## Opinion Paper

Evanthia Monogioudi*, Gustavo Martos, Dana Petronela Hutu, Heinz Schimmel, Pier Luigi Meroni, Joanna Sheldon and Ingrid Zegers

# Standardization of autoimmune testing – is it feasible?

**Abstract:** Correct measurement of autoantibodies is essential for the diagnosis of autoimmune diseases. However, due to the variability of autoantibody results and the heterogeneity of testing, wrong diagnosis is a reality. For this and more reasons, harmonization of testing is of the outmost importance. In this review we have summarized the factors contributing to this variability. The ways with which the working group on harmonization of autoantibody testing of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) has been trying to tackle the issue with the production and correct use of certified reference materials (CRMs), is discussed. Finally the advantages and the limitations of the use of CRMs are presented.

**Keywords:** autoimmune disease; certified reference materials; harmonization; standardization.

## Autoimmune disease

The term autoimmune diseases covers a rather wide range of diseases in which body tissues are damaged or dysfunction due to an abnormal immune response to self-antigens [1]. The classification of a disease as an autoimmune depends on a number of criteria, such as the target antigen and the presence of T-cells and/or antibodies in a target organ or the loss of B-cell tolerance [2–4]. Autoimmune diseases are thought to occur in genetically susceptible individuals after they have been exposed to environmental factors [5]. Independently of the underlying cause, the impact of having an autoimmune disease on the patient, their families and wider society can be immense [6].

There are more than 100 diseases currently characterized as autoimmune [6] and they are part of a spectrum from one affected organ (organ-specific) to systemic conditions. In many autoimmune diseases, antibodies against one specific antigen are involved, however, in some autoimmune diseases antibodies to several antigens may be involved. Autoimmune diseases such as rheumatoid arthritis and type 1 diabetes are well recognized but there are many rare autoimmune diseases such as juvenile dermatomyositis or myasthenia gravis. Taken together autoimmune diseases affect many people, and can have a major impact on a patient's wellbeing. Using prevalence data for 29 autoimmune diseases Cooper et al. [7] estimated that the global prevalence of autoimmune disease is between 7.6 and 9.4%.

The healthcare costs burdening governments around the globe for the treatment and follow-up of patients with an autoimmune disease is enormous. According to a report published in 2011 by the American Autoimmune Related Diseases Association (AARDA) and the National Coalition of Autoimmune Patient Groups (NCAPG) [8] more than $100 billion are spent yearly for their treatment in the United States.

*Corresponding author: Evanthia Monogioudi, European Commission – Joint Research Centre, Directorate F, Retieseweg 111, 02440-Geel, Belgium, Phone: +32-14-571906, E-mail: Evanthia.monogioudi@ec.europa.eu. http://orcid.org/0000-0001-8525-8921
Gustavo Martos, Dana Petronela Hutu, Heinz Schimmel and Ingrid Zegers: European Commission – Joint Research Centre, Directorate F, Geel, Belgium
Pier Luigi Meroni: University of Milan, IRCCS Instituto Auxologico Italiano, Milan, Italy
Joanna Sheldon: St. George's Hospital, Protein Reference Unit, London, UK

## Laboratory measurement of autoantibodies

The detection and quantification of disease associated autoantibodies is central to the diagnosis, treatment and monitoring of many autoimmune diseases. Early diagnosis and treatment improves the likelihood of remission in the majority of autoimmune diseases and reduces the possibility of permanent tissue damage [9]. The reliability

of autoantibody measurements is therefore crucial for optimum patient care.

Historically, manual, qualitative or semi-quantitative indirect immunofluorescence assays have been used for the detection or screening of autoantibodies. ELISA-based assays have been used as follow-up investigations to confirm and quantify the concentration of an autoantibody in blood serum. These measurements have traditionally been performed in specialized sections of pathology, e.g. immunology laboratories. However, autoantibody testing is now commonplace with an increasing tendency towards more automated methods providing numerical results with faster throughput. Currently, there are at least three fully automated platforms in the market for the detection of antibodies associated with systemic small vessel vasculitis.

A laboratory result must be of value to the requesting clinician who will in turn need appropriate information to interpret it. The specificity (negativity in health) and sensitivity (positivity in disease) of the test(s) must be considered to ensure that they are capable of contributing to the diagnosis, prognosis or monitoring of patients. The result will need to be related to a reference range or a cut-off value appropriate to the test and to the patient being investigated [10]. Increasingly, tests and their results are incorporated into protocols or best practice guidelines. To fulfill these requirements, a laboratory test should give a consistent result wherever it is measured and by which ever method.

