Sensitivity of three serum antibody tests in a large outbreak of Legionnaires’ disease in the Netherlands

Ed P. F. Yzerman, Jeroen W. den Boer, Kamilla D. Lettinga, Arnoud J. Schel, Joop Schellekens and Marcel Peeters

Correspondence
Ed P. F. Yzerman
eyzerman@streeklabhaarlem.nl

1Regional Laboratory of Public Health Haarlem, Boerhaavelaan 26, 2035 RC Haarlem, The Netherlands
2Academic Medical Center, Amsterdam, The Netherlands
3Municipal Health Service Kennemerland, Haarlem, The Netherlands
4National Institute for Public Health and the Environment, Bilthoven, The Netherlands
5Regional Laboratory of Public Health Tilburg, Tilburg, The Netherlands

Received 11 October 2005
Accepted 26 January 2006

In 1999, an outbreak involving 188 patients with Legionnaires’ disease (LD) occurred at a flower show in the Netherlands. This large outbreak provided the opportunity to evaluate serum antibody tests to assay anti-Legionella pneumophila, since limited data are available on the sensitivity of these tests. The sensitivities of an indirect serotype 1-6 immunofluorescence antibody test (IFAT), a rapid micro-agglutination test (RMAT) IgM serotype 1 antibody assay, and an ELISA to detect IgM and IgG serotype 1-7 antibodies, were evaluated using serum samples from LD patients related to the 1999 outbreak. Sensitivity was calculated using positive culture and/or a positive urinary antigen test as the gold standard in outbreak-related patients with radiographically confirmed pneumonia who fulfilled the epidemiological criteria. The IFAT, RMAT and ELISA showed sensitivities of 61, 44 and 64 %, respectively. The sensitivity of the three tests combined was 67 %. In epidemic situations, however, high standing titres may be included in the laboratory evidence of LD cases. In the study population, high standing titres were found in 16 % of cases. If the presence of high standing antibody titres was added to the criteria of a positive test, the sensitivities of IFAT, RMAT and ELISA were 86, 48 and 75 %, respectively. The sensitivity was 91 % for all tests combined. The higher sensitivity for the combined use of tests is offset by a reduction in specificity to 97 ∼ 6 %. The results of this study indicate that using a combination of serologic tests in pneumonia patients suspected to have LD does not substantially improve sensitivity. The results suggest that in the microbiological diagnosis of LD, both IFAT and ELISA are reasonably sensitive assays. In an epidemic situation, both tests are highly sensitive, the IFAT more so than the ELISA.

INTRODUCTION

Legionnaires’ disease (LD) is an acute pneumonia caused by Legionella spp., which are responsible for 2–5 % of community-acquired pneumonias (CAPs) (Marston et al., 1997; Bohte et al., 1995). Patients are infected by inhalation of aerosols containing Legionella spp. or by aspiration of water contaminated with Legionella spp. Over 90 % of LD cases are caused by Legionella pneumophila, of which 70–80 % belong to serogroup type I (Yu et al., 2002; Marston et al., 1997). Although over 90 % of LD cases are sporadic, much can be learned from outbreak-related LD cases, including their microbiological diagnosis. Diagnosis of LD is mainly based on culture, antigen detection in urine, and antibody detection in serum. Culturing of Legionella spp. from respiratory secretions has a sensitivity ranging from 25 to 75 % (Hoge & Breiman, 1991). Urinary antigen tests, both enzyme immunoassay (EIA) and immunochromatographic assay (ICT) combine a 99 % or higher specificity with a sensitivity ranging from 50 to 90 % (Birtles et al., 1990; Dominguez et al., 1998, 1999; Ruf et al., 1990). Serological tests to identify antibodies to L. pneumophila were developed from 1977 onwards, starting with the indirect immunofluorescence antibody test (IFAT) (McDade et al., 1977). Several less laborious serological tests

