Multiplex assay (Mikrogen re comBead) for detection of serum IgG and IgM antibodies to 13 recombinant antigens of Borrelia burgdorferi sensu lato in patients with neuroborreliosis: the more the better?

Ram B. Dessau,1 Jens K. Møller,2 Birte Kolmos2 and Anna J. Henningsson3

1Department of Clinical Microbiology, Slagelse Hospital, Slagelse, Denmark
2Department of Clinical Microbiology, Vejle Hospital, Vejle, Denmark
3Department of Clinical Microbiology, County Hospital Ryhov, Jönköping, Sweden

A multiplex-bead-based assay for the detection of serum antibodies to Borrelia burgdorferi sensu lato was evaluated. The assay contained 13 different antigens in both the IgG and the IgM assay; thus, a total of 26 measurement results were available from each sample. A total of 49 Danish patients with Lyme neuroborreliosis (LNB), 218 Danish blood donor controls, a set of 61 Swedish patients with LNB and 139 Swedish non-LNB patients investigated for suspected LNB were used. There are four parts developed in this study: a characterization of the sero-epidemiological antibody-response pattern, the construction of a diagnostic score, evaluation of the scoring method using an independent dataset and an assessment of the analytical quality of the multiplex assay. The VlsE IgG had the highest diagnostic value with an AUC (area under the curve) of 96% on the receiver operating characteristic curve. The OspC IgM had AUCs just above 80%. All the other antigens had both low quantitative reactivity and lower contrast in the patients with LNB compared to controls. The diagnostic value of the assay may be improved by using a logistic model giving a sensitivity of 90 and 79% for the specificities at 92 and 98%, respectively. Overall, the patients with LNB had serum reactivity in IgG VlsE, but modest antibody reactivity in the remaining 12 IgG and 13 IgM antibody measurements. Using a logistic regression model with five IgG and two IgM antigens, the sensitivity and specificity of the assay was improved; but the IgG VlsE component alone contributed most of the diagnostic contrast.

INTRODUCTION

Lyme borreliosis (LB) is a tick-borne bacterial infection with manifestations from different organ systems (Stanek et al., 2011, 2012). The diagnosis of LB is based on clinical features, supported by detection of specific IgG and IgM antibodies. Antibody detection is a very sensitive method, but as there may be a background level of natural immunity due to previous subclinical infection the positive predictive value may be low in clinical cases where the clinician may rather want to rule out LB. Serological testing for LB is being used extensively in many countries (Dessau et al., 2010a; Müller et al., 2012). In order to improve the specificity, a two-tier system has been widely used, where a positive result from the screening ELISA was further tested by an immunoblot assessing reactivity to several antigens. However, immunoblots require subjective interpretation, they are labour intensive and reproducibility has not been found to be optimal (Hunfeld et al., 2002; Müller et al., 2012). Therefore, an automated multiplex assay with objective quantitative cut-off calibration and reading of results should be a promising new development. The potential advantages could be rapid test results and improved analytic accuracy. The underlying theory is that a combination of biomarkers may provide a better diagnostic tool than any single test on its own (Dessau et al., 2010b; Pepe, 2003).

The purpose of the study was to investigate the seroreactivity to multiple antigens of Borrelia burgdorferi sensu lato in patients with Lyme neuroborreliosis (LNB) using a
commercial multiplex assay and to evaluate regression methods for scoring of the complex results. There are four topics developed in this study: (i) the sero-epidemiological antibody-response pattern, (ii) the construction of a diagnostic score, (iii) evaluation of the scoring method using an independent dataset, and (iv) the reproducibility and analytical quality of the multiplex assay.

METHODS

**Serum samples.** Serum samples from 49 patients with LNB and 218 blood donors (controls) were tested. This is the same patient material as described in a previous publication (Dessau, 2013). Samples were kept frozen at −20 °C and the analyses were performed in January 2012. The reference standard for patients with LNB was laboratory defined as requiring pleocytosis in the spinal fluid and a *Borrelia*-specific positive antibody index (IDEIA Lyme neuroborreliosis IgG/IgM assay; Oxoid). The LNB specimens were collected from consecutive routine sampling from 2006 to 2010, and serum samples collected from Danish blood donors in January 2011 were used as controls.