## Variability of autoantibody measurement results

As the number of available techniques for the measurement of autoantibodies is increasing, so is the concern with regard to the comparability of the obtained results. The detection and quantification of autoantibodies are fraught with problems. Laboratories around the world are required to participate in external quality assurance (EQA) schemes. It has been shown in various EQA studies [11] that both the within- and the between-method variation is very large even when the same units are used. The challenges of result comparability highlighted by EQA schemes have been summarized by Meroni et al. [12] in a review article published in *Nature Reviews*. Different laboratories using the same method may produce results varying up to a factor of 100. Additionally, different methods using the same units have been shown to in fact use scales that differ considerably, by up to a factor of 5 (e.g. for IgG

myeloperoxidase anti-neutrophil cytoplasmic autoantibodies [MPO ANCA] [13]). The problems are however not only numerical, but also qualitative. For MPO ANCA IgG it was shown that out of 30 routine samples only 10 samples had the same qualitative interpretation when tested in 11 different assays [13].

## Factors contributing to the variability of autoantibody results

– The growing number of in vitro diagnostics (IVD) methods with a wide variety of formats, reagents and signal detection systems may lead to *de facto* different method selectivities, even if they intend to measure the same analyte.
– Many methods, particularly the manual ELISAs have high coefficient of variability (CV) for repeatability and intermediate precision. Because of this, the uncertainty on a single measurement is high.
– Measurement scales used by the different methods for the same analyte can vary up to a factor of 100. In addition, antibody concentrations are often expressed in the same arbitrary units (e.g. "units" or "IU"), which further increases the confusion. In principle this does not prevent the proper use of the results, provided the scales are stable over time, and that the right cut-offs are used. However, it is clear that different classifications of patients are obtained even if the corresponding cut-offs provided by the manufacturers are used [13].
– In some cases different methods are using the same units (e.g. "units" or "IU") but in reality employ very different scales, which can lead to confusion. A consensus discussion on how to deal with different units and scales could improve this issue.
– Differences in test specificity and test efficiency can be seen, even in the presence of reliable standards. These differences may be between batches from one individual manufacturer or between kits from different manufacturers. These underlying causes are likely to be multi-factorial and include variability in the calibrators, enzyme substrates, antigens and antibodies between the different kits.
– The calibration of autoimmune measurements is not always optimal. The translation from the method signal (e.g. optical densities [ODs]) to method units is typically via a calibration curve. The calibration curve is usually non-linear with a sigmoid curve, at least for ELISAs. The method optimization is vital and should

focus on generating the most reliable concentration values at clinically critical values. However, the number of calibration points may be very low and cover a large concentration range. Any variation in any standard can skew the standard curve sufficiently to alter the interpretation of borderline results. The flattened curve seen at low and high concentrations can give high analytical imprecision both within and between batches and can therefore make monitoring patients difficult. This has been noted for MPO ANCA IgG analysis by Hutu et al. [13] where the coefficient of variation (CVs) for reported results in method units were much larger than seen for the raw OD values.

– The antibodies made by each individual patient are likely to be slightly different from those made by other patients; they may differ in class, subclass but most importantly in selectivity, affinity and avidity for the antigen. Methods that describe measuring the same analyte may actually be detecting only certain small parts of the three-dimensional antigen molecule. These factors contribute to the poor correlation that may be seen between methods. For example, one sample may give a high value in method A and a low value in method B, whereas a sample from a different patient will give a low value in method A and a high value in method B. Furthermore the reactivity of autoantibodies can change within a patient from the time of diagnosis, through treatment and into remission. Some methods may show good correlation with other usually similar methods but for other combinations of methods the scatter may be severe. It is important to remember that poor correlation between methods that is due to sample characteristics' and the differences in selectivities between the methods will not be improved by the introduction of a common reference material (RM). Laboratories should have a robust procedure to validate their assays when first adopted and for continual monitoring of the testing. Internal quality control (IQC) and EQA procedures are fundamental to these processes.

## Standardization, harmonization and certified reference materials

According to the ISO/IEC Guide 2:2004, standardization is the "activity of establishing, with regard to actual or potential problems, provisions for common and repeated use, aimed at the achievement of the optimum degree of order in a given context" [14].

The Clinical and Laboratory Standards Institute refers to a definition of "(method) harmonization" as "the process of recognizing, understanding, and explaining differences while taking steps to achieve uniformity of results, or at a minimum, a means of conversion of results such that different groups can use the data obtained from assays interchangeably" [15]. In terms of measurement results in clinical chemistry the focus of standardization is the term "repeated use". Standardization implies long-term comparability of measurement results, preferably through traceability of measurement results to a stable reference. On the other hand, the focus of the term harmonization is on uniformity of results.

