Comparative evaluation of three commercial Toxoplasma-specific IgG antibody avidity tests and significance in different clinical settings

Branko Bobić, Ivana Klun, Marija Vujanić, Aleksandra Nikolić, Vladimir Ivović, Tijana Živković and Olgica Djurković-Djaković

National Reference Laboratory for Toxoplasmosis, Serbian Centre for Parasitic Zoonoses, Institute for Medical Research, University of Belgrade, Belgrade, Serbia

Determination of the avidity of specific IgG antibodies has become a generally accepted diagnostic aid for dating Toxoplasma infection. In this study, the Labsystems, VIDAS and EUROIMMUN Toxoplasma IgG avidity assays were compared on a series of 133 Toxoplasma IgG- and IgM-positive sera from symptomatic patients (n=28), from pregnant (n=43) and non-pregnant (n=26) women, and on 18 IgG-positive and IgM-negative sera from chronically infected patients. The results showed excellent concordance between the Labsystems and VIDAS tests in both the IgM-positive (r=0.82, k=0.771) and IgM-negative (k=0.609) sera, whilst the agreement of the EUROIMMUN assay with both the Labsystems and VIDAS tests in the IgM-positive sera was moderate (k=0.575 and k=0.525, respectively) and in the IgM-negative sera was poor (k=0.000). Analysis of the kinetics of the maturation of avidity in 13 patients in whom follow-up sera were available showed that, despite a general trend of maturation, in two patients the avidity did not become high during 6 and 11 months of follow-up. In view of the clinical setting, in the symptomatic patients, despite one case of complete discrepancy and five cases of partial discrepancy, the Labsystems and VIDAS tests were in almost perfect agreement (k=0.812), whilst the agreement in pregnant and non-pregnant women was substantial (k=0.754 and k=0.708, respectively). In conclusion, the Labsystems and VIDAS tests are equally reliable for the measurement of Toxoplasma IgG avidity; the choice of test should depend on the laboratory set-up. The EUROIMMUN test may be an acceptable alternative in resource-limited settings, but should be used prudently.

INTRODUCTION

The discovery that the avidity of specific antibodies as a measure of their affinity for the antigen increases over time (Brown et al., 1984) has resulted in its clinical application for the diagnosis of a number of infectious diseases including viral infections and toxoplasmosis (Hedman & Seppälä, 1988; Hedman et al., 1989, 1993). In less than two decades, determination of the avidity of specific antibodies has become a standard supplementary tool for the dating of Toxoplasma infection. In symptomatic infection, IgG avidity is helpful in the differential diagnosis of lymphadenopathy (Montoya et al., 2004); however, its major use has been in the discrimination between recent and chronic asymptomatic infection in pregnancy (Lappalainen et al., 1993; Jenum et al., 1997; Logar et al., 1999; Liesenfeld et al., 2001). Because the clinical management of pregnancy depends on knowledge of the time of infection versus conception, determination of specific IgG avidity is particularly important in the case of detection of specific IgM antibodies, which may be present for a prolonged period of time (Bobić et al., 1991; Montoya & Remington, 1995). In such cases, a consensus has been reached on the value of avidity to help rule out infections acquired in the last few months (Liesenfeld et al., 2001; Remington et al., 2004; Petersen et al., 2005; Petersen, 2007; Flori et al., 2008).

Due to the technical ease with which the specific IgG avidity can be determined (by use of a dissociating agent such as urea to break up weak antigen–antibody associations), many laboratories developed in-house assays for its measurement, which was later followed by the introduction of a number of commercial tests. These tests differ in the antigen source and preparation, and in the use of serial or a single predetermined serum dilution, or one containing a specified amount of Toxoplasma IgG antibodies, as well as in the calculation and interpretation of results. Such variables may influence performance. However, few studies comparing commercial tests have been published to date (Barberi et al., 2001; Alvarado-Esquível et al., 2002; Horvath et al., 2005; Petersen et al., 2005; Soula et al.,...
2007). We thus performed a study to compare three such tests.

**METHODS**

**Study design.** The performance of three commercial assays for the measurement of the avidity of *Toxoplasma*-specific IgG antibodies was examined in a total of 151 serum samples (133 sera positive for both specific IgG and IgM, and 18 IgG-positive, IgM-negative sera).