For validation of the logistic scoring model, a set of Swedish results were used with 61 patients characterized as definite or possible LNB and 139 non-LB patients (Henningsson et al., 2014). In this study, the non-LB patients were a cohort investigated for suspected LNB. There are five more LNB patients and 20 more non-LB controls in the dataset than originally described (Henningsson et al., 2014). These extra samples were run in the Luminex serum assay, but could not be included in the cited study due to the low volume of spinal fluid. These cases have, however, been described in previous studies (Henningsson et al., 2011; Tjernberg et al., 2011). The Swedish and Danish case definitions were comparable as 57 cases had definite LNB defined by both pleocytosis and a positive spinal fluid/serum antibody index using the same assay based on flagella antigen (IDEIA). All 61 cases, including the four possible cases of LNB, were grouped together in this study.

**Assays.** The laboratory assays used were recomBead Borrelia IgG and IgM (Mikrogen, www.mikrogen.de), which are implemented on a platform based on coded beads allowing for simultaneous measurement of multiple analytes (Luminex multiplexing instruments; EMD Millipore). Recombinant test antigens were derived from borrelial proteins designated p100, VlsE, p38, p39, OspA, OspC and p18. In each assay there is a total of 13 recombinant antigens selected from strains of *Borrelia afzelii*, *B. garinii*, *B. spielmani*, *B. bavariensis* and *B. burgdorferi sensu stricto*. The IgG and IgM assays contain identical antigens; thus, a total of 26 antibody measurements were available from each tested sample. The selection of antigens conforms to the German standard recommendation (Wilske et al., 2000) and the assays meet various technical standards (EC directive 98/97/EC, DIN 58969-44). Computer software for interpretation was provided by the manufacturer. A positive result in any two antigens would be interpreted as positive for IgG and IgM, except that OspC IgM alone was also scored as a positive result. The rationale and the evidence on which the scoring system was based were not described in the documentation.

**Statistics.** R software was used (R Core Team, 2013) for statistical analysis and graphics. Details of the analysis are described in the Supplementary Material.

**Ethics statement.** This project was approved by the Regional Ethics Committee for Sjælland, Denmark (SF-202-31965). Informed consent was obtained from the blood donors, but not from the LNB cases due to the long-term retrospective nature of the sample collection (originally kept for laboratory quality control). This exemption was allowed by the ethics committee. The Swedish project was approved by the Regional Ethical Review Board in Linköping, Sweden (M83-05 T91-08, 2012/246-31).

RESULTS

**Qualitative results using the software supplied by the manufacturer**

The serum samples from 24 of the 49 Danish patients with LNB were found positive or borderline (Table 1) in IgM or IgG; thus, the sensitivity was 49% (95% binomial confidence interval 36–63; using binom:binom.confint). The specificity of IgG or IgM alone was 33 and 31%, respectively. The specificity measured in the 217 blood donors was 97% (94–99) for IgG and 99.5% (97–100) for IgM. Borderline results were only found in IgG and were counted as positive for calculation of the sensitivity. This seems reasonable as no borderline results were found in the few reactive blood donors. Concerning the Swedish data borderline results were also few. Using the Swedish dataset the sensitivity of serum was (52/61) 85% (74–92) and the specificity (114/139) 82% (75–88).

Thus, for IgG or IgM the sensitivity was 36% (19–51) higher in the Swedish dataset and the specificity was 15% (8–22) lower compared to the Danish set of samples (using Epi:ci.pd). Evaluation data presented in the package insert indicate (28/30) 93% sensitivity in clinically defined LNB and (165/200) 83% specificity in blood donors, which is similar to the Swedish controls. This indicates 14% (9–20) lower background seroreactivity in the Danish blood donors compared to the data from the manufacturer.

**Quantitative sero-epidemiological antibody-response pattern**

The receiver operating characteristic (ROC) curve displays the contrast of the LNB samples to the controls for each antigen (Fig. 1). The area under the curve (AUC) is the probability that the classifier will score a randomly drawn sample from a patient population with LNB higher than a randomly drawn blood donor or non-LNB control.

ROC-curve evaluation of the IgG assay shows that the only antigen with a high discriminatory ability is VlsE IgG (AUC=0.96, Fig. 1). The discriminatory power of the remaining 12 antigens in the IgG assay was lower. One group of six antigens with an intermediate AUC of 0.63 to 0.77 contained OspC, P58, P39 and P18 from *B. garinii*. The remaining six antigens do not have discriminatory power, with the AUC close to 0.5.