Standardization through the proper implementation of a reference system should ideally support the equivalence of measurement results over different procedures, laboratories and over time. However, equivalence of measurement results also depends to a large extent on all the other factors contributing to variability, like the repeatability of methods and sample-specific differences. It cannot be achieved by traceability to a reference system alone.

A reference system should provide traceability to a stable reference, such as an RM or a reference method. An RM is a material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process. A certified reference material (CRM) is characterized by a metrologically valid procedure for one or more specified properties, accompanied by a certificate that provides the value of the specified property, its associated uncertainty, and a statement of metrological traceability. A CRM thus provides a measurement scale, whether arbitrary or SI derived values are being used. The use of CRMs has as a major advantage the fact that reproducibility over longer times can be possible as long as the procedures followed by individual manufacturers are well controlled. These procedures would cover all aspects from the reconstitution of the CRM, to the preparation of dilutions and the use of valid protocols for the value transfer to intermediate calibrators.

The quantification of autoimmune biomarkers presents particular challenges because of the nature of the entity measured. The analyte is an antibody against a named antigen and, immunoassays rely on detecting the binding of the analyte to its natural antigen. The antigens associated with an autoimmune disease are typically large proteins with multiple epitopes but the lack of knowledge regarding an antigen's epitopes with respect to its analytical reactivity but also clinical importance is a significant issue that is hard to address. Consequently, autoantibodies

with different epitope specificities will often produce discrepant results depending on the method selectivity. Where evidence regarding clinically relevant epitopes is available, assay manufacturers should produce appropriate assays. Difference between the affinities of the antibodies in the patients samples and of the calibrator will have an impact on the accuracy of the assay [16]. It is desirable that calibration material has a comparable affinity to that of the typical patient sample. A common standard or RM must resemble routine samples with respect to the most relevant influence parameters ensuring commutability among different measurement procedures [12]. It is important to stress that the main requirement for an RM that is fit for purpose, is its commutability, i.e. the fact that it behaves in the same way as a patient sample with respect to the relevant methods.

According to the EU Directive on In Vitro Diagnostic Medical Devices (IVD-MD) (Directive 98/79/EC), any material intended to serve as a calibrant or as a control material, must be traceable to reference measurement procedures and/or RMs of higher order, if they are available.

## IFCC WG-HAT

It is well established that the detection and quantification of IgG antibodies to autoantigens are important for the diagnosis and monitoring of a number of autoimmune diseases. In 2009 the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) formed a new working group with a mandate for the Harmonization of Autoantibody Tests (known as WG-HAT). The Joint Research Centre of the European Commission is participating in this working group and is responsible for the development and production of CRMs based on a series of agreed targets. These initial targets for the standardization of autoantibody measurement results are set according to their acute, well-defined clinical use.

## The targets and their selection

The expert group forming the WG compiled a list of autoantibodies where the concentration of the antibody was important for diagnosis and more importantly, for disease or treatment monitoring and where harmonization of results could reduce clinical risk.

These are IgG autoantibodies to:
– Double stranded DNA – one of the classification criteria for systemic lupus erythematosus (SLE) and used for monitoring disease activity and response to treatment [17–19].
– Glomerular basement membrane – a pathogenic antibody that mediates anti-GBM disease. It is typically of an IgG type even though IgA and IgM types can also been present. The diagnosis and prognosis depends on early detection of the antibody [20–22] and during treatment, the concentration of the antibody is used to monitor disease activity and response to therapy.
– Anti-cardiolipin and anti-beta 2 glycoprotein 1 antibodies – antibodies associated with a higher risk of thrombosis in veins and/or arteries and pregnancy complications. Their presence is indicative for antiphospholipid syndrome (APS) [23–26].
– Proteinase 3 – one of the targets for anti-neutrophil cytoplasmic antibodies (ANCA) that is strongly implicated in the pathogenesis of ANCA small vessel associated vasculitis. PR3 ANCA IgG are found in about 80% of patients with granulomatosis with polyangiitis (GPA), and in about 35% of patients with microscopic polyangiitis (MPA), eosinophilic granulomatosis with polyangiitis (EGPA), and renal-limited rapidly progressive glomerulonephritis [27, 28]. Detection and quantification of PR3 ANCA IgG is important for diagnosis but also for monitoring disease activity and response to therapy.
– Myeloperoxidase – another target of ANCA, implicated in the pathogenesis of ANCA small vessel associated vasculitis, especially of patients suffering from MPA and EGPA. Detection and quantification of MPO ANCA IgG is similarly important.
– Anti-citrullinated protein antibodies (ACPA) – antibodies that are present in patients with rheumatoid arthritis and are therefore used as markers for diagnosis of the disease [29] but are also used alongside other markers for risk stratification and as prognostic indicators.