In addition, the kinetics of IgG avidity maturation was analysed in a subgroup of 13 patients in which follow-up sera were available.

Finally, the significance of the avidity results was analysed in the clinical context.

**Patients and sera.** The study series comprised a panel of 151 *Toxoplasma* IgG-positive serum samples obtained from 115 immunocompetent patients examined for toxoplasmosis at the National Reference Laboratory for Toxoplasmosis at the Institute for Medical Research, University of Belgrade, Serbia, between May 2004 and December 2006.

Of the 151 sera, 133 from 97 patients were positive for *Toxoplasma* IgM, whilst the 18 *Toxoplasma* IgM-negative sera originated from 18 randomly selected patients with chronic infection. Among the IgM-positive patients, 38 follow-up serum samples were available from 13 patients with a low or borderline avidity in their first sample; 33 of these (two samples in eight patients, three samples in four patients and five samples in one patient) were included in the study of the comparison of the tests. An additional five follow-up samples (from three patients), which arrived after the end of the comparative study, were analysed only by VIDAS.

According to the clinical setting, the IgM-positive patients comprised 28 with clinical infection (lymphadenopathy), whilst 69 were asymptomatic women tested for obstetric reasons. Of the latter, 43 women were pregnant at the time of testing and the remaining 26 were examined for history of miscarriage, premature delivery, etc., or as a pre-pregnancy control.

The study was approved by a local (Institute for Medical Research) Ethics Committee.

**Sero logical tests**

**Detection of specific IgM antibodies.** Specific IgM antibodies were detected using the commercial IgM-immunosor bent agglutination assay (bioMérieux) and/or the VIDAS Toxo IgM assay (bioMérieux), as recommended by the manufacturer. For the IgM-immunosorbent agglutination assay, results were expressed as an index on a scale of 0–12, where 0–5 was considered negative, 6–8 was borderline and 9 was positive. The cut-off value for the VIDAS Toxo IgM assay was 0.55, results between 0.55 and 0.65 were considered borderline, whilst results above 0.65 were considered positive.

**Avidity assays.** The avidity of the specific IgG antibodies was measured using three commercial assays, as follows: the *Toxoplasma gondii* IgG avidity enzyme immunoassay (Ani Labsystems), the VIDAS Toxo IgG avidity assay (bioMérieux) and an ELISA adapted for IgG avidity determination (EI 2410-9601-1 G; EUROIMMUN). Each was performed according to the instructions of the respective manufacturer. The avidity results for the Labsystems assay were calculated according to the distances (in mm) between the lines obtained by plotting the titration curves, and for the EUROIMMUN test by calculating the ratio of the absorbance values of the sample obtained with and without urea treatment. In contrast, the VIDAS test is a fully automated assay where the results are produced directly by the system software. Results were expressed as avidity percentages (Labsystems and EUROIMMUN) or indices (VIDAS). The respective manufacturer-recommended cut-off values with the interpretation of the results, classified as high, low or borderline avidity, are presented in Table 1.

**Data analysis.** The results of avidity tests were analysed per pair of tests, with borderline results grouped with either the low-avidity or the high-avidity results. Test results in this dichotomous form were classified into two-by-two contingency tables and the overall (observed) agreements were calculated. The concordance between the tests was determined by Pearson’s correlation coefficient (r) and, in addition, by κ statistics; κ values of 0.01–0.20 indicated slight agreement, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial and 0.81–0.99 almost perfect agreement (Landis & Koch, 1977).

In addition, where appropriate, the results obtained in different pairs of tests were analysed as being in agreement (low, borderline or high in both tests), partially discrepant (low or high in one, and borderline in the other test) or completely discrepant (low in one and high in the other test).

**RESULTS**

The results of a comparative investigation of three commercial tests for the measurement of specific IgG

<table>
<thead>
<tr>
<th>Test</th>
<th>Low avidity</th>
<th>Borderline avidity</th>
<th>High avidity</th>
<th>Interpretation for high avidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ani Labsystems</td>
<td>&lt;15%</td>
<td>15–30%</td>
<td>&gt;30%</td>
<td>Excludes infection in the last 3 months</td>
</tr>
<tr>
<td>EUROIMMUN</td>
<td>&lt;40%</td>
<td>40–60%</td>
<td>&gt;60%</td>
<td>Unspecified</td>
</tr>
<tr>
<td>VIDAS*</td>
<td>&lt;0.2</td>
<td>0.2–0.3</td>
<td>&gt;0.3</td>
<td>Excludes infection in the last 4 months</td>
</tr>
</tbody>
</table>

*Corresponds to <20%, <20–30% and >30%.

