

Performance of five serological chlamydia antibody tests in subfertile women

J.A.Land^{1,4}, A.P.Gijsen¹, A.G.H.Kessels², M.E.P.Slobbe³ and C.A.Bruggeman³

¹Department of Obstetrics and Gynaecology, ²Department of Clinical Epidemiology and Medical Technology Assessment and

³Department of Medical Microbiology, Academisch Ziekenhuis Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands

⁴To whom correspondence should be addressed. E-mail: jlan@sgyn.azm.nl

BACKGROUND: Micro-immunofluorescence (MIF) is widely used for chlamydia antibody testing (CAT). Recently a species-specific MIF and three enzyme-linked immunosorbent assay (ELISA) tests have been introduced. We compared five commercially available CAT tests, using laparoscopy as a reference, and evaluated whether combinations of tests could improve the predictive value of CAT. **METHODS:** In a consecutive cohort of 315 subfertile women, results of the five CAT tests were correlated to findings at laparoscopy. Likelihood and odds ratios (OR) were calculated for single tests and for combinations of tests. **RESULTS:** Of the tests evaluated, MIF Labsystems had the best diagnostic performance (OR 15.7), while pELISA Medac (OR 8.2) was the best of the three ELISA tests. Stepwise logistic regression analysis showed that performance of MIF Labsystems could not be improved by adding a second test. Significant cross-reactivity with *C. pneumoniae* antibodies was found in all tests evaluated, except in pELISA Medac. **CONCLUSIONS:** In screening for tubal factor subfertility, MIF Labsystems was superior to the ELISA tests evaluated, and combining two CAT tests did not improve its predictive value. Economic analysis will show whether serial testing by pELISA Medac, and retesting positive samples by MIF Labsystems, is most cost-effective. In CAT, cross-reactivity with *C. pneumoniae* antibodies is still a major concern.

Key words: Chlamydia antibody test/diagnostic test/screening/serology/tubal infertility

Introduction

Since the association between *Chlamydia trachomatis* antibodies in serum and tubal pathology was noticed (Punnonen *et al.*, 1979), chlamydia antibody testing (CAT) has been used in the fertility work-up as a screening test for tubal factor subfertility. A widely used test for CAT is the micro-immunofluorescence test (MIF), which has been considered the gold standard in the serological diagnosis of chlamydia infection (Dowell *et al.*, 2001). Initially, antigens from elementary bodies of each of the serotypes of *C. trachomatis* were included in the MIF test, which provided serotype-specific antibody testing (Wang and Grayston, 1970). Later, for practical reasons, the number of antigens has been reduced by pooling antigens of epidemiologically related serotypes, or by using one broadly reacting serotype (usually L2) (Treharne *et al.*, 1977). However, in modified MIF tests, cross-reactivity between *C. trachomatis* and *C. pneumoniae* occurs (Mannion *et al.*, 1991; Gijsen *et al.*, 2001). To overcome this, species-specific MIF tests have been developed. In these tests, the cross-reactivity between the different chlamydia species has been reduced by subtracting lipopolysaccharide (LPS), a common component of the outer membrane of all Chlamydiaceae, from *C. pneumoniae* and *C. trachomatis* antigens. Other disadvantages of the MIF tests are that they

are labour intensive, their reading is observer dependent, and interlaboratory variation is significant (Peeling *et al.*, 2000). To overcome these drawbacks, enzyme-linked immunosorbent assay (ELISA) tests with high specificity have been developed, using LPS-stripped elementary bodies as antigens (Ossewaarde *et al.*, 1994). Recently assays have become commercially available in which specific synthetic peptides are used. These peptides are based on the major outer membrane protein (MOMP) of *C. trachomatis*, which contains species-specific and serotype-specific epitopes (Närvänen *et al.*, 1997). These ELISA tests are easy to perform and are well standardized (Tuuminen *et al.*, 2000). The value of CAT using ELISA and species-specific MIF in predicting tubal factor subfertility has not been evaluated yet.

The first aim of this study was to compare five commercially available chlamydia IgG antibody tests in their accuracy to predict tubal factor subfertility, using laparoscopy as a reference. The performances of a species-specific MIF test and three ELISA tests were compared with findings obtained with a MIF test that had been used in our hospital for 10 years. Cross-reactivity with *C. pneumoniae* antibodies was evaluated in the five CAT tests as a possible cause of false positive CAT results. The second aim was to evaluate whether combinations of two tests could improve the predictive value of CAT.