The IFCC working group (WG-HAT) addressed a number of questions regarding these targets.

## Are there different methods measuring the same analyte – do the methods provide correlating results?

Prior to starting any standardization initiative, it is important to evaluate the consistency of the entities measured

by the different methods to be standardized. The numerical values may show marked variation but methods claiming to detect the same analyte should produce results that show a consistent relationship [30].

Practically, this can be done by analyzing a set of routine clinical samples covering the whole analytical range using all the methods available. The data can be plotted in a method A vs. method B graph providing a visual estimate of the correlation and linearity and statistical analysis will provide numerical values for these characteristics. Poor correlation between methods is a serious impediment for standardization and the underlying causes need thorough evaluation. If there is no consistent relationship in values between methods, recalibration alone will not be sufficient to achieve comparable results on the level of individual samples.

Data from EQA schemes show the marked variation in autoantibody results from different methods and this is supported in the scientific literature. Anti-cardiolipin antibodies (aCL) is an important test for the diagnosis of anti-phospholipid syndrome (APS) but there is a number of studies [31–35] showing the lack of agreement between the methods used for the analysis. These reasons vary from the preparation or formulation of the antigen, to the method characteristic and these lead to differences in the cut-off values and therefore to the final interpretation of the results. Kutteh and Franklin distributed samples from 20 patients to 10 different centers for measurement of antibodies to cardiolipin and other phospholipids. They showed that results from different centers were in agreement in only 45% of the cases [36]. Many studies suggest that, patient samples should be tested by more than one method whenever possible whether a common standard is used or not [36].

A similar situation has been described for MPO and PR3 ANCA by Bossuyt et al. [37], who found a lack of correlation between different methods for PR3 and MPO ANCA although the correlation was better between methods using a similar analytical process. Hutu et al. [13] showed a fairly good correlation for patient samples between most methods (Figure 1) although results from different methods (in terms of positive or negative) could be different. It was shown that recalibration and the use of a common cut-off could considerably reduce the variation of results although certain methods did not give comparable results, presumably due to different analytical selectivity. This finding does not mean that these methods are not valid, but it does suggest that they may be detecting antibodies in a different sub-set of patients. These findings also highlight the importance of a better understanding of the root cause of variation in autoantibody testing.
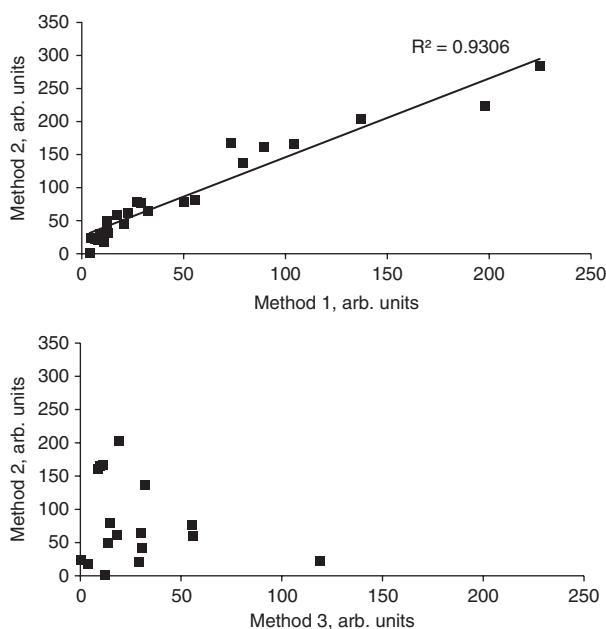


**Figure 1:** Examples of good (above) and poor (below) correlation between measurement results from three methods that target MPO ANCA IgG antibodies.

# Is it possible to produce a commutable reference material for autoantibody tests?

The selection of a raw material and the stabilization of a raw material are crucial steps in the production of an RM. There are different options for starting materials. These include plasmapheresis materials, pooled plasma or serum samples, and the use of monoclonal antibodies, eventually spiked into a serum-like background. It may seem that any matrix-matched (serum) material should be suitable. This is an assumption that has also been made for RMs for other analytes, and has been shown not always to be valid. There are different reasons why real serum material from patients could still be unsuitable for producing an RM:

