Autoimmune diagnostics by immunofluorescence: variability and harmonization

The antinuclear antibodies (ANA) determination is one of the most commonly used techniques in the autoimmunity clinical laboratory. Far from being outdated, indirect immunofluorescence (IF) is a powerful laboratory tool not only for clinical diagnostics, but for disease follow-up and prognosis estimation as well. Unlike other more precise quantitative techniques, IF can be affected by a significant amount of variability factors impacting repeatability and reproducibility of the results obtained. With the goal to minimize this variability and improve the quality of the results, there are different initiatives that can be undertaken to help obtain more accurate, precise and reliable results.

by Dr Petraki Munujos

Variability

It is well known that indirect immunofluorescence (IF) is a technology that expresses its results as presence/absence, reactive/non reactive or positive/negative, therefore, the test significance is focused on the decision point or cut-off that separates one condition from the other. The accuracy with which that point of decision has been set, determines the variability among the results obtained with reagents from different manufacturers, in different laboratories or in countries with different ethnic populations.

Noting this particular degree of variability, it is essential to reduce this analytical variability by standardizing the procedures and harmonizing the results obtained with this particular technique. The term standardization concerns analytical aspects and usually it is based on the availability of reference materials or procedures. However, the term harmonization reflects the consensus among the different actors involved in nomenclatures, reference values or diagnostic algorithms.

The factors that contribute to the variability of the results can be classified in three different categories (Table 1): biological factors, related basically to the reagent; procedural factors, associated with the analyst and the analyser used; and the factors related with the decision that is taken according to the test result.

The **biological factors** include:
- The diversity of HEp2 cells strains available in the market and in the numerous cell collections found all over the world. They all share the same name, but show differences in growth rate, morphology or even in

BIOLOGICAL FACTORS	examples	affectation
HEp2 cell strain	Growth rate, morphology	Sensitivity and localization
Culture conditions	Antibiotics, medium, drugs, time, temperature	Antigen expression (sensitivity)
Adherence proteins	poly-L-Lysina, collagen (type I, II or IV), fibronectin, gelatin, condroitin.	Morphology, antigen expression
Animals used	Sex, age, strain, diet	Sensitivity and specificity
Slides processing	Organs and tissue preservatives, fixatives (acetone, methanol, formalin, none)	Sensitivity, specificity, stability
Conjugates	Isotype, species, nature of immunogen, purification method, fluorochrome, fluorescein / protein ratio (F/P), concentration, anti-fading	Sensitivity and specificity
PROCEDURAL FACTORS	examples	affectation
Samples	Collection and storage temperature, freeze-thawing cycles, records on anomalous sera, ambiance temperature, dust free areas	Repeatability and reproducibility
Processing	Manual vs. automated, faithfulness to insert, errors in results transcription	Repeatability and reproducibility
Microscope	Leds vs. Mercury lamp, optical quality, camera sensitivity	Sensitivity
Image interpretation	Expertise, training, computer assisted	Diagnostic capability
Cut-off verification	Collection of sera classified by clinical criteria	Diagnostic capability
DECISION FACTORS	examples	affectation
Starting dilution	Diverging recommnedations, differences in ethnicity and target populations	Diagnostic capability
Patterns nomenclature	Ambiguous descriptions, different names for the same antibody pattern	Reproducibility
Diagnostic strategy	Excessive diagnostic kits, diverse diagnostic algorithms, non-suitable requests	Time, cost
Reports	non-suitable requests, diverse information, limited lab-clinician communication	Time, cost
Guidelines & recommendations	Diverging criteria, insufficient difusion, limited implementation	Reproducibility

Table 1. Factors of variability involved in the design, execution and performance of IVD Indirect Immunofluorescence tests.

the expression of certain proteins.

- The effects of being exposed to antibiotics, different culture media, cell cycle synchronizing agents or different culture conditions can induce differences in the level of protein expression and, consequently, can affect the sensitivity of the assay.
- In tests based on cultured cells, the use of extracellular matrix molecules (ECM) is common. Many of these ECM contain amino acid sequences that are recognized and react with the cell membrane receptors. This binding, together with the growing factors present in the medium, causes a proliferative response, promoting the growth and spreading of the cells on the growing surface. Modifying these contacts may result in coordinated changes in the cell, the cytoskeleton and the nucleus morphology. Thus, the manufacturer must assure the right choice of ECM since it can have a direct effect on the final staining pattern.
- In a similar way to cell culture, IF on animal tissue sections can be affected by differences due the strain, sex, age or diet of the animal.
- Moreover, the use of specific solutions to prevent loss of antigenicity during the tissue extraction or the use of certain fixatives can dramatically alter the staining pattern.