With respect to the IgM assay, none of the antigens had a very high discriminatory ability. The three OspC antigens had AUCs from 0.82 to 0.84. There was also an intermediate group, with VlsE, P18 *B. garinii*, P58 and P39 that had AUCs from 0.63 to 0.71. The remaining antigens had
little useful discriminatory ability, with AUCs near 0.5, ranging from 0.46 to 0.55 (Fig. 2). Boxplots to examine the distribution of the measurement values and ROC curves for the Swedish data are displayed in the Figs S1, S2 and S3.

Construction of a diagnostic score

In the recomBead Borrelia IgG and IgM assays the same 13 antigens are used; thus, for each sample 26 results are produced. A multivariate logistic model was used to assess which antigens had significant odds ratios (ORs). The seven IgG and seven IgM antigens with an AUC above 0.60 were used in the initial model. Only IgG for VlsE, P18 and OspC B. afzelii had significant estimates. No IgM results were significant. After the automated stepwise removal of model components, seven variables were kept and, of these, five had significant ORs (Table 2). The significant predictors in the model were IgG VlsE, P18 B. garinii and IgM OspC B. afzelii. The IgG P58 and OspC B. afzelii had negative logistic estimates (OR <1). Predictions from the model using the same dataset show, as expected, high sensitivity and specificity (Fig. 3). Having VlsE reactivity combined with a high P58 reactivity has a negative correlation coefficient of 0.83 (Fig. 2). Also VlsE and OspC B. afzelii IgG, and P58 and P18 B. garinii had negative correlations of 0.43 and 0.51, respectively. Positive correlations were 0.3 or below. Thus, no two antigens seem to be strongly correlated in the group of patients with LNB.

### Table 1. Number of samples evaluated as positive, borderline or negative using the software provided with the assay

<table>
<thead>
<tr>
<th>IgG and IgM (combination of results)</th>
<th>Danish samples</th>
<th>Swedish samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LNB patients</td>
<td>Blood donors</td>
</tr>
<tr>
<td>Pos pos</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Pos neg</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Bor pos</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Bor neg</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Neg pos</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Neg bor</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Neg</td>
<td>25</td>
<td>210</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>217</td>
</tr>
</tbody>
</table>

Bor, Borderline; neg, negative; pos, positive.

### Fig. 1. ROC curve of all 13 antigens in the IgG and IgM assay. AUC values are listed in decreasing order (from the top) for each of the 13 antigens. The same type of figure for the Swedish data is shown in Fig. S3. The red diagonal broken line denotes no difference between the distribution of measurements in the blood donor controls compared to LNB. The abbreviations used are from the data output of the assay system. There are two variations of P18 from B. garinii designated 1 and 2. BAFZ, B. afzelii; BGAR, B. garinii; BSP, B. spielmani; BSS, B. burgdorferi sensu stricto; OSPC, outer surface protein C.
Evaluation of the scoring method using the independent Swedish dataset

Using the regression model based on the Danish data, the Swedish patients have nearly the same ROC curve, except for an area in the upper left corner of the graph (Fig. 3). A sensitivity of 79 and 90% was obtained at the chosen decision threshold at specificities of 98 and 92%, respectively (Table 3). Thus, 11% of the LNB cases fall in the intermediate group. This may be compared to the scoring algorithm provided by the manufacturer where the sensitivity and specificity were both 82% (Table 1). The VlsE IgG assay alone provides nearly the same diagnostic formation as the more complex regression model. However, the specificity was lower using VlsE alone as 28 (46%) fall in the intermediate group. The ROC curves were similar within the 95% confidence intervals of the ROC curve from the predicted Swedish samples and the difference in sensitivity would depend on small changes in the choice of specificity (Dessau, 2013).

Using the logistic model generated from the Danish data (Table 2), the calculated sensitivity for the Swedish data is 79% (Table 3) at the 98% specificity level. Simply using VlsE antigen alone gave a lower specificity for around 50% of the samples. Thus, using the regression model, a higher specificity may be obtained without compromising sensitivity on the independent Swedish dataset. However, the decision threshold values have to be optimized because of the higher background seropositivity. As shown in Table 3, the numerical values of the logistic prediction score are quite different at a fixed specificity of 92 and 98%.