– The raw material could contain antibodies that are outliers in their analytical behavior, with respect to antibodies in the majority of patient samples. For example, in Figure 2 the plasmapheresis materials shown as a filled triangle or a filled large square show a bias of about 30% with respect to the regression line of the patient samples. If this plasmapheresis material would be used for calibration of the two methods it would result in an average bias of 30% for the patient samples.
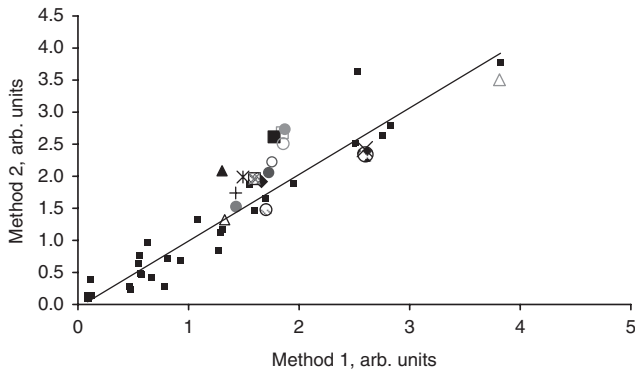
**Figure 2:** Correlation between optical density (OD) responses for two PR3 ANCA IgG methods.
Results for routine patient samples are shown as small squares, candidate reference materials as larger symbols.

- The material could have stability problems, changing in properties over time.
- Processing of the material, like freeze-drying, can change the oligomeric state of antibodies, or lead to lack of between-vial homogeneity.

The ability of an RM to perform comparably to tested routine samples is called commutability and it is one of its most important characteristics [38]. Reference preparations made from serum in its close to native state may have issues with the complexity of the constituent proteins and even protein interactions that may all generate interferences and poor commutability. Reference preparations made from purified proteins may show a lack of commutability because of degradation or structural modifications made to the proteins during isolation and purification, or because of a different matrix composition. There are also advantages and disadvantages to using samples from one patient, e.g. with plasmapheresis fluids or to using mixtures of a large number of donations from individual patients. There is no a priori rationale for choosing the one or the other strategy for the selection of raw material and processing conditions. Instead the selected material, or preferably different candidate materials, should be tested for commutability.

Despite the challenges, finding a material that is commutable is not impossible. Hutu et al. [13] completed a study for myeloperoxidase MPO ANCA IgG, where 18 forms of candidate materials were analyzed with different manufacturer assays. The results for both patient samples and candidate RMs were plotted in pairs. It was shown that results correlated reasonably well for a subset of the methods, while for other combinations there was little correlation (Figure 3). The group continued with a
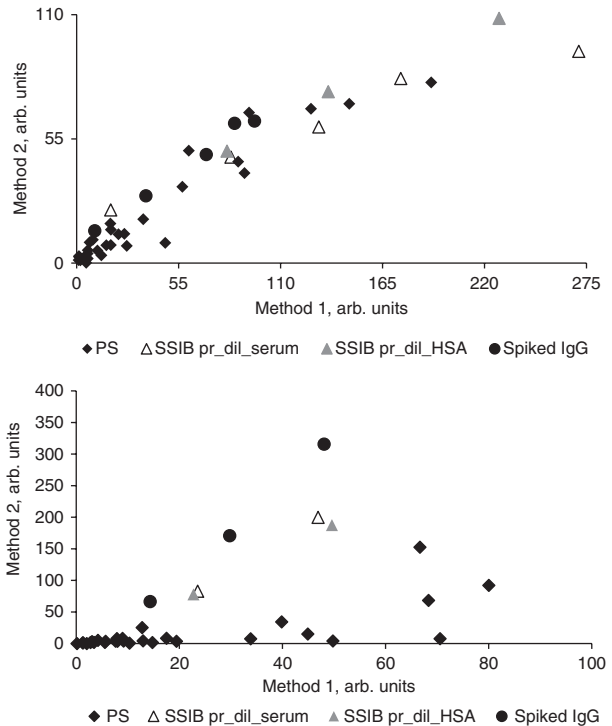




**Figure 3:** Example of good (above) and poor (below) correlation between measurement results from three methods that target MPO ANCA IgG antibodies for both patient samples (PS) and various candidate reference materials (SSIB is a random code given for a PS) these samples were processed and diluted whether in human serum or in human serum albumin.
Purified PR3 ANCA IgG was spiked into filtered unprocessed serum.