Materials and methods

The study was performed in female patients who sought treatment for subfertility in our clinic, a tertiary care centre. Since 1992, CAT using MIF has been a routine procedure in our fertility work-up. In all female patients, blood is drawn at their initial visit and a MIF test for *C. trachomatis* IgG antibodies (Biomerieux, The Netherlands) is performed, and all spare serum samples are cryopreserved at -20°C . CAT is used for selecting patients at high risk for tubal factor subfertility. Patients who have antibody titres <64 are considered low risk and tubal factor is evaluated initially by hysterosalpingography (HSG). In cases of an abnormal HSG or failure of pregnancy to occur in the 6 months following HSG, the fertility work-up is concluded by laparoscopy and tubal testing with methylene blue dye. In patients who have antibody titres ≥ 64 , no HSG is performed, but tubal factor is evaluated primarily by laparoscopy.

Between June 1992 and September 2001, 315 successive patients underwent laparoscopy as part of their fertility work-up. All took part in this study. Patients who had undergone previous pelvic surgery (except for an uneventful appendectomy or Caesarean section) were excluded.

All laparoscopy reports were scored by two gynaecologists, who were blinded for the CAT results. For the sake of the study, tubal pathology at laparoscopy was defined as extensive periannexal adhesions and/or distal occlusion of both tubes, since this has been shown to reflect chlamydia associated tubal disease most accurately (Land *et al.*, 1998).

After thawing the cryopreserved sera of the participating patients, the five different CAT tests were performed. A number of patients had participated in a previous study (Land *et al.*, 1998), but for the present study all serum samples were retested. Finally, data obtained at laparoscopy were correlated to the serological test results.

Serological methods

All assays and calculations were performed according to the manufacturer's instructions. For each test, all sera were tested simultaneously using the same batch of test kits, to avoid inter-test variability.

MIF Biomerieux

Chlamydia trachomatis-spot IF test (Biomerieux). In this indirect fluorescent IgG antibody test, *C. trachomatis* L2 is used as a group antigen. A positive reaction is characterized by specific fluorescence of the *C. trachomatis* elementary bodies. For quantitative determination serial dilutions are made. In this study all slides were evaluated independently by two readers. In case of disagreement, the judgement of a third reader was decisive. In this study a third reader was needed in four cases.

MIF Labsystems

Chlamydia pneumoniae IgG/IgM micro-IF test (Labsystems, Finland). In this species-specific MIF test *C. pneumoniae*, *C. trachomatis* and *C. psittaci* elementary bodies are used as antigens. In order to diminish cross-reactivity, the immunological activity of chlamydia LPS in *C. pneumoniae* and *C. trachomatis* antigens has been reduced. In this study the test was used to detect anti-*C. trachomatis* and anti-*C. pneumoniae* IgG antibodies. A positive reaction is characterized by specific fluorescence of the respective elementary bodies. For quantitative determination serial dilutions are made. In this study all slides were evaluated independently by two readers. In case of disagreement, the judgement of a third reader was decisive. In this study a third reader was needed in three cases.

ELISA Labsystems

Chlamydia trachomatis IgG EIA (Labsystems). In this indirect enzyme immunoassay synthetic peptides are used, which are derived from MOMP of *C. trachomatis* (Närvänen *et al.*, 1997). The optical density of the test plates is read in a spectrophotometer, and the signal to cut-off indices (SCI) are categorized as negative, equivocal, positive or highly positive.

pELISA Medac

Chlamydia trachomatis-IgG-pELISA (Medac, Germany). Medac has a LPS-based, genus-specific test (rELISA) and a species-specific ELISA (pELISA) on the market. The pELISA has been used in this study, and according to the manufacturer's information it is based on 'a synthetic peptide' from MOMP of *C. trachomatis*. The optical density of the test plates is read in a spectrophotometer, and the SCI are categorized as negative, equivocal, or positive.

ELISA Savyon

Sero-CT-IgG (Savyon, Israel). In this ELISA *C. trachomatis* species-specific peptides, derived from 'different serotypes' are used. The optical density of the test plates is read in a spectrophotometer, and the SCI are categorized as negative, equivocal, or positive.

Statistical methods

The outcomes of the five tests were compared with the findings at laparoscopy in order to determine their prognostic value of tubal factor subfertility. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio (LR), odds ratio (OR) and 95% confidence interval (CI) were calculated.

For comparison of LR of the results of the CAT tests, the χ^2 -test was used. To test the difference between OR of the different tests, the bootstrap technique was used (Efron and Tibshirani, 1993). Two-sided $P < 0.05$ was considered significant. A stepwise logistic regression analysis was used to select the best combination of tests. For evaluation of agreement between tests, κ was calculated, and $\kappa > 0.7$ was considered to indicate good agreement. For comparison of the distributions of *C. pneumoniae* titres in patients with a positive and negative chlamydia antibody test respectively, the Mann-Whitney U -test was used. $P < 0.05$ was considered significant.