Table 2. Results of logistic regression showing ORs with 95% confidence intervals

<table>
<thead>
<tr>
<th>Antigen</th>
<th>OR</th>
<th>95% CI</th>
<th>P value from the reduced model</th>
<th>P value in original complete model with 14 antigens</th>
</tr>
</thead>
<tbody>
<tr>
<td>VlsE IgG</td>
<td>11.4</td>
<td>5.1–39</td>
<td>0.000001</td>
<td>0.00008</td>
</tr>
<tr>
<td>P58 IgG</td>
<td>0.09</td>
<td>0.02–0.29</td>
<td>0.0008</td>
<td>0.007</td>
</tr>
<tr>
<td>OspC. B. afzelii IgG</td>
<td>0.06</td>
<td>0.01–0.33</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>OspC. B. garinii IgG</td>
<td>3.12</td>
<td>0.89–18</td>
<td>0.13*</td>
<td>0.27</td>
</tr>
<tr>
<td>P18. B. garinii IgG</td>
<td>5.1</td>
<td>1.5–32</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>OspC. B. afzelii IgM</td>
<td>3.5</td>
<td>1.6–9.8</td>
<td>0.006</td>
<td>0.16</td>
</tr>
<tr>
<td>P58 IgM</td>
<td>9.7</td>
<td>0.91–171</td>
<td>0.08*</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*Included in the reduced model, but not significant.
In order to summarize and compare the diagnostic potential for the Danish and Swedish data of each antigen, the AUCs were also calculated for all the antigens using the Swedish data. It may be seen that antigens are pairwise similar in AUCs (Fig. 4). However, the Swedish data have a trend toward higher AUCs compared to the Danish data for most antigens except the OspC IgM for all three species and VlsE IgG (Fig. 4). None of the AUCs of the Swedish data was below 0.5.

Reproducibility and analytical quality of the multiplex assay

Concerning the individual antigens, the mean interassay coefficient of variation was 30% for both IgG and IgM; however, this was much less for the logistic score (Fig. S4). For details of the assessment of analytical quality see the Supplementary Material.

DISCUSSION

This study shows that using VlsE antigen alone would provide most of the diagnostic information. Adding more antigens and calculating a logistic regression score enhanced specificity. Thus, a combination of biomarkers may provide a better diagnostic tool than a single test can on its own. However, it was not shown that using more than 7 of the 26 provided antigens would provide better diagnostic utility. The scoring additive algorithm provided by the manufacturer was suboptimal concerning both sensitivity and specificity. There was a remarkable difference in sensitivity of 48% in the Danish set compared to 85% in the Swedish set of LNB cases.

We have not found a similar detailed analysis of seroreactivity in the literature, and a previous model found for the analysis and presentation of such complex data was not found. Thus, this study experimented with the presentation

| Table 3. Comparison of the 49 Danish patients and the 61 Swedish patients standardizing the specificity at 98 and 92% |

<table>
<thead>
<tr>
<th>Sample</th>
<th>Decision threshold values</th>
<th>Unit</th>
<th>Probability from logistic prediction</th>
<th>Index units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish samples</td>
<td>0.27, 0.04</td>
<td>98%</td>
<td>0.94, 0.64</td>
<td>95%</td>
</tr>
<tr>
<td>Swedish samples</td>
<td>0.54, 0.64</td>
<td>99%</td>
<td>0.51, 0.34</td>
<td>96%</td>
</tr>
<tr>
<td>VlsE Danish</td>
<td>1.24, 18.0</td>
<td>96%</td>
<td>1.24, 18.0</td>
<td>97%</td>
</tr>
<tr>
<td>VlsE Swedish</td>
<td>1.24, 18.0</td>
<td>97%</td>
<td>1.24, 18.0</td>
<td>97%</td>
</tr>
</tbody>
</table>

Fig. 4. Pairwise comparison of the AUCs for the ROC curve for the 13 antigens in IgG and IgM for the Danish and Swedish data. The diagonal line is the expected identity. Black circles, IgG; red circles, IgM.
of complex data and the main objective was to provide a rationale for clinical interpretation of this complex assay for the diagnosis of infection with *B. burgdorferi sensu lato*. As demonstrated in this study, both the sensitivity and specificity were high in LNB patients using a blood donor control group or a cohort of patients with suspected LNB. The important issue is to use evidence-based interpretation as the chosen Boolean interpretation rules specified by the manufacturer were not found to be optimal.