second commutability study, where the five best forms of the candidate material were analyzed using 11 assays (manual ELISAs, a multiplex based assay, a fluorescence assay and a chemiluminescent assay). That study resulted in the selection of a freeze-dried form that was finally chosen by the group as the material to be developed and certified [13].
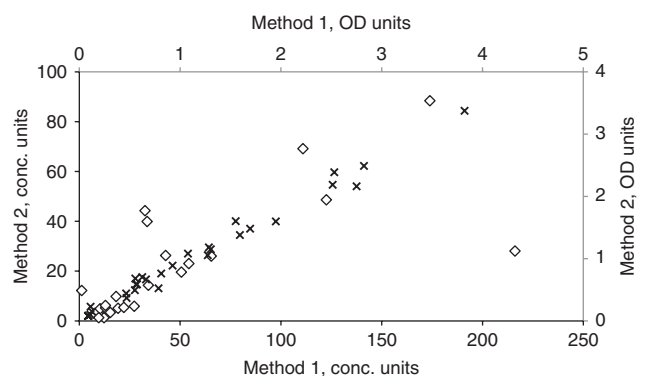


**Figure 4:** Correlation between optical density (OD) responses (x) and concentration estimates (rhombus) of two PR3 ANCA IgG methods.

PR3 ANCA IgG candidate RMs have also been evaluated in feasibility studies comprising methods from several manufacturers. Some serum materials were commutable for all but one method that where compared with respect to their assay response (i.e. optical density). Interestingly, correlation between methods significantly worsened when considering the calculated PR3 ANCA IgG titers (Figure 4), which points at potential calibration deficiencies. In case these issues would be related to the use of non-commutable calibrators, a common suitable RM could improve methods' comparability. It is important to remember that the preparation of an RM is only the first step; other areas that must be robust include the value transfer to intermediate calibrators, selecting the appropriate curve fitting model and ensuring linearity over a clinically useful concentration range.

## Can values be assigned in a manner that they are traceable to a stable reference?

A pivotal requirement for a calibrator according to ISO/IEC 17511 is the metrological traceability of the material [39]. It is defined as the "property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty". The mass concentration values assigned to CRMs for MPO and PR3 ANCA (ERM-DA476/IFCC and ERM-DA483/IFCC, respectively), are operationally defined by the immunoassay procedures used to characterize them and are traceable to the stated value of the mass concentration of total IgG in United States National Reference Preparation (USNRP) 12-0575C [40]. The certification report accompanying the release of any CRM will contain a graphical presentation and detailed explanation of the traceability chain [41].

The RMs for both MPO ANCA and PR3 ANCA enable consistent results to be generated using a large number of methods. However, some methods show outliers in both concentration and clinical interpretation. It is likely that the relative success of these materials originates in the initial selection of the raw material, which was chosen because it behaved as an "average" patient sample within the evaluation methods. However, as previously discussed, the properties of IgG ANCA in individual patient samples can be variable so we are aware that some patient samples and in some methods will be impossible to calibrate using the existing RM.

## Progress in the development of CRMs

A number of CRMs has been developed, certified and released by the JRC in collaboration with the IFCC for standardization of the mass concentrations of various proteins. These include ERM-DA470k/IFCC for 12 human serum proteins including total IgG, ERM-DA471/IFCC for cystatin C and ERM-DA474/IFCC for C-reactive protein. More recently and as briefly mentioned above, the JRC released two serum protein materials for the standardization of measurements of PR3 ANCA IgG (ERM-DA483/IFCC) [41] and for the measurement of MPO ANCA IgG (ERM-DA476/IFCC) [42]. These materials represent a significant step forward in the harmonization of autoantibody tests as the first CRMs with values assigned in mass units for a specific IgG antibody. The production of these materials has taken approximately 6 years but in addition to producing and certifying them, we have gained considerable experience in the area of autoantibody standardization which should enable the development of further materials using the same well-defined protocols.

Preparation of a CRM certified for the mass concentration of β-2-glycoprotein I antibodies in human serum that are associated with occurrence of arterial and/or venous thrombosis as well as with recurrent miscarriage [43] is in progress. The IFCC committee on harmonization of autoantibody testing and the JRC are planning the production of a certified reference preparation for IgG antibodies to the glomerular basement membrane and considering producing an RM for IgG and IgA anti-tissue transglutaminase antibodies. Other groups have expressed their intention to prepare RMs for IgG anti-double stranded DNA, anti-cyclic citrullinated peptide and rheumatoid factor.

## Advantages and limitations of the use of certified reference materials for testing for autoimmune antibodies

Tests for autoantibodies are important in the diagnosis and monitoring of autoimmune disease with many tests forming part of the diagnostic criteria. Many methods for autoimmune antibodies show good clinical significance but there are still important issues with limited RMs and high variability between methods, making robust result interpretation difficult.