Abbreviations: CAP, community-acquired pneumonia; EWGLI, European Working Group on Legionella Infections; IFAT, immunofluorescence antibody test; LD, Legionnaires’ disease; RMAT, rapid micro-agglutination test.
followed, among them numerous (mostly experimental) ELISAs and a micro-agglutination test (MAT). The first publication on an ELISA was in 1978 (Farshy et al., 1978). Although commercial ELISA kits have become widely used, few studies of their sensitivity and specificity have been published. The same holds for the MAT, which was also developed in 1978; this was followed by a rapid version, the rapid micro-agglutination test (RMAT), with identical test characteristics, in 1982 (Harrison & Taylor, 1982). The sensitivity of all three assays has been tested using sporadic cases of LD and not in outbreak situations. A recent outbreak in the Netherlands (den Boer et al., 2002) provided a unique opportunity to evaluate the sensitivity of three serum antibody tests. This outbreak was caused by a \textit{L. pneumophila} serogroup 1 strain. The nationwide alert for LD cases ensured optimal case finding, thereby decreasing patient selection bias. Furthermore, the outbreak situation allowed us to take the clinical spectrum into account, since all available clinical data for the LD patients related to this outbreak were recorded centrally. In addition, the conditions for a gold standard were favourable: two national reference laboratories performed all microbiological tests and a large collection of samples from a national serum bank was available to investigate the distribution of antibodies to \textit{L. pneumophila} serogroup 1 among the Dutch population.

To assess the sensitivity of the detection of serum \textit{Legionella} antibodies, we performed three tests on serum specimens from outbreak-related LD cases: a serotype 1-6 antibody IFAT, an IgM serotype 1 antibody RMAT, and an ELISA to detect IgM and IgG serotype 1-7 antibodies (Boshuizen et al., 2003). The aim of our study was to compare the sensitivities of the three tests using confirmative and epidemic LD criteria, to see if combined use of the three tests would be justified, and to compare our results to published sensitivities of IFAT, RMAT and ELISA in both epidemic and non-epidemic settings.

**METHODS**

**The outbreak.** In February 1999, an outbreak involving 133 confirmed and 55 probable cases of LD, according to the European Working Group on Legionella Infections (EWGLI) criteria (Anonymous, 1998), occurred in Bovenkarspel in the Netherlands. Of these patients, 89\% were hospitalized. The outbreak investigation indicated that a whirlpool displayed at the consumer product division of an annual flower show was the most likely source of infection. Two \textit{L. pneumophila} serogroup 1 strains were isolated from sputum of 29 patients. One of the strains was of the Allentown/France serogroup 1 subtype (Helbig et al., 2002) and was genotyped 010-London according to the EWGLI criteria (Vry et al., 2000). This strain was identified in 28 of the 29 culture-positive patients. Genotyping of the environmental \textit{L. pneumophila} strains revealed that one \textit{L. pneumophila} serogroup 1 strain found in the implicated whirlpool was of the 010-London genotype as well (Den Boer et al., 2002).

**Patients.** In this study, 133 patients with a confirmed \textit{Legionella} pneumonia according to the EWGLI criteria were included after having obtained written consent from them or from their relatives. Of these patients, 104 fulfilled the criteria of the gold standard used in this study. The gold standard (identical to the EWGLI criteria, but excluding serological evidence of infection) was defined as a patient who had symptoms compatible with pneumonia, with radiological signs of infiltration, who had visited the site of the outbreak during the incubation period, and who had laboratory evidence of infection with \textit{L. pneumophila}. Laboratory evidence included at least one of the following criteria:

1. Isolation of \textit{L. pneumophila} from a respiratory sample cultured on buffered charcoal yeast extract supplemented with 2-ketoglutarate (BCYA–2), followed by genotyping and subsequent comparison to the environmental strain of the Bovenkarspel outbreak.

2. Presence of \textit{L. pneumophila} antigens in urine specimens, as shown in an antigen test. The three tests used in our study were two enzyme immunoassays, the Binax (Binax, Portland, ME, USA) and Biotest (Biotest AG, Dreieich, Germany) \textit{Legionella} urinary antigen tests, and the BinaxNOW test, a qualitative immunochromatographic assay. All tests were used according to the manufacturers’ instructions. Urine was concentrated and reread after 1 h, as recommended (Dominguez et al., 1996; Helbig et al., 2001).

