Assays for measurement of TNF antagonists in practice

Niels Vande Casteele 1, 2

ABSTRACT

Tumour necrosis factor (TNF) antagonist drug exposure is correlated with clinical, endoscopic and pathophysiological outcomes during induction and maintenance therapy. Measuring drug concentrations is therefore a useful tool when treating to target and optimising therapy. One of the main factors leading to suboptimal drug exposure is the formation of antidrug antibodies (ADAs), due to an immunogenic reaction of the immune system towards the non-self protein. The development of ADA does pose important concerns for drug efficacy and for safety as ADAs have been associated with acute infusion reactions, hypersensitivity reactions and serum sickness. Various assays exist to measure serum drug and ADA concentrations, either offered as a service in a specialised laboratory or commercially available as a kit. It is unclear how the performance of these assays relates to each other, until recently various comparative studies were carried out. The majority of these studies show that indeed a good correlation exists between the assays that measure drug, but that absolute concentrations can differ across tests. This is particularly relevant in clinical practice when a specific threshold or drug concentration range is targeted. For ADA assays, drug sensitivity or the ability of the assay to measure ADA in the presence of drug remains an important issue, especially for drugs with a higher dosing frequency. In addition, standardisation across ADA assays is difficult, making it hard to compare quantitative or semiquantitative (low/medium/high) results across assays and across studies.

INTRODUCTION

Over the past decade tumour necrosis factor (TNF) antagonists, such as infliximab, adalimumab, certolizumab pegol and golimumab, have been used to induce and maintain clinical remission in patients with inflammatory bowel diseases. 1 All TNF antagonists currently on the market are parenterally administered monoclonal antibodies or antibody fragments. It is well known that these protein-based drugs exert interindividual and intraindividual variability in drug pharmacokinetics (PK) and pharmacodynamics. 2 One major factor attributing to this variability is the formation of antidrug antibodies (ADAs) in a subset of patients, irrespective of disease indication or degree of humanisation of the antibody. 3 Measuring ADA and monitoring of drug PK are essential for drug development. Therapeutic drug monitoring (TDM) with the measurement of drug and ADA in serum is also used in the clinic as a tool to rationalise treatment decisions at the time of loss of response and to optimise dosing in patients on maintenance therapy with infliximab. 4 5 Many assays and different assay formats exist to measure drug and ADA in serum. The goal of this review article is to give an overview of the different assay formats that are being used for TDM in clinical practice and to review different comparative studies that evaluated the similarities and differences between these assays.

TYPES OF ASSAYS

The three most commonly used assay formats are the ELISA, radioimmunoassay (RIA) and homogenous mobility shift assay (HMSA). Within each format, different assay types exist, owing to differences in the preanalytical steps, capture and detection of antibody.

Enzyme-linked immunosorbent assay

To measure drug concentrations, as coating either a mouse monoclonal ADA 6 7 or TNF is used to capture the TNF antagonist from the serum. In case of the latter, TNF can be directly

1 Division of Gastroenterology, IBD Center, University of California San Diego, La Jolla, California, USA
2 Department of Pharmaceutical and Pharmacological Sciences, Therapeutic and Diagnostic Antibodies, KU Leuven—University of Leuven, Leuven, Belgium

Correspondence to
Dr Niels Vande Casteele, Division of Gastroenterology, School of Medicine, University of California San Diego, IBD Center, 9500 Gilman Drive #0956, La Jolla, CA 92039, USA; nvandecasteele@ucsd.edu

Received 2 February 2016
Accepted 17 February 2016
Published Online First
8 March 2016
coated, or captured by a monoclonal antibody against TNF that was first coated on the plate. The theoretical advantage of the latter is that TNF is always oriented in the same way. As detecting antibody, either antihuman IgG, monospecific polyclonal ADA (from immunised goats or rabbits) or monoclonal ADA (murine origin) is used. The advantage of a monoclonal or monospecific polyclonal ADA is the specificity towards the TNF antagonist, resulting in lower aspecific binding and a lower risk for false positives.