Table 1. Cut-off values and result interpretation of the commercial IgG antibody avidity tests as recommended by the respective manufacturer
antibody avidity on a series of 133 sera positive for both *Toxoplasma* IgG and IgM are presented in Fig. 1. All three tests analysed per pair correlated well, with a strong correlation between Labsystems and VIDAS tests \((r=0.82)\), and good correlation between the EUROIMMUN and both Labsystems and VIDAS tests \((r=0.69\) and \(0.66\), respectively). Similarly, analysis of the results by \(\kappa\) statistics showed at least moderate agreement between all pairs of tests (Table 2). However, whilst the EUROIMMUN assay showed only moderate agreement with both the Labsystems and VIDAS tests, the agreement between the Labsystems and VIDAS tests was substantial. The agreement depended somewhat on whether the borderline values were grouped with the low- or high-avidity values. Although all statistical calculations are presented, it should be borne in mind that, given that interpretation of test results is generally limited to high-avidity values ruling out recent infections (precisely how long depends on the particular test), grouping of borderline values with the high-avidity values does not seem meaningful. When borderline results were grouped with the low-avidity ones, the EUROIMMUN test agreed similarly with both Labsystems \((\kappa=0.575)\) and VIDAS \((\kappa=0.525)\) tests. When these were analysed together with the high-avidity results, the EUROIMMUN test agreed significantly better with the Labsystems test \((\kappa=0.648)\) than with the VIDAS test \((\kappa=0.496)\) (a finding that seems to bear no real significance in view of the above consideration). In contrast, the Labsystems and VIDAS tests agreed substantially irrespective of the grouping \((\kappa>0.7)\).

To examine the agreement of the avidity tests in a situation in which IgG avidity is expected to be high, a group of 18 specific IgM-negative sera was analysed (Table 3). Although no low-avidity value was obtained in any of the tests, not all samples were shown to be of high avidity. Thus, with the EUROIMMUN test, there were nine (50%) borderline results, whilst the Labsystems and VIDAS tests gave four (22%) and two (11%) borderline results, respectively. Accordingly, the statistical agreement of the EUROIMMUN test with either of the latter tests was not above the probability of chance \((\kappa=0.000, P=1)\). In contrast, the agreement between the Labsystems and VIDAS tests was substantial \((\kappa=0.609)\). In two patients, the tests were partially discrepant in that the Labsystems result was still borderline and the VIDAS result was high, whilst in another two patients, both tests were borderline. These latter two patients obviously indicated delayed IgG avidity maturation.

Having thus established the Labsystems and VIDAS tests to be superior to the EUROIMMUN assay, only results from these two tests were considered in further analysis. The kinetics of IgG avidity maturation was examined in a series of patients who presented with low or borderline avidity. This subgroup comprised 13 patients whose initial sample was found to be of low (seven and ten patients in the Labsystems and VIDAS tests, respectively) or borderline (six and three patients in the Labsystems and VIDAS tests,

**Fig. 1.** Correlation of *Toxoplasma* IgG antibody avidity levels determined in a series of 133 IgM-positive patient sera, using the Ani Labsystems (AniL) and VIDAS (a), Ani Labsystems and EUROIMMUN (Elmmun) (b) and VIDAS and EUROIMMUN (c) tests. Data for individual sera fit a linear regression line as shown.
respectively) avidity (Fig. 2). The patients were followed for a mean of 3.38 months (range 1–12 months) for the 33 samples analysed using both the Labsystems and VIDAS tests, and 5.46 months (range 1–14 months) when the five later follow-up samples (analysed by VIDAS test alone) were added. Whilst there was a general trend of maturation, the Labsystems test showed that, of the four patients followed for greater than 3 months, three matured to high avidity (two from low and one from borderline) and one matured to borderline avidity (however, this patient was not tested beyond 4 months). On the other hand, the VIDAS test showed that, of the seven patients followed for more than 4 months, three patients went from low to high avidity, one from low to borderline and two from borderline to high, whilst one remained borderline throughout the 6 months of follow-up. This patient, a pregnant woman, represents an actual case of delayed maturation of IgG antibody avidity. In addition, a similar phenomenon was noted in one symptomatic patient (with three samples) in whom the avidity was very low in the first two samples taken 3 months apart, whilst the final sample taken 8 months after the second one was borderline.