An RM can provide an anchor point for working on those issues. They can be used for quality control, to detect and reduce within and between-laboratory variation. When methods provide correlating results they can be used to put these methods on the same scale, provided that they are commutable. Importantly, they can also be used by IVD manufacturers to keep scales stable over time. They will however not solve all issues, and will not reduce variability between non-correlating methods. Some methods will certainly generate different numbers and different interpretations for some samples but provided the laboratory is consistent in the method used for their patients, and the clinicians are aware that results for the same test name are not necessarily comparable, clinical risk should be limited. A better understanding of the underlying causes of variability in autoantibody testing should enable us to find ways to inform our users; this may involve more detailed nomenclature of the tests, e.g. PR3 ANCA capture method or PR3 ANCA direct ELISA. It may also lead to better definition of the antigens involved in the testing, e.g. defining the epitopes that are reactive in a particular assay. We are aware that true harmonization or standardization of autoantibody testing may be a multi-step process but having one stable "point of reference" in an analytical system is the vital first step in improving the quantification of autoantibody testing.

# References

1. Mackay IR, Burnet FM. Autoimmune diseases, pathogenesis, chemistry and therapy. Springfield, IL: John Wiley & Sons, Inc., 1963.
2. Rose NR, Bona C. Defining criteria for autoimmune diseases (Witebsky's postulates revisited). Immunol Today 1993;14:426–30.
3. Pillai S, Mattoo H, Cariappa A. B cells and autoimmunity. Curr Opin Immunol 2011;23:721–31.
4. Giltiay NV, Chappell CF, Clark EA. B-cell selection and the development of autoantibodies. Arthritis Res Ther 2012;14(Suppl 4):S1.
5. Mackay IR. Tolerance and autoimmunity. West J Med 2001;174:118–23.
6. Shoenfeld Y, Cervera R, Gershwin E. Diagnostic criteria in autoimmune diseases. Totowa: Humana Press, 2008.
7. Cooper GS, Bynum ML, Somers EC. Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. J Autoimmun 2009;33:197–207.
8. American Autoimmune Related Diseases Association, National Coalition of Autoimmune Patient Groups. The cost burden of autoimmune disease: the latest front in the war on healthcare spending. Eastpointe, MI: American Autoimmune Related Diseases Association, Inc., 2011.
9. Sen D, Brasington R. Tight disease control in early RA. Rheum Dis Clin North Am 2012;38:327–43.
10. Wiik A, Cervera R, Haass M, Kallenberg C, Khamashta M, Meroni P, et al. European attempts to set guidelines for improving diagnostics of autoimmune rheumatic disorders. Lupus 2006;15:391–6.
11. Van Blerk M, Van Campenhout C, Bossuyt X, Duchateau J, Humbel R, Servais G, et al. Current practices in antinuclear antibody testing: results from the Belgian External Quality Assessment Scheme. Clin Chem Lab Med 2009;47:102–8.
12. Meroni PL, Biggioggero M, Pierangeli SS, Sheldon J, Zegers I, Borghi MO. Standardization of autoantibody testing: a paradigm for serology in rheumatic diseases. Nat Rev Rheumatol 2014;10:35–43.
13. Hutu DP, Tuddenham E, Monogioudi E, Meroni P, Schimmel H, Sheldon J, et al. First steps in the standardization of immunoglobulin IgG myeloperoxidase-anti-neutrophil cytoplasmic antibody measurements. Clin Exp Immunol 2016;183:193–205.
14. ISO/IEC Guide 2:2004. Standardization and related activities – General vocabulary. Geneva, Switzerland, 2004.
15. Harmonized Terminology Database. Accessed: 12 Jan 2018.
16. Griswold WR, Chalquest RR. Theoretical analysis of the accuracy of calibrated immunoassays for measuring antibody concentration. Mol Immunol 1991;28:727–32.
17. Isenberg DA, Ehrenstein MR, Longhurst C, Kalsi JK. The origin, sequence, structure, and consequences of developing anti-DNA antibodies. A human perspective. Arthritis Rheum 1994;37:169–80.

18. van Bruggen MC, Walgreen B, Rijke TP, Tamboer W, Kramers K, Smeenk RJ, et al. Antigen specificity of anti-nuclear antibodies complexed to nucleosomes determines glomerular basement membrane binding in vivo. Eur J Immunol 1997;27:1564–9.

19. Khalil M, Spatz L, Diamond B. Anti-DNA antibodies. In: Lahita RG, editor. Systemic lupus erythematosus. New York: Academic Press, 1999:197–217.

20. Lahmer T, Heemann U. Anti-glomerular basement membrane antibody disease: a rare autoimmune disorder affecting the kidney and the lung. Autoimmun Rev 2012;12:169–73.