Results

In 315 subfertile women, results of five different CAT tests and laparoscopy were available for analysis. In 51 patients (16%) tubal pathology was found at laparoscopy. In Table I results are given for each test, reflecting its accuracy in predicting tubal factor subfertility. The effect of using different cut-off levels for a positive test is shown. Within each test, increase of the cut-off titre or SCI increased specificity and PPV at the expense of sensitivity. When using 32 or 64 as the cut-off level for a positive test, the OR of MIF Labsystems differed significantly from all OR of the other four CAT tests. For further analyses, not all different cut-off levels and SCI were used. MIF Biomerieux and MIF Labsystems were considered positive if the IgG titre was ≥ 32 , ELISA Labsystems was considered positive if SCI was positive or highly positive, and pELISA Medac and ELISA Savyon if SCI was positive.

When MIF Biomerieux was used, 132 women had a positive CAT test, 52 women had a positive test by MIF Labsystems, 53 by ELISA Labsystems, 62 by pELISA Medac and 87 women had a positive test when ELISA Savyon was used (Table I).

Table I. Chlamydia antibody testing and tubal pathology at laparoscopy in 315 patients, using five different serological tests and different cut-off levels

| Test | Cut-off* | No. of patients with positive test | No. of patients with positive test and TP | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | LR+ | LR- | OR | 95% CI |
|------------------|---------------|------------------------------------|---|-----------------|-----------------|---------|---------|------|-----|-------------------|----------|
| MIF Biomerieux | 8 | 231 | 45 | 88 | 30 | 19 | 93 | 1.3 | 0.4 | 3.1 | 1.2–9.9 |
| | 16 | 149 | 39 | 76 | 58 | 26 | 93 | 1.8 | 0.4 | 4.6 | 2.1–10.3 |
| | 32 | 132 | 37 | 73 | 64 | 28 | 92 | 2.0 | 0.4 | 4.7 ^a | 2.3–10.2 |
| | 64 | 104 | 36 | 71 | 74 | 35 | 93 | 2.7 | 0.4 | 6.9 | 3.3–14.9 |
| | 128 | 60 | 29 | 57 | 88 | 48 | 91 | 4.8 | 0.5 | 9.9 | 4.7–21.1 |
| MIF Labsystems | 8 | 91 | 32 | 61 | 77 | 34 | 91 | 2.7 | 0.5 | 5.3 | 2.6–10.8 |
| | 16 | 75 | 32 | 61 | 83 | 41 | 92 | 3.6 | 0.5 | 7.8 | 3.7–16.2 |
| | 32 | 52 | 30 | 59 | 92 | 58 | 92 | 7.1 | 0.4 | 15.7 ^b | 7.1–35.1 |
| | 64 | 37 | 24 | 47 | 95 | 65 | 90 | 9.6 | 0.6 | 17.2 | 7.1–42.4 |
| | 128 | 19 | 13 | 25 | 98 | 68 | 87 | 11.2 | 0.8 | 14.7 | 4.6–52.4 |
| ELISA Labsystems | Equivocal | 84 | 23 | 45 | 77 | 27 | 88 | 2.0 | 0.7 | 2.7 | 1.3–5.4 |
| | Positive | 53 | 19 | 37 | 87 | 36 | 88 | 2.9 | 0.7 | 4.0 ^c | 1.9–8.4 |
| | High positive | 26 | 12 | 24 | 95 | 46 | 87 | 4.4 | 0.8 | 5.5 | 2.0–14.4 |
| pELISA Medac | Equivocal | 74 | 28 | 55 | 83 | 38 | 90 | 3.2 | 0.5 | 5.8 | 2.8–11.8 |
| | Positive | 62 | 28 | 55 | 87 | 45 | 91 | 4.3 | 0.5 | 8.2 ^d | 3.9–17.3 |
| ELISA Savyon | Equivocal | 99 | 26 | 51 | 72 | 26 | 88 | 1.8 | 0.7 | 2.7 | 1.4–5.4 |
| | Positive | 87 | 25 | 49 | 77 | 29 | 89 | 2.1 | 0.7 | 3.1 ^e | 1.6–6.3 |

Tubal pathology (TP) was found in 51 patients.

*IgG titre in micro-immunofluorescence (MIF) tests, and signal to cut-off index in enzyme-linked immunosorbent assay (ELISA) tests.

^{b–a}, ^{b–c}, ^{b–d}, ^{b–e}, ^{d–c}, ^{d–e} Differences statistically significant.

PPV = positive predictive value; NPV = negative predictive value; LR+ = likelihood ratio of positive test; LR- = likelihood ratio of negative test; OR = odds ratio; CI = confidence interval.