Finally, the avidity results in the IgM-positive patients were analysed in the clinical context, i.e. according to whether the infection was clinically manifest or not (Table 4). The substantial agreement between the tests as shown in Table 2 varied according to the clinical setting. In the 44 samples analysed from 28 patients with clinically manifest acute toxoplasmosis, complete discrepancy was found in a single sample (2.3 %). Five more samples were partially discrepant (11.4 %), but the agreement was nevertheless statistically almost perfect ($\kappa = 0.812$). More importantly, of these five, the partial discrepancy in four where the findings in the VIDAS test were still low (in three) and borderline (in one) whilst being borderline and high in the Labsystems test may be attributable to a 1-month difference in interpretation of high-avidity results (4 and 3 months, respectively). Although there was no instance of complete discrepancy in the asymptomatic patients, the number of partial discrepancies amounted to 15.7 % in pregnant women and 26.3 % in non-pregnant women. This resulted in a statistically lower although still substantial agreement ($\kappa = 0.754$ and $\kappa = 0.708$, respectively). The proportion of low-avidity findings was highest in symptomatic patients (47.4 %), followed by pregnant (20.9 %) and non-pregnant (3.6 %) women. Also, when the borderline findings were added to those of low avidity, the distribution trend remained similar (50 vs 32.6 vs 25 % in symptomatic patients, pregnant women and non-pregnant women, respectively). This is not surprising as it reflects the fact that symptomatic patients tend to be examined serologically earlier in the course of infection, whilst pregnant women...
women tend to be more regularly examined than those tested for other obstetric reasons.

**DISCUSSION**

We compared three commercial tests for the measurement of *Toxoplasma* IgG avidity, including a fully automated assay on the VIDAS system, a well-known manual test by Labsystems and a less well-known test by EUROIMMUN. This latter test was included because it is readily available and, due to its competitive price, is often used in Serbia. It should be noted that, unlike the study of Barberi et al. (2001), we compared the tests only per pair, in order not to mask potential concordances between two of the three tests as occurred in their study. An excellent concordance was found in the results of the Labsystems and VIDAS tests in both *Toxoplasma* IgM-positive and IgM-negative patients. The strong correlation (r=0.82) in the IgM-positive patients was effectively the same as that (r=0.8) in the single previous study by Alvarado-Esquível et al. (2002) comparing these two tests, whereas the overall agreement of 89% in our larger series was better than the value of 75% (48/64) that they obtained. Our study is the first to date to evaluate the EUROIMMUN test. The concordance of the EUROIMMUN test with either of the other two tests in the *Toxoplasma* IgM-positive patients was weaker but statistically good; however, this was undermined by a poor agreement in the IgM-negative patients. Although IgM-negative patients are not normally the target population for IgG avidity determination, its measurement in a small group of randomly selected patients with chronic infection showed somewhat disappointing results, particularly with the EUROIMMUN test; although no complete discrepancies were obtained, as many as 50% of the IgM-negative patients had borderline avidity, compared with 22 and 11% in the Labsystems and VIDAS tests, respectively. These findings raise the issue of how to set the cut-off point in different tests, as discussed previously (Flori et al., 2004, 2008; Horvath et al., 2005).

**Fig. 2.** Kinetics of IgG antibody avidity maturation over time in a series of 13 patients as measured by the Ani Labsystems (a) and the VIDAS (b) avidity assays. The cut-off points for low, borderline and high avidity are indicated by horizontal lines.

**Table 4.** IgG avidity findings as measured by the Ani Labsystems and VIDAS tests on a series of 133 *Toxoplasma* IgM-positive sera according to clinical setting