To measure ADA concentrations, as coating and detection antibody, typically the TNF antagonist itself is used. The drug then forms a bridge between the capture and detection antibody. Others have used an anti-λ antibody as ADA detecting antibody, whereby the ADA does not require two free binding sites and the assay is less drug sensitive (see section ‘Differences in performance’).

**Homogenous mobility shift assay**

To measure drug concentrations, an excess of fluorescent-labelled TNF is added to the serum sample after which TNF in complex with drug can be separated from free TNF using high pressure liquid chromatography in combination with size exclusion chromatography. The degree of fluorescence of the fluorescent-labelled TNF in complex with the TNF antagonist is relative to the concentration of TNF antagonist present in the serum. ADA concentrations are measured in a similar way, except that fluorescent-labelled TNF antagonist is used instead of fluorescent-labelled TNF.

**Radioimmunoassay**

To measure drug concentrations, serum is typically diluted with protein A, after which non-bound serum components are washed off and radiolabelled TNF is added. After incubation, non-bound radiolabelled TNF is washed away and agarose-bound radioactivity is measured. ADAs are measured in a similar way, except radiolabelled drug or F(ab)2 is used instead of radiolabelled TNF.

**Others**

A gene reporter assay was developed by Lallemand et al in which both the TNF antagonist as well as anti-idiotype ADA can be quantified using human erythroleukaemic K562 cells transfected with an NFκB regulated firefly luciferase reporter gene construct. Briefly, TNF will activate the NFκB pathway leading to luciferase production (readout in relative luciferase units, RLUs). If TNF activity is blocked by TNF antagonist presence, RLU decreases whereas if anti-idiotype ADA is present, this will neutralise the TNF antagonist activity, resulting in free TNF and an increase in RLU. A cell-based assay based on the same concept was developed by Gils et al, in which human fibrosarcoma cells were used that express interleukin-6 via the NFκB pathway, which is stimulated upon addition of TNF to the cell medium. The advantage of the latter two cell-based assays is that they mimic conditions in vivo by measuring TNF-mediated effects on TNF receptor-positive cells.

An electrochemiluminescence immunoassay was used in the Programme evaLuating the Autoimmune disease iNvestigational drug cT-p13 in ankylosing spondylitis patients (PLANETAS) and Programme evaLuating the Autoimmune disease iNvestigational drug cT-p13 in rheumatoid arthritis patients (PLANETRA) studies, to compare the immunogenicity between infliximab and the infliximab biosimilar CT-P13. The assay uses the technology from Meso Scale Discovery (MSD; Rockville, Maryland, USA), where serum samples are first acidified and then neutralised together with ruthenylated infliximab and biotinylated infliximab. ADAs that are bound to both sulfo-tagged and biotinylated infliximab can then bind to the streptavidin-coated MSD plate and subsequently quantified by measuring the chemiluminescent signal. The assay thus relies on the bivalency of the ADA, but is less drug sensitive because of the acid dissociation of endogenous drug–ADA complexes (see section ‘Differences in performance’).

An important trend for the future will be the transformation of some of these assays to a point of care format, allowing for a quantitative, rapid readout of the result. Van Stappen et al assessed the feasibility of measuring infliximab in serum with two methods: fibre-optic surface plasmon resonance (FO-SPR) and lateral flow test (LFT). FO-SPR measures in real-time local refractive index changes due to interactions between infliximab in solution and a specific anti-infliximab monoclonal antibody, functioning as a capture antibody that is immobilised on the optical fibre. The signal was amplified by employing gold nanoparticles functionalised with a specific anti-infliximab monoclonal antibody as detecting antibody. The LFT used TNF as capture antigen and a specific anti-infliximab monoclonal antibody as detecting antibody. The colorimetric signal was enhanced with conjugated gold nanoparticles and measured with a portable LFT reader.