Table II. Number of positive chlamydia antibody tests in 315 patients, and the likelihood of tubal pathology at laparoscopy

| No. of positive tests | MIF Biomerieux* | MIF Labsystems* | ELISA Labsystems† | pELISA Medac‡ | ELISA Savyon‡ | LR ^d | 95% CI |
|-----------------------|-----------------|-----------------|-------------------|---------------|---------------|------------------|----------|
| 0 | 183 | 263 | 262 | 253 | 228 | 0.5 | 0.3–0.7 |
| 1 | 46 | – | 4 | 3 | 14 | 0.4 ^a | 0.2–0.9 |
| 2 | 22 | 5 | 4 | 5 | 14 | 1.0 | 0.4–2.6 |
| 3 | 17 | 6 | 10 | 9 | 12 | 2.0 ^b | 0.8–5.0 |
| 4 | 20 | 14 | 8 | 18 | 20 | 4.2 ^b | 1.9–9.4 |
| 5 | 27 | 27 | 27 | 27 | 27 | 6.5 ^b | 3.2–12.7 |

*IgG titre ≥ 32 .

†Signal to cut-off index positive or highly positive.

‡Signal to cut-off index positive.

^{a–b} Difference statistically significant.

LR = likelihood ratio; CI = confidence interval.

The OR of MIF Labsystems (15.7) was significantly better than the OR of all other tests evaluated. pELISA Medac had the highest OR (8.2) of the three ELISA tests.

In 158 patients all five CAT tests were negative, 67 patients had one positive test, 25 patients had two positive tests, 18 patients had three positive tests, 20 patients had four positive tests and in 27 patients all five tests were positive. Table II gives the number of positive CAT tests in 315 patients. The likelihood of tubal factor subfertility was correlated to the number of positive tests. A statistically significant difference was found between the LR of three, four and five simultaneously positive tests as compared with the LR of one positive test only.

The impact of combining two different tests on the predictive value for tubal factor subfertility was studied. Three different strategies were evaluated. First, a stepwise logistic regression analysis was performed to investigate whether combinations of tests would perform better than the best single CAT test (i.e. MIF Labsystems). From the logistic

regression analysis it could be concluded that the diagnostic performance of MIF Labsystems could not be improved significantly by adding another CAT test.

As a second strategy the CAT tests were divided into two groups: the ELISA tests, as they can be automated, and the more laborious MIF tests. Based on stepwise logistic regression analysis with the three ELISA tests in the first block and the two MIF tests in the second block, pELISA Medac was chosen as the first test, and MIF Labsystems as the second test (Table III). In retesting of initially positive sera, the serial set was considered positive if the second test was also positive, and in retesting of initially negative sera the serial set was considered negative if the second test was also negative. Using pELISA Medac on all serum samples, 62 patients (20%) had a positive test. After retesting these 62 sera by MIF Labsystems, 44 sera (14%) remained positive. An initial negative test was found in 253 patients (80%), and after retesting these 253 sera, 245 were negative (78%). The results of serial testing with the chosen set of CAT tests (i.e. pELISA

Table III. Chlamydia antibody testing and tubal pathology at laparoscopy in 315 patients: results for serial testing with two different sets of tests

| Sets of tests | No. of patients with positive test | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | LR+ | LR- | OR | 95% CI |
|--------------------------|------------------------------------|-----------------|-----------------|---------|---------|-----|-----|------|----------|
| Serial set* | | | | | | | | | |
| Retest positive results† | 44 | 51 | 93 | 59 | 91 | 7.5 | 0.5 | 14.2 | 6.2–32.6 |
| Retest negative results† | 70 | 61 | 86 | 45 | 92 | 4.2 | 0.5 | 9.2 | 4.4–19.5 |
| Low κ set‡ | | | | | | | | | |
| Retest positive results† | 47 | 35 | 89 | 38 | 88 | 3.2 | 0.7 | 4.4 | 2.0–9.6 |
| Retest negative results† | 138 | 75 | 62 | 28 | 93 | 2.0 | 0.4 | 4.8 | 2.3–10.6 |

*pELISA Medac (signal to cut-off index positive) and MIF Labsystems (IgG titre ≥ 32).

†Positive, retesting of all positive test results; negative, retesting of all negative test results.

‡ELISA Labsystems (signal to cut-off index positive or highly positive) and MIF Biomerieux (IgG titre ≥ 32).

PPV = positive predictive value; NPV = negative predictive value; LR+ = likelihood ratio of positive test; LR- = likelihood ratio of negative test; OR = odds ratio; CI = confidence interval.