The main principle of analysis is to find the contrast pattern between the case and control samples. Both ROC curves and the binary regression model are tools to display this contrast.

**Sero-epidemiological antibody-response pattern**

A relatively low serum antibody reactivity was found. This supports the use of the antibody index comparing reactivity in the spinal fluid and serum for the diagnosis of LNB (Henningsson *et al.*, 2014; Mygland *et al.*, 2010; Stanek *et al.*, 2011). Only the VlsE antigen gave a high reactivity with a good contrast to the controls in this study on serum antibodies in patients with LNB. Notably, only OspC from the IgM assay gave any potentially useful diagnostic contrast; otherwise the IgM assay was quite non-reactive.

In this study, patients with LNB are (conveniently) used as a model for systemic early LB. This assumption is questionable and studies reporting on the use of this multiplex assay in patients with other manifestations of LB are needed. However, samples from other early manifestations, such as Lyme carditis and lymphocytoma, are difficult to collect as they have a relatively low incidence. Concerning erythema migrans one could expect more IgM reactivity, and concerning acrodermatitis chronica atrophicans or Lyme arthritis, reactivity to most of the IgG antigens could be expected. The LNB patients could all have a short duration of clinical disease; this is, however, unlikely as they were preselected by a positive antibody index in the spinal fluid and most were VlsE IgG reactive. The Swedish set of cases had a median duration of symptoms of 21 days, with a range of 1-224 days, and it must be noted that all the Swedish LNB cases were positive when the antibody index was calculated for the recomBead assay (Henningsson *et al.*, 2014). Another explanation for the low reactivity could be technical as it must be a difficult to optimize the detection of each individual antigen as buffer conditions and serum dilution are the same. This assumption was corroborated by the fact that positive outliers (Fig. S1) have low reactivity in both LNB patients and controls (excepting the VlsE IgG and the three OspC IgMs). The majority of patients with LNB should be expected to be infected with *B. garinii* (Stanek & Reiter, 2011). However, the OspC antigens of *B. afzelii* and *B. burgdorferi sensu stricto* were equally reactive, and with similar ROC curves (Figs 1 and 4) for IgM. Thus, these three antigens appear interchangeable without the need for including all three. Another reason for low reactivity in the Danish or the Swedish samples could be due to geographical variation. Possible geographical variations of antibody development in Europe have been found in other studies (Hauser *et al.*, 1999; Robertson *et al.*, 2000), but there are no published data concerning the recomBead assay. Another reason for low reactivity to most of the antigens could be that the recombinant antigens may not quite resemble the natural proteins and, thus, may to some extent miss the targeted antibodies. An interesting finding is that some P58 IgG and OspC *B. afzelii* IgG have a relatively high frequency in the controls and could thus be markers of old immunity.

**Construction of a diagnostic score**

The manufacturer has chosen to set a common cut-off for each antigen, and then scores each positive result as 4 points and intermediate results as 1 point. Then a sum of 8 points is positive and 5-7 points is borderline. No rationale for the Boolean scoring algorithm or the choice of decision threshold was provided. The same type of Boolean scoring algorithm as provided by the manufacturer of the recomBead assay is typically specified in immunoblots. However, using Boolean operators may be problematic when more than just two tests are considered (Pepe, 2003). In this case where 26 biomarkers were measured it is not feasible to find the best sets of rules among $2^{26}$ possible combinations. Thus, simple and reliable Boolean rules are impossible to construct for combinations of multiple biomarkers from a theoretical point of view.

Instead a risk score may be used and logistic regression is a practical method to optimize a combination of biomarkers. This score may be determined a risk score defined as an increasing function. The larger the score, the larger the probability is of belonging to the diseased group. Based on this one score, thresholds may then be chosen in order to aid clinical interpretation. In this case, a specificity at 92 and 98% was standardized to define the decision threshold (Dessau, 2013). The strategy to use the full quantitative range including the weaker reactivity was inspired by previous measurements of the flagella antigen in serum (Dessau *et al.*, 2010b; Dessau, 2013). It was noted that combining IgM and IgG of the same assay could increase diagnostic accuracy of weak serum reactivity also (Dessau *et al.*, 2010b). Thus, to exploit the full potential of the recomBead assay the weaker antibody responses should be included in the scoring algorithm without setting a cut-off for each individual antigen. The logistic regression score using the complete range of quantitative results was shown to increase specificity without loss of sensitivity. Certain antigens or antigen combinations correlated negatively and thus should be subtracted, as they have a relatively high reactivity in the control population. The regression methodology allows for scoring based on evidence from patients and the utility of this approach was proven by similar results for a second independent dataset.