**Differences in performance**

There are known advantages for measuring ADA in a fluid phase system (RIA, HMSA) over the solid phase system (ELISA): endogenous IgG4 ADA can exchange halve molecules in vivo leading to monospecific ADA that cannot be detected in the double-antigen ELISA (requiring bivalency). Fluid phase techniques can detect low affinity ADA because of less wash steps, in contrast to solid phase assays. From a practical perspective, fluid phase assays are often more labour intensive and in case of the RIA (using radioisotopes), less sustainable.
Drug sensitivity, or the inability to detect ADA in the presence of drug, is an important limitation of first generation ADA assays. Each ADA assay format (ELISA, RIA and HMSA) is drug sensitive to a certain extent as the majority of these assays rely on the capture of the ADA by drug. Most of the second generation, drug resistant, ADA assays therefore include a preanalytical step using acid to lower the pH of the sample to pH 2.5-3 so the complex of ADA in the sample with endogenous drug is dissociated and the ADA can subsequently be captured by the drug in the assay. There are now several drug-tolerant ADA ELISA, RIA and HMSA in use.

**COMPARISON OF ASSAYS**

Various studies compared assays that measure infliximab, adalimumab, golimumab and etanercept drug and ADA concentrations.

**Drug**

For infliximab drug concentrations, linear correlations (Pearson r) between the various assays tested were found to be good and ranged from 0.73 to 0.99, irrespective of the assay format (table 1). Similarly, non-linear correlations (Spearman r) ranged from 0.95 to 0.97. Good correlations were also observed for adalimumab, golimumab and etanercept. One study specifically looked at the reactivity of various infliximab assays to measure infliximab biosimilar (CT-P13) concentrations and also here good correlations (range 0.86–0.96) between the assays were found.

The correlation between two assays does not take into account the difference in absolute concentration measured. Agreement between two assays can be assessed graphically in a Bland-Altman plot or expressed quantitatively by calculating the intraclass correlation coefficient (ICC). The ICC ranged from 0.59 to 0.98 across comparative studies and showed that there is a discrepancy in absolute concentration reported by some of the assays. In the study by Steenholdt et al mean differences in infliximab concentrations of up to −3.44 μg/mL were found when an inhouse reporter gene assay was compared with the Prometheus ELISA. This can impact the implementation of TDM in clinical practice, when specific thresholds or drug concentration ranges are targeted. Differences in analytical sensitivity were found between assays, although the ability to measure extremely low drug concentrations is likely clinically irrelevant and all methods are able to detect subtherapeutic drug concentrations. However, regarding specificity, one blinded round robin experiment performed at three laboratories throughout Belgium and the Netherlands reported the detection of false-positive infliximab concentrations in 11 samples using one assay, including quality control samples from healthy volunteers spiked with different amounts of antibody to infliximab.

**Antidrug antibodies**

When comparing ADA titres across assays, linear and non-linear correlations were acceptable and ranged from 0.71 to 0.99 and from 0.54 to 0.93, respectively (table 2). Because of different measures of quantification and a lack of international standards, the agreement between ADA assays cannot be calculated. Differences in sensitivity were mainly attributed to the interference of drug in the detection of ADA. Assays that allow for the detection of low affinity ADA and IgG4 are typically more sensitive, but the clinical relevance of detecting these subtypes of ADA remains to be shown. Indeed, a comparison of a bridging ELISA, RIA and HMSA found these assays to provide overall similar guidance for clinical practice in most patients with loss of response.

**DISCUSSION**

A wide range of assays and assay formats were developed for therapeutic drug and immunogenicity monitoring of TNF antagonists. The majority of the studies comparing the performance of these assays show a good correlation in drug concentrations and ADA titres. Nonetheless, the agreement between the absolute drug concentrations reported by some of the assays is poor. This can have an impact on clinical practice if a specific threshold or concentration range is targeted, especially when using a different assay than the one that was used to establish the threshold. It is not feasible to suggest one universal assay as various tests have shown a clinical correlation and it is therefore up to the discretion of the physician or laboratory technician to choose the most appropriate assay. Cost and turnaround time are important factors to take into account as well. Eventually, the field will evolve towards the use of rapid assays that will allow fast and quantitative determination of the drug concentration at or close to site of point of care, allowing for an incorporation of the pharmacological result in the treatment decision ad hoc. This would also facilitate the incorporation of dosing-based-on-exposure treatment strategies to clinical practice.