21. Bolton WK. Goodpasture's syndrome. Kidney Int 1996;50:1753–66.

22. Pusey CD. Anti-glomerular basement membrane disease. Kidney Int 2003;64:1535–50.

23. Brusch A. The significance of anti-beta-2-glycoprotein I antibodies in antiphospholipid syndrome. Antibodies 2016;5:16.

24. Raby A, Moffat K, Crowther M. Anticardiolipin antibody and anti-beta 2 glycoprotein I antibody assays. Methods Mol Biol 2013;992:387–405.

25. Miyakis S, Lockshin MD, Atsumi T, Branch DW, Brey RL, Cervera R, et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). J Thromb Haemost 2006;4:295–306.

26. Meroni PL, Borghi MO, Raschi E, Tedesco F. Pathogenesis of antiphospholipid syndrome: understanding the antibodies. Nat Rev Rheumatol 2011;7:330–9.

27. Hoffman GS, Specks U. Antineutrophil cytoplasmic antibodies. Arthritis Rheum 1998;41:1521–37.

28. van Der Woude FJ, Lobatto S, Permin H, Van Der Giessen M, Rasmussen N, Wiik A, et al. Autoantibodies against neutrophils and monocytes: tool for diagnosis and marker of disease activity in Wegener's granulomatosis. Lancet 1985;325:425–9.

29. Wigerblad G, Bas DB, Fernades-Cerqueira C, Krishnamurthy A, Nandakumar KS, Rogoz K, et al. Autoantibodies to citrullinated proteins induce joint pain independent of inflammation via a chemokine-dependent mechanism. Ann Rheum Dis 2016;75:730–8.

30. Schimmel H, Zegers I, Emons H. Standardization of protein biomarker measurements: is it feasible? Scand J Clin Lab Invest Suppl 2010;242:27–33.

31. Audrain MA, Colonna F, Morio F, Hamidou MA, Muller J-Y. Comparison of different kits in the detection of autoantibodies to cardiolipin and beta2glycoprotein 1. Rheumatology (Oxford) 2004;43:181–5.

32. Robert JM, Macara LM, Chalmers EA, Smith GC. Inter-assay variation in antiphospholipid antibody testing. Br J Obstet Gynaecol 2002;109:348–9.

33. Coulam C, McIntyre J, Wagenknecht D, Roth N. Interlaboratory inconsistencies in detection of anticardiolipin antibodies. Lancet 1990;335:865.

34. Peaceman AM, Silver RK, MacGregor SN, Socol ML. Interlaboratory variation in antiphospholipid antibody testing; discussion. Am J Obstet Gynecol 1992;166(6 Pt 1):1780–4.

35. Lakos G, Bentow C, Mahler M. A clinical approach for defining the threshold between low and medium anti-cardiolipin antibody levels for QUANTA flash assays. Antibodies 2016;5:14.

36. Kutteh WH, Franklin RD. Assessing the variation in antiphospholipid antibody (APA) assays: comparison of results from 10 centers. Am J Obstet Gynecol 2004;191:440–8.

37. Bossuyt X, Rasmussen N, van Paassen P, Hellmich B, Baslund B, Vermeersch P, et al. A multicentre study to improve clinical interpretation of proteinase-3 and myeloperoxidase anti-neutrophil cytoplasmic antibodies. Rheumatology (Oxford) 2017;56:1533–41.

38. Vesper HW, Emons H, Gnezda M, Jain CP, Miller WG, Rej R, et al. Characterisation and qualification of commutable reference materials for laboratory medicine: approved guideline. CLSI Document C53-A. Wayne, PA: CLSI, 2010.

39. Standardization ISO. ISO 17511. In vitro diagnostic medical devices – measurement of quantities in biological samples – metrological traceability of values assigned to calibrators and control materials. Geneva, Switzerland, 2003.

40. Reimer CB, Smith SJ, Wells TW, Nakamura RM, Keitges PW, Ritchie RF, et al. Collaborative calibration of the U. S. National and the College of American Pathologists reference preparations for specific serum proteins. Am J Clin Pathol 1982;77:12–9.

41. Monogioudi E, Hutu DP, Charoud-Got J, Sheldon J, Schimmel H, Trapmann S, et al. Certification report: the certification of anti-proteinase 3 immunoglobulin G in human serum ERM® – DA483/IFCC. Belgium: Publications Office of the European Union, 2017.

42. Monogioudi E, Hutu DP, Martos Sevilla G, Tuddenham E, Sheldon J, Schimmel H, et al. Certification report: the certification of anti-myeloperoxidase immunoglobulin G in human serum ERM® – DA476/IFCC. Belgium: Publications Office of the European Union, 2015.

43. de Groot PG, Urbanus RT. The significance of autoantibodies against β2-glycoprotein I. Blood 2012;120:266.