Published studies on the detailed reactivity pattern to multiple antigens are scarce. Studies on immunoblots are not quantitative and are usually quite small. A substantial
control group comparing the contrasts for each individual antigen has been included in the present study. When using the two-tier algorithm with preselection of only ELISA positives a larger disease-negative control group has usually not been characterized. It is important to establish which antigens have diagnostic utility. Even though a two-step procedure using immunoblots has been recommended to improve the diagnostic accuracy for many years, studies trying to establish evidence-based scoring algorithms are not discussed in literature reviews (Aguero-Rosenfeld et al., 2005; Aguero-Rosenfeld, 2008). The concept and recommendation for the two-step procedure was originally established in a consensus conference in 1995 (CDC, 1995). Attempts were subsequently made to establish a similar recommendation for Europe, but the presented evidence did not support a single scoring algorithm, presumably due to regional strain heterogeneity (Hauser et al., 1997, 1999; Wilske et al., 1998). It was concluded that immunoblot interpretation should be adapted in relation to the characteristics of LB in local areas (Wilske et al., 2000). There is a continuing development of recombinant antigens that are implemented in commonly used commercial assays on the market in Europe (Hunfeld & Kraiczy, 2009). Immunoblotting has not been used routinely in Denmark (Dessau et al., 2010a, 2011; Dessau, 2013). However, as shown in this study, the heterogeneity problem may be due to the chosen scoring algorithm rather than differences in regional strains of the bacterium or study population.

**Evaluation of the proposed scoring method using an independent dataset**

The Swedish dataset was used to validate the logistic scoring algorithm, and was shown to provide sensitivities of 79 and 90% at a specificity of 98 and 92%, respectively. The AUCs were comparable around 98%. For the individuals antigens (Fig. 4) the AUCs were somewhat higher in the Swedish samples for most of the antigens except VlsE IgG and OspC IgM. It is possible that the higher background immunity in the Swedish population could activate a broader range of antibody reactivity in the average LNB patient after a shorter duration of clinical disease. Laboratory background problems are possible but somewhat unlikely, as not all the antigens have higher reactivity or contrast in the data from the Swedish laboratory (Figs S1 and S2).

**Perspectives**

The study may also be read as a more general method to score multiplex assays in a standardized manner whether for LB or other diseases. LB antigen patterns appear to vary according to geography, but this has little impact on the ROC curve with the logistic regression score. We have shown that the same assay works quite well both in Denmark and in Sweden, but decision thresholds may have to be adjusted to suit local needs. This is probably more due to differences in background immunity in the human populations than variation in the antigen expression by the bacterium. We propose this hypothesis from looking at Fig. 4, where the size of the pairwise AUCs is comparable, indicating antibody development to the same antigens on the average. This is also consistent with features of *B. burgdorferi sensu lato* antigens like the native flagella, VlsE and OspC are well conserved between the pathogenic species; thus, it should be possible to use the same assays all over Europe as is done with the commonly used ELISAs.

**Conclusion**

In contrast to VlsE IgG, there was surprisingly little antibody reactivity and diagnostic contrast in all the 25 other IgG or IgM antibody measurements. These findings should be investigated in patients with systemic or disseminated manifestation of LB outside the central nervous system as these patients could have a stronger and more diverse antibody-reactivity pattern in serum. Reactivity in both P58 and VlsE IgG was relatively frequent in the blood donors and this combination indicates background seroreactivity rather than active disease. The scoring algorithm provided with the recomBead assay was not optimal. When using a logistic regression model seven antigens were identified to maximize diagnostic accuracy. Thus, the total of 26 antigens in both IgG and IgM do not appear necessary to improve the diagnostics. The same risk score model may be used in Denmark and Sweden; however, decision thresholds should be adjusted to the local background seroreactivity.

**ACKNOWLEDGEMENTS**

The project was supported by local funding from the Research Committee of Region Zealand. The first author is a member of the ESCMID (European Society of Clinical Microbiology and Infectious Diseases) study group for Lyme borreliosis (ESGBOR). Discussions in the ESGBOR network have been important in providing background knowledge for the present study.

**REFERENCE**