3. All isolates from sputum samples were sent to the National Institute for Public Health and the Environment (RIVM) for genotyping of \textit{L. pneumophila}. The Regional Laboratories of Public Health of Tilburg and Haarlem were the reference laboratories for the detection of \textit{L. pneumophila} antigen in urine.

**Classification of severity of disease.** To evaluate exclusion bias by differences in severity of disease, the patients were divided into three clinical categories for CAP. Clinical data were collected from the hospital chart using a standardized case record form. Severity of pneumonia was scored on a five-point scale (0–4), according to the minor criteria for severity of CAP (Niederman et al., 1993), using the following clinical criteria at hospital admission: (1) respiratory frequency above 30 breaths per minute, (2) PaO₂ below 60 mm Hg or O₂ saturation below 92\%, (3) bilateral or multilobar infiltration on chest X-ray, 4) systolic blood pressure below 90 mm Hg or diastolic below 60 mm Hg.

Patients with a radiographically proven unilateral unilobar pneumonia, but without signs or symptoms according to the above-mentioned criteria, were classified as CAP category 1 (mild pneumonia). CAP category 2 consisted of patients who fulfilled only one of the criteria. According to our definition, these patients were suffering from a moderately severe pneumonia. CAP category 3 consisted of patients who presented themselves with two or more criteria and were considered to suffer from a severe pneumonia.

**Serum samples and serum antibody tests.** All 62 laboratories of clinical microbiology involved in diagnosis and treatment of the patients related to this outbreak sent all available serum samples to the Regional Laboratory of Public Health in Haarlem. After collection, serum samples were stored in portions at −70°C.

The criteria for a paired serum sample confirmative for LD were defined as:

1. A fourfold rise in IgG, IgM and/or IgA antibodies to \textit{L. pneumophila} in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera, with final titre $\geq 1:128$, using a commercial serotype 1-6 antibody IFAT (Polyvalent \textit{Legionella} IFA Test Serogroups 1-6, Meridian Bioscience Europe Srl.) (McDade et al., 1977).

2. A fourfold rise in IgM antibodies to \textit{L. pneumophila} in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera, with final titre $\geq 1:32$ in accordance with the 99\% cut-off value found in a serosurvey among healthy volunteers representative of the Dutch population (Boshuizen et al., 2003), using an IgM, serogroup 1 Philadelphia-I subtype, antibody RMAT (Farshy et al., 1978).
(3) Seroconversion to positive IgM or IgG antibodies to *L. pneumophila* in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera, in accordance with the age-specific 99% cut-off values found in a serosurvey among healthy volunteers representative of the Dutch population (Boshuizen *et al.*, 2003), using a commercial ELISA to detect IgM and IgG serotype 1-7 antibodies (Serion classic ELISA, Institut Virion/Serion GmbH, Würzburg, Germany) (Etienne *et al.*, 1999). The antigens in this test are prepared from cultures of different *L. pneumophila* serogroups and subtypes: serogroup 1 Knoxville and Philadelphia subtypes, serogroup 2 Togus-1 subtype, serogroup 3 Bloomington-2 subtype, serogroup 4 Los Angeles-1 subtype, serogroup 5 Dallas-1 subtype, serogroup 6 Chicago-2 subtype, and serogroup 7 Chicago-8 subtype.

The criteria for a paired serum sample suggestive for LD used in an epidemic situation were defined as:

1. High standing titres of IgG, IgM and/or IgA antibodies to *L. pneumophila* in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera, with final titre \( \geq 1 : 128 \) using a serotype 1-6 antibody IFAT, as described above.

2. High standing titres of IgM antibodies to *L. pneumophila* in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera, with final titre \( \geq 1 : 32 \) using an IgM antibody RMAT, as described above.

3. High standing titres of IgM or IgG antibodies to *L. pneumophila* in paired acute-phase (0–15 days after disease onset) and convalescent-phase sera using an ELISA to detect IgM and IgG serotype 1-7 antibodies, as described above.