The second kind of elements contributing to variability are the **procedural factors** and can be noticeable at any stage of the analytical process.

- Several tasks in the *pre-analytical stage* (samples management) may exert an effect on the reproducibility of the final results: the obtainment, treatment, handling or storage of the samples; the results-patient traceability; records on incidences or potential interferences (hemolysis, lipemia, bilirubin, rheumatoid factor) that can later explain certain results.
- Regarding the *analysis of the results* under the microscope, the expertise and the analyst skills in interpreting the images are probably some the most influencing elements for a correct diagnosis. In addition, the characteristics of the microscope (LED versus conventional mercury lamps), the use of computer-aided diagnostic systems and the sensitivity of

Figure 1. Diversity in the recommendations on the information to be included in the laboratory report, according to different societies and organizations. ⊗*, not mentioned;* ✓ *defined;* ✓*, optional; BEL, Belgian recommendations; BRA, Brazilian consensus; NED, Netherlands; GER, Germany; ITA, Firma (*Forum Interdisciplinare per la Ricerca nelle Mallattie Autoimmuni*); EASI, European Autoimmunity Standardization Initiative; ACR, American College of Rheumatology. Adapted from Damoiseaux et al., 2016 [1].*

the camera play an important role.

- As for the *test procedure*, differences can be observed when comparing manual versus automated processing. Other factors like pipetting, handling or mistakes in transferring data along the process may impact in the final result. Special attention should be given to the lack of compliance to the instructions for use: it is a common practice in many laboratories to modify the instructions provided with the kit (starting dilution, incubation times…) at their convenience.
- When a laboratory faces the *validation* of a commercial kit to be acquired in the routine, it is common, or it should be, to verify the soundness of the cut-off recommended by the manufacturer. The selection of the sera used for this purpose is critical. It looks obvious that the quality and the composition of the collection of classified sera are crucial to confirm the diagnostic capability of the test.

The third group of variability circumstances is those related to the decision taken with the analytical results in view of guidelines, standards, recommendations or in the absence of any of them. They could be called **decision factors**:

- The *sample dilution,* the actual cut-off, still is a controversial issue. There are several recommended dilutions depending on the manufacturer, the guidelines edited by scientific and medical societies, or even local habits.
- The *nomenclature* and description of the immunofluorescence staining patterns are different all over the world: different names describing the same pattern, ambiguous descriptions which lead to diagnostic inaccuracies.
- The lack of harmonized *diagnostic algorithms* results in different practices according to the laboratory or the local health system. Currently, there is an excess of tests in the market often with no added value but with a significant impact in the cost of health services and in the heterogeneity of the existing diagnostic profiles.
- When the *test request* is not justified (and this can be the case of numerous ANA screening requests), the prediction values and even the credit of the test can be affected, because very different conclusions are drawn when the results are confronted to the diverse clinical situations that generated the request.
- There is no consensus about the information that the *report* of the results must contain. This is specially important in the case of the ANA results due to the ambiguity of the term *antinuclear*, that for some professionals includes the cytoplasmic antigens but for some others doesn't. This implies that a same result (e.g. antimitochondrial) may be ANA positive for some, and ANA negative for others. Moreover, it is not clear whether the pattern and the titre should be reported. The national and international recommendations on the reporting of ANA test results show an important lack of consistency (Figure 1) [1,2,3].

Harmonization initiatives

Given all these aspects of variability –biologic, procedural and decision factors –it is quite obvious that there is a need for considering strategies and proposals to reduce such degree of discrepancies. And in this task, all the players in the field of autoimmune in vitro diagnostics (IVD) should be involved, i.e. IVD manufacturers, laboratory staff, physicians, scientific and medical societies, academia and regulatory organizations. In this paper, three initiatives or approaches are presented: a) the first

international consensus in the *nomenclature* and description of ANA patterns; b) the participation of laboratories in *external quality assessment programmes* as a tool to correct deviations in the performance of the participants; and c) building a common criteria to validate and verify the *analytical performance* of the IF IVD tests before manufacturers launch the tests in the market and laboratories acquire them for routine use.

Consensus in ANA patterns

The correct identification of the staining patterns on HEp2 cells is often difficult due to the diversity in the way of naming the patterns and the complexity in the interpretation of the observed images.