Table IV. κ -Values for agreement in predicting tubal pathology at laparoscopy between five chlamydia antibody tests

| | MIF Biomerieux* | MIF Labsystems* | ELISA Labsystems† | pELISA Medac‡ | ELISA Savyon‡ |
|-------------------|-----------------|-----------------|-------------------|---------------|---------------|
| MIF Biomerieux* | | 0.43 | 0.35 | 0.42 | 0.45 |
| MIF Labsystems* | | | 0.49 | 0.72 | 0.48 |
| ELISA Labsystems† | | | | 0.50 | 0.55 |
| pELISA Medac‡ | | | | | 0.59 |
| ELISA Savyon‡ | | | | | |

*IgG titre ≥ 32 .

†Signal to cut-off index positive or highly positive.

‡Signal to cut-off index positive.

Medac and MIF Labsystems) are given in Table III. The OR of both serial sets (14.2 and 9.2 respectively) were lower than the OR of MIF Labsystems as a single test (15.7; Table I).

In the third strategy, it was hypothesized that the two CAT tests which differed most from each other may cover a wider spectrum of serotypes, and act complementarily in identifying women with tubal pathology. In order to select the two tests to be used in the model, the agreement between the five tests was established (Table IV). For evaluation of the third strategy the set with the lowest κ value, i.e. ELISA Labsystems and MIF Biomerieux ($\kappa = 0.43$), was performed on all serum samples. The result of the set was defined as positive if both tests were positive, whereas all other combinations of test results were considered negative. The result of the set was considered negative if both tests had a negative test result, and all other combinations of test results were considered positive. The results of serial testing with the low κ set are given in Table III. The OR of the low κ set (4.4 and 4.8 respectively) did not differ significantly from the OR of the single tests of which the sets were composed (4.0 for ELISA Labsystems and 4.7 for MIF Biomerieux respectively; Table I). This indicates that serial testing with the two CAT tests which differed most from each other could not improve the predictive value of CAT.

In order to study cross-reactivity with *C. pneumoniae* antibodies in the five CAT tests, in all 315 patients *C. pneumoniae* IgG antibody titres were determined by MIF Labsystems. In 41 patients the IgG titre was < 8 , in five patients it was 8, in 33 patients it was 16, in 61 patients it was 32, in 83 patients it was 64, in 56 patients it was 128, in 20 patients it was 256 and in 16 patients it was > 256 . Based on the results of the CAT test, patients were divided into two groups, i.e. those with

a positive test result (ELISA positive or highly positive, or MIF titre ≥ 32) and those with a negative test result. The distribution of the *C. pneumoniae* antibody titres in the CAT positive patients was compared with the distribution in the CAT negative patients. This comparison was made for all five CAT tests separately. Significant differences were found between the distributions of *C. pneumoniae* antibodies in the CAT positive and CAT negative patients when CAT was performed by MIF Biomerieux ($P < 0.00001$), MIF Labsystems ($P = 0.001$), ELISA Labsystems ($P = 0.02$) or ELISA Savyon ($P = 0.02$), whereas no significant difference was found by pELISA Medac ($P = 0.06$). This suggests significant cross-reactivity with *C. pneumoniae* antibodies in all CAT tests evaluated, except in pELISA Medac.

Discussion

Active genital tract infections with *C. trachomatis* can be diagnosed by direct detection of the micro-organism from the infected site. After the acute episode however, the organism may not be detectable any longer, and chlamydia antibodies in serum may be the only indication of previous chlamydia involvement. The aim of screening subfertile women by CAT is to identify patients with previous *C. trachomatis* infections, who are at increased risk for tubal pathology. It has become evident, however, that not all women develop *C. trachomatis* antibodies after chlamydia infection (Schachter *et al.*, 1979), and that not all women with antibodies have tubal pathology. Although the immunopathology underlying chlamydia infection is still poorly understood, antibody tests have been developed for clinical application. The information provided

by the manufacturers about the antigenic epitopes used in their tests is very limited, and is restricted to statements that the test contains 'a' or 'some' specific peptides. Furthermore, manufacturers modify the epitopes used in their tests, which may remain unnoticed by the customers who therefore cannot anticipate changes in test performance. The present study is a clinical comparison of five commercially available chlamydia antibody tests in their ability to predict tubal pathology in subfertile women. Four of these tests had not been evaluated as screening tests for tubal factor subfertility before. We are aware that the results of the study show the diagnostic performances of the tests at a given moment in time only.

In the present study, CAT results were compared with the findings at laparoscopy, and therefore only women who had undergone a laparoscopy were included. Patients with low chlamydia antibody titres are less likely to have tubal pathology and were less likely to be included in the study, since many will have conceived before laparoscopy can be done. This verification bias (Mol *et al.*, 1999) will influence predictive values of CAT, which are dependent on the prevalence of disease. Verification and selection bias is hard to prevent in clinical studies, however, unless one is prepared to perform the complete fertility investigation of a patient on a single day. Although the bias will result in overestimation of the LR, the overestimation will be similar for all tests investigated.