**Sensitivity.** Sensitivity was defined as the number of patients with a positive test result divided by the total number of patients tested. Sensitivity for confirmed LD was calculated using the confirmative criteria for LD. Sensitivity for an epidemic situation was calculated using the confirmative and the suggestive criteria for LD.

**Statistics.** Statistical analysis was performed with the statistical program SPSS version 12.0 (Statistical Product and Service Solutions). Univariate analysis was used to identify factors potentially associated with test positivity: age, gender, clinical severity, number of days passed between onset of disease and collection of the first serum sample, and number of serum samples taken.

### RESULTS AND DISCUSSION

Two or more properly timed serum samples with complete sampling dates were available for evaluation from 45 of the 104 hospitalized patients who fulfilled the criteria of the gold standard used in this study (see Fig. 1 for patient selection). Table 1 shows that excluded patients died significantly more often from LD, although their CAP score

![Fig. 1. Selection of LD patients in this study. UAG, urinary antigen.](http://jmm.sgmjournals.org)

<table>
<thead>
<tr>
<th>Table 1. Differences between 45 included and 59 excluded LD patients for factors possibly associated with exclusion bias</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor</strong></td>
</tr>
<tr>
<td>Proportion of male gender</td>
</tr>
<tr>
<td>Mean age in years (range)</td>
</tr>
<tr>
<td>CAP category 1</td>
</tr>
<tr>
<td>CAP category 2</td>
</tr>
<tr>
<td>CAP category 3</td>
</tr>
<tr>
<td>Admission to intensive care unit</td>
</tr>
<tr>
<td>Fatal outcome</td>
</tr>
</tbody>
</table>

*Difference is statistically significant.
distribution was very similar to that of the included patients. Of the 45 included patients, 26 were men, with a mean age of 64 years (range 46–78 years) and 19 were women with a mean age of 60 years (range 21–77 years). The microbiological diagnosis was established by culture alone in one case, by culture and by urinary antigen test in nine cases and by urinary antigen test alone in 35 cases. The first serum to be tested for antibodies against \textit{L. pneumophila} was collected 0–15 days (mean 8 days) after the first day of signs and symptoms of pneumonia. From 22 patients two sera were available, from 13 patients three sera were available and from 10 patients four to eight sera were available for evaluation. Table 2 shows the timing of the convalescent sera. For one of the patients, the volume of serum was too small to perform all three tests, so only the RMAT test was performed. Patients with positive test results on average had their first serum taken 2 days earlier in the acute phase than patients with high standing titres: 7·1 versus 9·4 days after disease onset. This finding may suggest that in LD cases with high standing antibody titres, seroconversion had already occurred. In total, 7/45 (16 %) of patients showed high standing titres in all assays. Table 3 shows the sensitivities of the three tests and the sensitivity for the combination of all tests for both the non-epidemic and epidemic criteria.

Using the epidemic criteria, the sensitivities were 61 and 64 %, respectively.

It seems that the superior sensitivity of the IFAT and ELISA compared to the RMAT was mostly due to the identification of antibodies other than IgM. In this outbreak, the combination of tests was of limited value.

It has been suggested that in LD, any diagnostic test with a specificity below 99 % is unacceptable (Hoge & Breiman, 1991). Most IFAT sensitivity studies use this criterion and show similar results in non-outbreak patient series. However, the sensitivity for the IFAT in the Philadelphia outbreak was 42 % for a fourfold rise in titre, given a specificity of 99·1 %. With high standing titres added to the criteria of an LD case, sensitivity was 68 % (McDade et al., 1977). The IFAT used in our study seems more sensitive, probably due to the use of a polyvalent conjugate directed against IgG, IgM and IgA antibodies.