The 12th International Workshop on Autoantibodies and Autoimmunity (IWAA) held in Sao Paulo in 2014 saw the first international consensus on nomenclature and descriptions of ANA staining patterns on HEp2: ICAP (International Consensus on ANA Patterns (*www.ANApatterns.org*)

ANA react with molecules typical of the cell nucleus. However, the term also refers to cytoplasmic structures. For this reason, ICAP proposes naming the staining patterns anti-cell patterns (AC) and identify each of them by AC followed by a number from 1 to 28.

According to the expertise of the observer, the patterns are classified in two levels: competent level, easily recognizable and with clinical relevance; and expert level, difficult to recognize unless a considerable experience level has been reached. Patterns are also classified according to the cellular compartment stained: nuclear, cytoplasmic and mitotic patterns (Figure 2).

As an example, figure 3 shows the differences in the nomenclature and description of one cytoplasmic pattern, according to the new nomenclature under the ICAP consensus, and nomenclature inspired in the Cantor project and the glossary published by Wiik in 2010 [4].

Universal performance evaluation

When a quantitative method has to be evaluated, several metrology parameters need to be verified: precision, accuracy, linearity, detection capability, procedure comparison and bias estimation, interferences. However, when dealing with qualitative methods like IF, many of these metrological characteristics become meaningless. Only diagnostic accuracy, which includes diagnostic sensitivity and specificity, and procedure comparison are suitable for verification in IF techniques. The CLSI EP12 guideline can be followed for these purposes [5]. In addition, estima-

tions of the detection capability and the analytical specificity in IF tests (i.e. interferences studies), can also be undertaken.

For the estimation of the *diagnostic accuracy*, the clinical or physiological condition to be studied with the test is taken as a reference. Therefore, a collection of well characterized sera classified by clinical criteria is needed as the true value in the assessment of sensitivity and specificity. As a common rule, these collections include a group of patients with clinically confirmed diagnosis and a group of healthy donors. But it would be advisable to add a group of patients suffering from related diseases, but that are supposed to be negative for the test being evaluated.

When it is not possible to have clinically classified sera, a *method comparison* can be carried out. In this case, the value obtained with the reference test is taken as the true value. In these cases, the result is expressed as the degree of concordance between the two tests.

To approach the estimation of the *detection capability*, there are no guidelines for techniques like IF. However, reference sera are available (ANA-CDC/AF, Centers for Disease Control and Arthritis Foundation) and can be tested and titrated. Therefore, two different tests can be com-
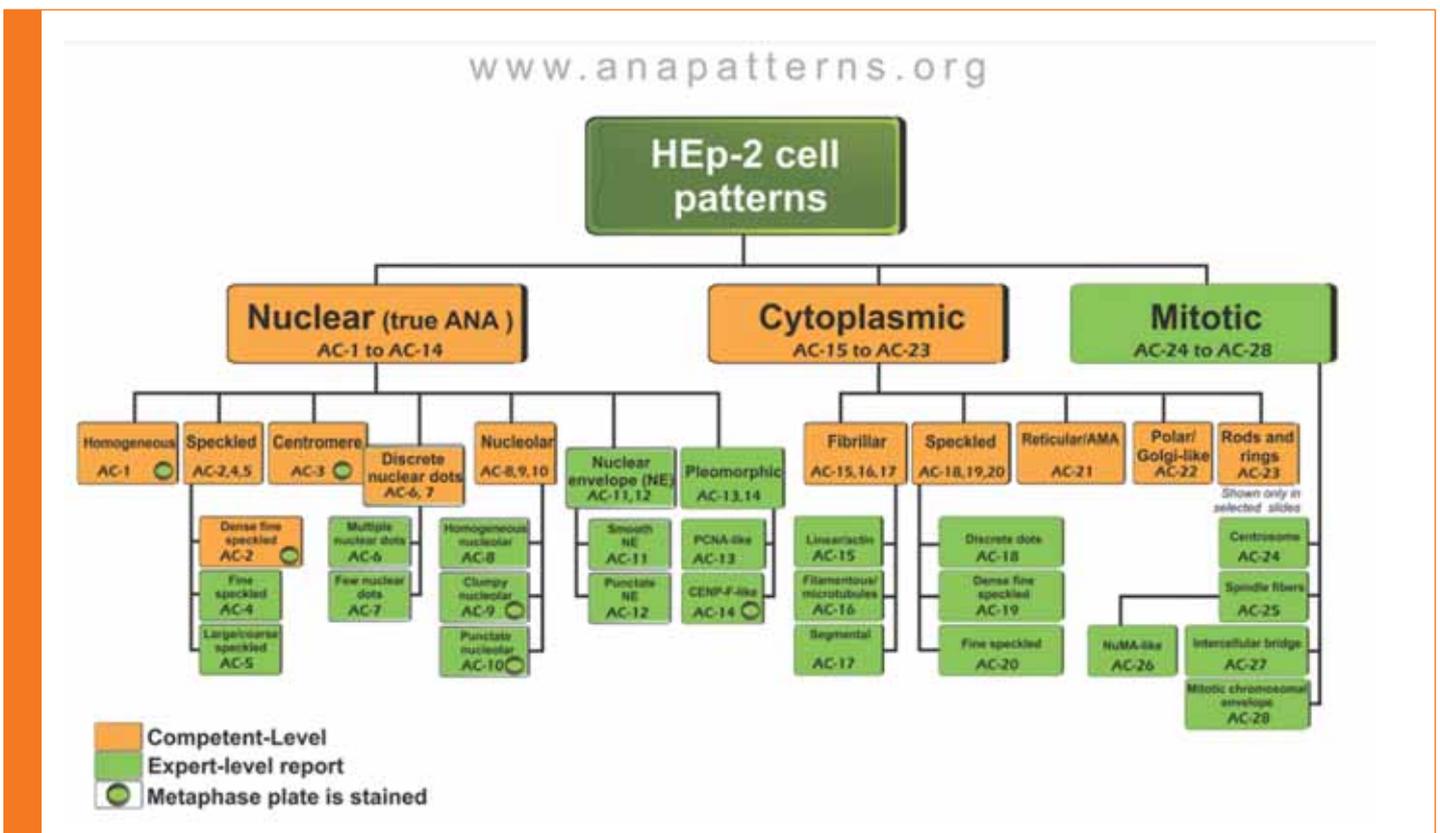