<table>
<thead>
<tr>
<th>Clinical setting</th>
<th>$\kappa$ value</th>
<th>$P$ for $\kappa$ value</th>
<th>Agreement</th>
<th>Partial discrepancy</th>
<th>Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td></td>
<td>Samples ($n$)</td>
<td>Total (%)</td>
<td>Samples ($n$)</td>
</tr>
<tr>
<td></td>
<td>$L/L$</td>
<td>$B/B$</td>
<td>$H/H$</td>
<td>$L/B$</td>
<td>$B/L$</td>
</tr>
<tr>
<td>Lymphadenopathy</td>
<td>0.812</td>
<td>0.000</td>
<td>18 1 19 38 (86.3%)</td>
<td>0 3 1 1 5 (11.4%)</td>
<td>0 1 1 (2.3%)</td>
</tr>
<tr>
<td>Pregnant women</td>
<td>0.754</td>
<td>0.000</td>
<td>9 5 29 43 (84.3%)</td>
<td>3 1 1 3 8 (15.7%)</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Non-pregnant women</td>
<td>0.708</td>
<td>0.000</td>
<td>1 6 21 28 (73.7%)</td>
<td>4 0 1 5 10 (26.3%)</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>28 12 69 109 (81.9%)</td>
<td>7 4 3 9 23 (17.3%)</td>
<td>0 1 1 (0.8%)</td>
</tr>
</tbody>
</table>

B. Bobić and others 362 Journal of Medical Microbiology 58
According to the clinical setting, the Labsystems and VIDAS tests showed the best agreement, as expected, in symptomatic patients, and, in descending order but still substantially, in pregnant and non-pregnant women. Symptomatic patients are those in which the time of onset of infection is known, which is a prerequisite for the evaluation of the performance of *Toxoplasma* IgG avidity tests. However, even among these patients, of the 44 serum samples tested, there was one sample with complete discrepancy and five samples with partial discrepancy, explicable at least in part by the 1-month difference in the high-avidity results interpretation. Among samples in agreement in both tests, the proportion of low-avidity findings in symptomatic patients was 47.4 %, and of low together with borderline was 50 %. This reflects a delay in presentation for serology after the onset of infection in about half of the symptomatic patients in our series.

In the asymptomatic patients, the time of infection is obviously unknown and such infections are discovered only upon presentation for serology. In the specific IgM-positive asymptomatic women in our series, the proportion of low-avidity findings was 20.9 % in pregnant women and only 3.6 % in non-pregnant women, and rose to 32.6 and 25 %, respectively, when borderline findings were added to those of low avidity. This decreased proportion of diagnosed recent infections in non-pregnant versus pregnant women probably reflects the fact that pregnant women are examined more regularly than those tested for other obstetric reasons; the latter, particularly if tested following a pathological outcome of a previous pregnancy, often present for serology months later, which is why, even in cases where toxoplasmosis may have been the cause of an adverse outcome of pregnancy, they are not tested early enough during the course of infection. Such cases also underline the high prevalence of specific IgM antibodies that do not indicate infection in the immediate past.

Analysis of the kinetics of IgG avidity maturation showed delayed maturation in 15.4 % of patients (2/13). Further evidence for delayed maturation was presented by 11 % (2/18) of the borderline findings in specific IgM-negative patients obtained in both tests (two more with Labsystems, amounting to 22 %). The persistence of low/borderline avidity during a prolonged period of time in certain individuals has been noted in many studies (Jenum et al., 1997; Montoya et al., 2004; Remington et al., 2004; Petersen et al., 2005). In a recent study of delayed maturation of IgG avidity in pregnant women, Lefevre-Pettazzonni et al. (2006) attributed this phenomenon to factors including individual variations, state of pregnancy and possibly treatment, but also to the assay system. However, as maturation of IgG avidity is not necessarily linear over time, it has also been considered to be individual rather than assay-related (Petersen et al., 2005).

In conclusion, based on the results of this study, the EUROMMUN test may have its place, particularly in low-resource settings, but should be used prudently. On the other hand, the excellent concordance between the Labsystems and VIDAS tests shows that these are equally reliable for the measurement of *Toxoplasma* IgG avidity. Any limitations of the interpretation of the avidity results are intrinsic to the IgG avidity per se rather than to these two tests. One advantage of the VIDAS assay is that the results are produced by the system software, but the appropriate apparatus is required, whereas the Labsystems test requires no specific equipment but its performance and calculation of the results are cumbersome and time-consuming. Thus the choice of test should depend on the particular laboratory framework.

ACKNOWLEDGEMENTS

The expert technical assistance of Mrs Jordanka Djurović is gratefully acknowledged. The study was supported by grant no. 145002 from the Ministry of Science and Technological Development of Serbia. The authors do not have any commercial interest in the tests evaluated.

REFERENCES