The sensitivity of the RMAT was 63 % in a study evaluating sera from 119 patients with LD using a positive culture for \textit{L. pneumophila} or its presence shown by direct fluorescence antibody as the gold standard (Harrison et al., 1987). The substantially higher sensitivity of the RMAT assay found in that study (63 versus 44 % in our study) is difficult to explain. A different strain of \textit{L. pneumophila} serogroup 1 was used, but the influence of different strains on RMAT

### Table 2. Timing of convalescent serum sample(s) and presence of antibodies to \textit{L. pneumophila}, as demonstrated by any of three assays, for 45 LD patients for whom two or more properly timed sera were available

<table>
<thead>
<tr>
<th>Time</th>
<th>No. of patients with available sera</th>
<th>Standing titre*</th>
<th>Fourfold rise in titre or seroconversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2–6 weeks</td>
<td>37</td>
<td>6/37</td>
<td>22/37</td>
</tr>
<tr>
<td>6–12 weeks</td>
<td>12</td>
<td>0/12</td>
<td>4/12</td>
</tr>
<tr>
<td>More than 12 weeks</td>
<td>20</td>
<td>1/20</td>
<td>4/20</td>
</tr>
</tbody>
</table>

*Acute-phase serum was positive.

### Table 3. Sensitivity of three serum \textit{L. pneumophila} antibody tests using sera from 45 LD patients

<table>
<thead>
<tr>
<th>Test criterion</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourfold rise in IFAT</td>
<td>61 % (27/44)</td>
</tr>
<tr>
<td>Fourfold rise in RMAT</td>
<td>44 % (20/45)</td>
</tr>
<tr>
<td>Seroconversion in IgM or IgG ELISA</td>
<td>64 % (28/44)</td>
</tr>
<tr>
<td>Combination of the three tests above</td>
<td>67 % (30/45)</td>
</tr>
<tr>
<td>Fourfold rise or high standing titres in IFAT</td>
<td>86 % (38/44)</td>
</tr>
<tr>
<td>Fourfold rise or high standing titres in RMAT</td>
<td>58 % (26/45)</td>
</tr>
<tr>
<td>Seroconversion or high standing titres in IgM or IgG ELISA</td>
<td>75 % (33/44)</td>
</tr>
<tr>
<td>Combination of the three tests above</td>
<td>91 % (41/45)</td>
</tr>
</tbody>
</table>
sensitivity is unknown. There was a difference in the timing of the first serum sample (0–28 days after disease onset in the study of Harrison et al., 1987), but late sampling tends to underestimate the sensitivity. There were differences in study population: higher clinical severity, higher mortality, and over 50% immunocompromised patients in the study of Harrison et al. (1987).

To our knowledge, only two reports are available describing data from which an ELISA sensitivity based on a 99% or higher specificity could be inferred (Bangsberg et al., 1994; Wreghitt et al., 1982). Bangsberg et al. (1994) studied convalescent sera from 20 culture-proven LD patients and described an ELISA based on a lipopolysaccharide antigen with sensitivities of 70% for IgM and 75% for IgG, respectively. Although in the Bangsberg et al. (1994) study the sensitivity of an IgM and IgG ELISA combination is not described, the reported individual sensitivities are 10% higher than those found in our study. This may be due to a difference in the antigen used or a difference in study population. Wreghitt et al. (1982) studied sera from 11 patients and found sensitivities of 58 and 41% for IgM and IgG ELISA, respectively. In the Wreghitt et al. (1982) study, another serum antibody test was included in the gold standard, and no details were given concerning the timing of serum samples under study, complicating comparison to our data.

In conclusion, to diagnose L. pneumophila pneumonia, serology is still a valuable tool. This study is the first to compare the sensitivities of an IFAT, an RMAT and an ELISA in an outbreak situation. The IFAT and ELISA seem to yield similar results using confirmative LD criteria. In an epidemic situation, the IFAT seems superior. Combination of assays in an outbreak situation improves sensitivity only marginally, from 86 to 91%. The data from this study suggest that early sampling may lead to a higher sensitivity for all tests.

ACKNOWLEDGEMENTS

We thank all hospital clinicians and microbiologists for requesting patients’ permission and allowing us to collect clinical data and serum samples. Special thanks to Yvonne Boelens and Jacob P. Bruin, who performed the serum antibody tests.

REFERENCES


bacterium and demonstration of its role in other respiratory disease. 

**Niederman, M. S., Bass, J. B., Jr, Campbell, G. D. & 7 other authors.**


**Yu, V. L., Plouffe, J. F., Pastoris, M. C. & 8 other authors.**