Figure 2. ANA patterns classification tree, according to the compartment stained (nuclear, cytoplasmic and mitotic) and to the expertise level. (www.ANApatterns.org)
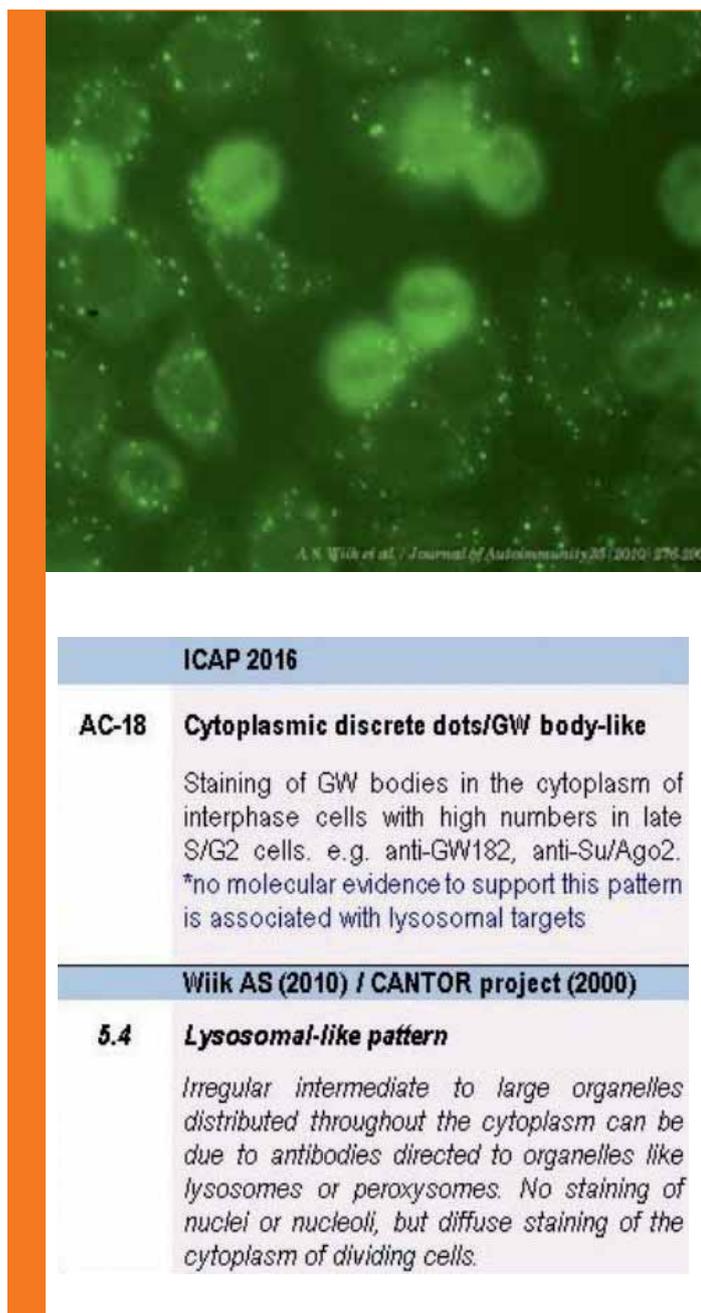
One of the goals of these organizations is the promotion of external quality assessment programs (EQA) to develop an international common concept on standardization with the following aims:

- To provide participants with an objective assessment of the laboratory performance in comparison to other laboratories
- To provide information on the relative performance of the kits and methods available
- To identify the factors associated with good and poor performance

In most of the EQA schemes, 3 important aspects are taken into consideration and usually requested: the slides manufacturer, the conjugate specificity and the starting dilution. When analysing the results of the 2014 BioSystems PREVECAL ANA EQA programme, the diversity of practices of the 95 participating laboratories became apparent. Around 35% of the labs reported the use of a polyvalent conjugate instead of the anti-IgG conjugate which is recommended by all the manufacturers involved.

Also it was noticed that the starting dilution reported by 30% of the participating labs was 1/40, whereas the rest worked at 1/80 and 1/160. The most striking was that only 5% used slides from manufacturers that recommend the 1/40 dilution in their insert. This means that up to 25% of the labs did not follow the recommendations of the manufacturer, which is quite disturbing.

Regarding the staining patterns, the highest number of correct responses corresponded to the ANA homogeneous pattern, followed by the anti-centromere pattern. The highest number of errors was reported with the fine punctate nucleolar pattern, with not a single correct response. When taking a closer look at the overall results, monospecific patterns matched with the most correctly reported sera, while the worst rate of correct responses corresponded to samples with multiple specificities. This fact opens the discussion whether more polyspecific sera should be included in the IF EQA programmes to improve the analyst's skills and correct the poor results obtained with this kind of samples.

References

1. Damoiseaux, J., von Mühlen, C.A., Garcia-De La Torre, I., Carballo, O.G., de Melo Cruvinel, W., Francescantonio, P.L.C., Fritzler, M.J., Herold, M., Mimori, T., Satoh, M., Andrade, L.E.C., Chan, E.K.L., Conrad, K. International Consensus on ANA Patterns (ICAP): the bumpy road towards a consensus on reporting ANA results. Autoimmunity Highlights, 2016; 7:1

2. Sociedad Española de Bioquímica Clínica y Patología Molecular (SEQC). Actualización en el manejo de los anticuerpos anti-nucleares en las enfermedades autoinmunes sistémicas. Documento de la SEQC, Junio 2014.

3. The American College of Rheumatology ad hoc Committee on Immunologic Testing Guidelines. Guidelines for immunologic laboratory testing in the Rheumatic Diseases: An introduction. Arthritis Rheum (Arthritis Care Research) 2002; 47(4): 429-433

4. Wiik, A.S., Hoier-Madsen, M., Forslid, J., Charles, P. & Meyrowitsch, J. Antinuclear antibodies: a contemporary nomenclature using HEp-2 cells. Journal of autoimmunity 2010; 35, 276-290.

5. Clinical and Laboratory Standards Institute (CLSI). User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline_Second Edition EP-12-A2 Vol. 28 No.3 (2008)

The author

Petraki Munujos, PhD
BioSystems S.A. Barcelona, Catalonia, Spain

Figure 3. AC-18 pattern on HEp2 cells. Changes not only affect the name of the pattern, but the description and even the antigenic target recognized by the antibodies responsible of the pattern. [www.ANApatterns.org and 4]

pared based on the amount of antibody they are able to detect in the reference sera.

A common source of systematic error is the one caused by the presence of interfering agents in the sample. The interferences study is performed not only to determine the systematic error, but to be able to prevent it. The study of interferences can be considered as an estimation of the *analytical specificity* of the test.

External Quality Assessment Programmes

There are several national and international organizations devoted to the standardization of immunological methods in the clinical laboratory and the harmonization of the results obtained worldwide. These include the World Health Organization (WHO), the International Union of Immunological Societies (IUIS), the European Autoimmune Standardization Initiative (EASI) or the American College of Rheumatology (ACR).