Assays for measuring drug concentrations can be standardised since the reference product is readily available. False-negative results have been described for some assays, but the clinical relevance of measuring extremely low drug concentrations is nil. False-positive results, on the other hand, are an important problem as these might erroneously impact treatment decisions. Assays for measuring ADA are harder to compare as different standards and outcome measures are used to quantify the result. This has implications for treatment algorithms as often ambiguous terms such as ‘low’, ‘intermediate’ or ‘high’ ADA are used that...
are dependent on the assay. Not the (semi)quantitative result or threshold and the range cannot be compared across assays. There is a need for international harmonisation of assays that measure ADA, for example by using a monoclonal ADA standard that can be made recombinantly and of which the concentration can be determined quantitatively as suggested by Gils et al.26 However, as shown by van Schouwenburg et al.,45 the output of ADA assays depends on antibody characteristics such as affinity and avidity in non-uniform ways, thereby inherently limiting the comparability of results across assays. This might preclude universal standards from being used across various assay formats, but may hold value within a certain ADA assay category (eg, bridging ELISA). Cut-offs could then be compared across assays, as well as the lower and higher limit of detection, indicating the measurement range of the ADA assay.

An immunogenic response towards a biologic typically evolves from low titre, affinity and avidity ADA to stronger binding immunoglobulins of higher titres.46 There is insufficient knowledge about the maturation of this immune response and why some patients develop transient and others develop persistent ADA.47–49 With the advent of biosimilars, drug-tolerant ADA assays will become standard to compare immunogenicity rates between molecules, especially for those drugs that are administered with a higher frequency.50

One of the limitations of this review and the various comparative studies that were performed is that there are currently no reporting requirements when changes are made to the set-up or protocol of assays for TDM. Various inhouse or commercially available kits have been optimised and thus prior comparative results may not be representative of the current performance of these assays. One should