In comparing the five CAT tests, considerable variation was found in the number of patients with a positive test (Table I). An IgG titre ≥ 32 was found in 132 women using MIF Biomerieux and in 52 using MIF Labsystems, 53 had a positive test by ELISA Labsystems, 62 by pELISA Medac and 87 by ELISA Savyon. The tests with the highest number of positive test results had the lowest PPV, indicating the highest rates of false positive test results. False positive CAT results (i.e. positive CAT tests in patients without tubal pathology at laparoscopy) may be explained by cross-reactivity with *C. pneumoniae* antibodies, which can be found in ~70% of subfertile women (Gijzen *et al.*, 2001). In this study a significant correlation was found between *C. pneumoniae* antibody titres and CAT results in four tests, in particular in MIF Biomerieux ($P < 0.00001$). This confirms the findings of our previous study using a *C. pneumoniae* IgG ELISA, in which we found cross-reactivity to occur between *C. pneumoniae* and *C. trachomatis* antibodies in MIF Biomerieux. Since the only CAT test (pELISA Medac) in which no significant correlation with the distribution of *C. pneumoniae* titres was found did not have the highest PPV, cross-reactivity with *C. pneumoniae* antibodies does not seem to be the only explanation for the high rates of false positive test results obtained in the CAT tests. Since not all women with chlamydia antibodies have tubal pathology at laparoscopy, it has been suggested that genetic factors in the host may also play a role, by modulating immune defence mechanisms and the development of late sequelae (Kinnunen *et al.*, 2002).

Comparing MIF and ELISA, the results (Table I) suggest that ELISA tests tend to have lower sensitivity and NPV, i.e. more false negative test results. In ELISA tests chosen for this study, specific synthetic peptides are used which are considered

analogous to the serotype-specific antigenic determinants of MOMP of *C. trachomatis* (Närvänen *et al.*, 1997). These serotype-specific determinants differ between tests, and may explain the differences found in test performances between ELISA tests of different manufacturers. Furthermore, tests based on highly specific peptides may be so specific that they are not able to detect all relevant antigens (Bas *et al.*, 2001). Variants of serotypes have been identified in urogenital isolates (Morré *et al.*, 1998) and mutations have been shown to occur in positions within MOMP (Dean *et al.*, 2000). Consequently, highly specific tests may not be able to identify all serotypes involved in chlamydia infection, and cause false negative CAT results (i.e. negative CAT tests in patients with tubal pathology at laparoscopy).

There are a few reports in the literature on the diagnostic accuracies of different antibody tests for *C. trachomatis*, in which the tests evaluated in our study have been included. In these studies either tubal pathology, or the direct demonstration of the micro-organism in the genital tract, have been used as reference standards for the serological tests. Each reference standard has its limitations, however. Since tubal pathology can be caused by other micro-organisms in addition to *C. trachomatis*, it is obvious that tests based on chlamydia antibodies will be imperfect in predicting all tubal pathology. Studies in which the direct detection of the micro-organism is used as a reference have also limited diagnostic accuracy (Chernesky *et al.*, 1998; Bas *et al.*, 2001), since superficial infections may provide a poor stimulus for antibody formation. Paukku *et al.* (1998) did not find a significant difference between the presence of IgG antibodies in 78 patients with tubal factor subfertility, using a modified MIF test and ELISA Labsystems. In a serological follow-up study of 16 women with *C. trachomatis* positive cervical swabs, ELISA Labsystems, ELISA Savyon and a MIF test have been used (Clad *et al.*, 2000). ELISA Labsystems was found to be the most sensitive test, and it was concluded that ELISA Savyon did not cover all chlamydia serotypes. Morré *et al.* (2002) studied the IgG prevalences in 43 women with PCR positive cervical swabs and 106 PCR negative women. Results obtained by two in-house MIF tests were compared with results by ELISA Labsystems, pELISA Medac and ELISA Savyon. The authors concluded that the ELISA tests performed as well as the MIF tests.

In the present study the likelihood of tubal factor subfertility improved as the number of positive CAT tests in a patient increased (Table II). Compared with one positive CAT test, the LR improved significantly in patients in whom three, four or five positive tests were found. But from a clinical point of view, performing more than two CAT tests in patients is impractical and expensive. Therefore, we evaluated whether two serially performed tests could be of any benefit in predicting tubal factor subfertility. First, from a stepwise logistic regression analysis it was concluded that the diagnostic performance of the best single CAT test (i.e. MIF Labsystems) could not be improved by adding a second CAT test. Second, we evaluated the performance of a serial set of tests from a laboratory perspective, by constructing a model in which pELISA Medac was performed as the first test and MIF Labsystems as the second test. Although pELISA Medac was found to have a

