in patients with rheumatoid arthritis. Ann Rheum Dis 2011;70:909–20.
20. Quartuccio L, Fabris M, Salvin S, Atzeni F, Saracco M, Benucci M, et al. Rheumatoid factor positivity rather than anti-CCP positivity, a lower disability and a lower number of anti-TNF agents failed are associated with response to rituximab in rheumatoid arthritis. Rheumatology (Oxford) 2009;48:1557–9.
21. Klein F, Janssens MB. Standardisation of serological tests for rheumatoid factor measurement. Ann Rheum Dis 1987;46:674–80.
22. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
23. Harindranath N, Goldfarb IS, Ikematsu H, Burastero SE, Wilder RL, Notkins AL, et al. Complete sequence of the genes encoding the VH and VL regions of low- and high-affinity monoclonal IgM and IgA1 rheumatoid factors produced by CD5+ B cells from a rheumatoid arthritis patient. Int Immunol 1991;3:865–75.
24. Corper AL, Sohi MK, Bonagura VR, Steinitz M, Jefferis R, Feinstein A, et al. Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. Nat Struct Biol 1997;4:374–81.
25. Burastero SE, Casali P, Wilder RL, Notkins AL. Monoreactive high affinity and polyreactive low affinity rheumatoid factors are produced by CD5+ B cells from patients with rheumatoid arthritis. J Exp Med 1988;168:1979–92.
26. Steinitz M, Izak G, Cohen S, Ehrenfeld M, Flechner I. Continuous production of monoclonal rheumatoid factor by EBV-transformed lymphocytes. Nature 1980;287:443–5.
27. Sutton B, Corper A, Bonagura V, Taussig M. The structure and origin of rheumatoid factors. Immunol Today 2000;21:177–83.
28. Ameratunga R, Musaad S, Sugrue C, Kyle C. Rheumatoid factor measurement–continuing problems 70 years after discovery. Clin Rheumatol 2011;30:1215–20.
29. van der Linden MP, Batstra MR, Bakker-Jonges LE, Foundation for Quality Medical Laboratory D, Detert J, Bastian H, et al. Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis: is it sensible to look at levels of rheumatoid factor? Arthritis Rheum 2011;63:1190–9.
30. Bas S, Perneger TV, Kunzle E, Vischer TL. Comparative study of different enzyme immunoassays for measurement of IgM and IgA rheumatoid factors. Ann Rheum Dis 2002;61:505–10.
31. Tuomi T. Which antigen to use in the detection of rheumatoid factors? Comparison of patients with rheumatoid arthritis and subjects with 'false positive' rheumatoid factor reactions. Clin Exp Immunol 1989;77:349–55.
32. Ioan-Facsinay A, Willemze A, Robinson DB, Peschken CA, Markland J, van der Woude D, et al. Marked differences in fine specificity and isotype usage of the anti-citrullinated protein antibody in health and disease. Arthritis Rheum 2008;58:3000–8.

33. Randen I, Thompson KM, Natvig JB, Forre O, Waalen K. Human monoclonal rheumatoid factors derived from the polyclonal repertoire of rheumatoid synovial tissue: production and characterization. Clin Exp Immunol 1989;78:13–8.

34. Robbins DL, Kenny TP, Snyder LL, Ermel RW, Larrick JW. Allotypic dependency of the specificity and avidity of human monoclonal IgM rheumatoid factors derived from rheumatoid synovial cells. Arthritis Rheum 1993;36:389–93.

35. Pascual V, Randen I, Thompson K, Sioud M, Forre O, Natvig J, et al. The complete nucleotide sequences of the heavy chain variable regions of six monospecific rheumatoid factors derived from Epstein-Barr virus-transformed B cells isolated from the synovial tissue of patients with rheumatoid arthritis. Further evidence that some autoantibodies are unmutated copies of germ line genes. J Clin Invest 1990;86:1320–8.

36. Fang Q, Kannapell CC, Gaskin F, Solomon A, Koopman WJ, Fu SM. Human rheumatoid factors with restrictive specificity for rabbit immunoglobulin G: auto- and multi-reactivity, diverse VH gene segment usage and preferential usage of V lambda IIIb. J Exp Med 1994;179:1445–56.

37. Kaplan S, Hyman K, Brooks R, Wakai M, Hashimoto S, Furie R, et al. Monoclonal IgM, IgG, and IgA human rheumatoid factors produced by synovial tissue-derived, EBV-transformed B cell lines. Clin Immunol Immunopathol 1993;66:18–25.

38. Goddard DH, Moore ME. Common tests for rheumatoid factors: poorly standardized but ubiquitous. Arthritis Rheum 1988;31:432–5.

39. Besada E, Nikolaissen C, Nossent H. Should rheumatoid factor in rheumatoid arthritis be sent to Davy Jones's Locker? Scand J Rheumatol 2012;41:85–8.