---

### Table 1 Drug concentration comparative studies

<table>
<thead>
<tr>
<th>Drug</th>
<th>Comparison</th>
<th>Pearson r correlation</th>
<th>Spearman r correlation</th>
<th>Agreement ICC (95% CI)</th>
<th>Mean difference, μg/mL (95% CI)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infliximab</td>
<td>Sanquin ELISA–Leuven ELISA</td>
<td>0.91</td>
<td>—</td>
<td>0.91 (0.86 to 0.95)</td>
<td>—</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Sanquin ELISA–Theradiag</td>
<td>0.83</td>
<td>—</td>
<td>0.73 (0.58 to 0.83)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leuven ELISA–Theradiag</td>
<td>0.73</td>
<td>—</td>
<td>0.59 (0.39 to 0.73)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Infliximab</td>
<td>Biomonitor RIA–Prometheus ELISA</td>
<td>0.95</td>
<td>0.97</td>
<td>0.76 (0.29 to 0.90)</td>
<td>−3.12 (−3.98 to −2.25)</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Biomonitor RIA–Prometheus HMSA</td>
<td>0.96</td>
<td>0.97</td>
<td>0.82 (0.41 to 0.93)</td>
<td>−2.48 (−3.18 to −1.77)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biomonitor RIA–RGA</td>
<td>0.94</td>
<td>0.96</td>
<td>0.94 (0.90 to 0.96)</td>
<td>0.32 (−0.09 to 0.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RGA–Prometheus ELISA</td>
<td>0.91</td>
<td>0.95</td>
<td>0.72 (0.23 to 0.88)</td>
<td>−3.44 (−4.39 to −2.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RGA–Prometheus HMSA</td>
<td>0.95</td>
<td>0.95</td>
<td>0.80 (0.30 to 0.92)</td>
<td>−2.80 (−3.54 to −2.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prometheus ELISA–Prometheus HMSA</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96 (0.94 to 0.98)</td>
<td>0.64 (0.15 to 1.12)</td>
<td></td>
</tr>
<tr>
<td>Infliximab</td>
<td>ELISA–RIA</td>
<td>0.98</td>
<td>—</td>
<td></td>
<td>—</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>ELISA–RGA</td>
<td>0.99</td>
<td>—</td>
<td></td>
<td>—</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RIA–RGA</td>
<td>0.97</td>
<td>—</td>
<td></td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Infliximab</td>
<td>Promonitor ELISA–Sanquin ELISA</td>
<td>0.98</td>
<td>—</td>
<td></td>
<td>—</td>
<td>38</td>
</tr>
<tr>
<td>Infliximab</td>
<td>Promonitor ELISA V1–Promotion ELISA V2</td>
<td>–</td>
<td>–</td>
<td></td>
<td>–</td>
<td>41</td>
</tr>
<tr>
<td>Infliximab</td>
<td>Promonitor ELISA–Immunodiagnostic ELISA</td>
<td>–</td>
<td>0.94</td>
<td>–</td>
<td>2.09 (1.41 to 2.77)</td>
<td>39</td>
</tr>
<tr>
<td>Infliximab</td>
<td>Leuven ELISA–Leuven FO-SPR</td>
<td>0.998</td>
<td>–</td>
<td>0.98 (N/A)</td>
<td>–</td>
<td>30</td>
</tr>
<tr>
<td>Infliximab (CT-P13)</td>
<td>Matriks ELISA–Theradiag</td>
<td>–</td>
<td>0.96</td>
<td></td>
<td>–</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Matriks ELISA–R-Biopharm</td>
<td>–</td>
<td>0.86</td>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theradiag ELISA–R-Biopharm</td>
<td>–</td>
<td>0.92</td>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Adalimumab</td>
<td>Promonitor ELISA–Sanquin ELISA</td>
<td>0.99</td>
<td>–</td>
<td></td>
<td>–</td>
<td>38</td>
</tr>
<tr>
<td>Etanercept</td>
<td>Promonitor ELISA–Sanquin ELISA</td>
<td>0.98</td>
<td>–</td>
<td></td>
<td>–</td>
<td>38</td>
</tr>
<tr>
<td>Golimumab</td>
<td>Promonitor ELISA–Sanquin ELISA</td>
<td>0.98</td>
<td>0.98</td>
<td>−0.03 (−0.09 to 0.03)</td>
<td>–</td>
<td>40</td>
</tr>
</tbody>
</table>

FO-SPR, fibre-optic surface plasmon resonance; HMSA, homogenous mobility shift assay; ICC, intraclass correlation coefficient; RGA, reporter gene assay; RIA, radioluminoassay.
perform rigorous validation experiments and compare the performance of the optimised versus previous protocol, for example when switching from a polyclonal to a monoclonal detecting antibody in the assay.17

In conclusion, TDM assays can be standardised and although a good correlation is observed for the majority of assays, a difference in absolute drug concentration can be seen. Because of a lack of universal standards, ADA titres cannot be quantitatively compared across assays. For most treatment algorithms based on TDM, measuring the serum drug concentration is the first step. Subsequent measurement of ADA can be useful to explain undetectable or low drug concentrations. The dynamics of ADA in relation to drug concentrations can be more informative than a sole ADA measurement because of known limitations of sensitivity of the ADA assay in the presence of drug and incomparable cut-offs or thresholds across assays.

**Contributors**
NVC is the sole contributor to the design of the work, the acquisition and interpretation of the data, drafting of the manuscript and critically revising it for important intellectual content. NVC approves the final published version of the manuscript and critically revising it for important intellectual content. NVC approves the final published version of the manuscript and critically revising it for important intellectual content. NVC approves the final published version of the manuscript and critically revising it for important intellectual content. NVC approves the final published version of the manuscript and critically revising it for important intellectual content.

**Funding**
NVC is a Postdoctoral Fellow of the Research Foundation—Flanders (FWO), Belgium; grant number 1260714N.

**Competing interests**
NVC received speakers fees from AbbVie and consultancy fees from MSD, Janssen Biologics, Pfizer, UCB and Takeda.

**Provenance and peer review**
Commissioned; internally peer reviewed.

**REFERENCES**