significantly lower OR (8.2) compared with MIF Labsystems (15.7), ELISA tests have the advantage over MIF tests of being less laborious. From our results it can be concluded that if pELISA Medac is performed as the first test on all samples, and all samples with positive test results (i.e. 20% of all samples) are retested with MIF Labsystems, the predictive value of the set (OR 14.2) is comparable to the predictive value of MIF Labsystems as a single test (OR 15.7). Cost-effectiveness analysis has to demonstrate which strategy is to be preferred: MIF Labsystems as a single test on all samples, or pELISA Medac on all samples and retesting of the positive ones with MIF Labsystems. Third, we hypothesized that two tests with poor agreement (low κ) might have different antigenic properties, and might react with a greater number of serotypes and therefore complement each other. From the results presented in Table III it can be concluded that the low κ set of tests (ELISA Labsystems and MIF Biomerieux) did not improve the predictive value of CAT significantly, compared with the predictive values of the single tests of which the set was composed. Therefore the hypothesis of complementary testing with the set of tests with low κ had to be rejected.

The results of the present study show that there still is no excellent screening test for tubal pathology in subfertile women. In order to develop more accurate tests for the prediction of chlamydia-associated tubal pathology, future research should focus on the immunopathology of chlamydia infections. Evidence exists that patients with chronic, persisting chlamydia infections are particularly at risk for developing late sequelae. Although IgG antibodies are markers of a previous infection, they do not reflect an ongoing chronic inflammation properly. Candidates to be introduced into screening for tubal factor subfertility, in addition to specific *C. trachomatis* IgG antibodies, are anti-HSP60 and anti-LPS antibodies. HSP60 has been shown to play a prominent role in chronic inflammation and scarring (Claman *et al.*, 1997), and anti-LPS antibodies might be indicators of ongoing chlamydia infection (Tuuminen *et al.*, 2000).

In conclusion, although ELISA tests have been claimed to be highly sensitive and specific, in the present study they were not superior in predicting tubal factor subfertility. Of the five CAT tests evaluated, MIF Labsystems had the best diagnostic performance, and among the three ELISA tests, pELISA Medac performed best. MIF Biomerieux had the largest number of false positive test results, probably due to cross-reactivity with *C. pneumoniae* antibodies. Combining two different CAT tests did not improve the predictive value for tubal factor subfertility. Health care evaluation from an economic perspective has to prove whether serial testing with the automated pELISA Medac as a first test, and retesting of all positive serum samples with the more laborious MIF Labsystems, is to be preferred to testing of all samples with MIF Labsystems only.

Acknowledgements

The laboratory assistance of Gert Grauls is gratefully acknowledged. This study was supported by a grant from ZON MW (945-12-015), The Netherlands.

References

- Bas, S., Muzzin, P., Ninet, B., Bornant, J.E., Scieux, C. and Vischer, T.L. (2001) Chlamydia serology: comparative diagnostic value of immunoblotting, microimmunofluorescence test, immunoassays using different recombinant proteins as antigens. *J. Clin. Microbiol.*, **39**, 1368–1377.
- Chernesky, M., Luinster, K., Sellors, J., Schachter, J., Moncada, J., Caul, C., Paul, I., Mikaelian, L., Toye, B., Paavonen, J. and Mahony, J. (1998) Can serology diagnose upper genital tract *Chlamydia trachomatis* infections? Studies on women with pelvic pain, with or without chlamydial plasmid DNA in endometrial biopsy tissue. *Sex. Transm. Dis.*, **25**, 14–19.
- Clad, A., Freidank, H.M., Kunze, M., Schnoeckel, U., Hofmeier, S., Flecken, U. and Petersen, E.E. (2000) Detection of seroconversion and persistence of *Chlamydia trachomatis* antibodies in five different serological tests. *Eur. J. Clin. Microbiol. Infect. Dis.*, **19**, 932–937.
- Claman, P., Honey, L., Peeling, R.W., Jessamine, P. and Toye, B. (1997) The presence of serum antibody to the chlamydial heat shock protein (CHSP60) as a diagnostic test for tubal factor infertility. *Fertil. Steril.*, **67**, 501–504.
- Dean, D., Suchland, R.J. and Stamm, W.E. (2000) Evidence of long-term cervical persistence of *Chlamydia trachomatis* by *omp1* genotyping. *J. Infect. Dis.*, **182**, 909–916.
- Dowell, S.F., Peeling, R.W., Boman, J., Carlone, G.M., Fields, B.S., Guarner, J., Hammerschlag, M.R., Jackson, L.A., Kuo, C.C., Maass, M. *et al.* (2001) Standardizing Chlamydia pneumoniae assays: recommendations from the Centers for Disease Control and Prevention (USA) and the Laboratory Centre for Disease Control (Canada). *Clin. Infect. Dis.*, **33**, 492–503.
- Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*. Chapman & Hall, New York.
- Gijzen, A.P., Land, J.A., Goossens, V.J., Leffers, P., Bruggeman, C.A. and Evers, J.L.H. (2001) *Chlamydia pneumoniae* and screening for tubal factor subfertility. *Hum. Reprod.*, **16**, 487–491.
- Kinnunen, A.H., Surcel, H.M., Lehtinen, M., Karhukorpi, J., Tiitinen, A., Halttunen, M., Bloigu, A., Morisson, R.P., Karttunen, R. and Paavonen, J. (2002) HLA DQ alleles and interleukin-10 polymorphism associated with *Chlamydia trachomatis*-related tubal factor infertility: a case-control study. *Hum. Reprod.*, **17**, 2073–2078.
- Land, J.A., Goossens, V.J. and Evers, J.L.H. (1998) How to use *Chlamydia* antibody testing in subfertility patients. *Hum. Reprod.*, **13**, 1094–1098.
- Mannion, P.T., Mallinson, H. and Treharne, J.D. (1991) Serological diagnosis with the *Chlamydia* spot-IF test. *J. Med. Microbiol.*, **35**, 244–248.
- Mol, B.W.J., van der Veen, F. and Bosschuyt, P.M.M. (1999) *Chlamydia* antibody titres. *Hum. Reprod.*, **14**, 268–270.
- Morré, S.A., Ossewaarde, J.M., Lan, J., van Doornum, G.J.J., Walboomers, J.M.M., MacLaren, D.M., Meijer, C.J.L.M. and van den Brule, A.J.C. (1998) Serotyping and genotyping of genital *Chlamydia trachomatis* isolates reveal variants of serovar Ba, G and J as confirmed by *omp1* nucleotide sequence analysis. *J. Clin. Microbiol.*, **36**, 345–351.
- Morré, S.A., Munk, C., Persson, K., Krüger-Kjaer, S., van Dijk, R., Meijer, C.J.L.M. and van den Brule, A.J.C. (2002) Comparison of three commercially available peptide-based immunoglobulin G (IgG) and IgA assays to microimmunofluorescence assay for detection of *Chlamydia trachomatis* antibodies. *J. Clin. Microbiol.*, **40**, 584–587.
- Närvänen, A., Puolakkainen, M., Hao, W., Kino, K. and Suni, J. (1997) Detection of antibodies to *Chlamydia trachomatis* with peptide-based species-specific enzyme immunoassay. *Infect. Dis. Obstet. Gynecol.*, **5**, 349–354.
- Ossewaarde, J.M., de Vries, A., van den Hoek, J.A.R. and van Loon, A.M. (1994) Enzyme immunoassay with enhanced specificity for detection of antibodies to *Chlamydia trachomatis*. *J. Clin. Microbiol.*, **32**, 1419–1426.
- Paukku, M., Närvänen, A., Puolakkainen, M., Dreesbach, K., Tiitinen, A., Hao, W., Anttila, T.I. and Paavonen, J. (1998) Detection of *Chlamydia trachomatis* antibodies in 2 novel tests: rELISA and peptide EIA. *Int. J. STD AIDS*, **9**, 604–607.
- Peeling, R.W., Wang, S.P., Grayston, J.T., Blasi, F., Boman, J., Clad, A., Freidank, H., Gaydos, C.A., Gnarp, J., Hagiwara, T. *et al.* (2000) *Chlamydia pneumoniae* serology: interlaboratory variation in microimmunofluorescence assay results. *J. Infect. Dis.*, **181** (Suppl. 3), s426–429.
- Punnonen, R., Terho, P., Nikkanen, V. and Meurman, O. (1979) Chlamydial serology in infertile women by immunofluorescence. *Fertil. Steril.*, **31**, 656–659.
- Schachter, J., Cles, L., Ray, R. and Hines, P.A. (1979) Failure of serology in

- diagnosing chlamydial infections of the female genital tract. *J. Clin. Microbiol.*, **10**, 647–649.
- Trehanne, J.D., Darougar, S. and Jones, B.R. (1977) Modification of the microimmunofluorescence test to provide a routine serodiagnostic test for chlamydial infection. *J. Clin. Pathol.*, **30**, 510–517.
- Tuuminen, T., Palomäki, P. and Paavonen, J. (2000) The use of serologic tests for the diagnosis of chlamydial infections. *J. Microbiol. Meth.*, **42**, 265–279.
- Wang, S.P. and Grayston, J.T. (1970) Immunologic relationship between genital TRIC, lymphogranuloma venereum, and related organisms in a new microtiter indirect immunofluorescence test. *Am. J. Ophthalmol.*, **70**, 367–374.

Submitted on March 14, 2003; resubmitted on May 14, 2003; accepted on August 19, 2003