Study of a Cohort of 1,886 Persons To Determine Changes in Antibody Reactivity to *Borrelia burgdorferi* 3 Months after a Tick Bite

Ram B. Dessau, Linda Fryland, Peter Wilhelmsson, Christina Ekerfelt, Dag Nyman, Pia Forsberg, Per-Eric Lindgren

Department of Clinical Microbiology, Slagelse Hospital, Region Zealand, Slagelse, Denmark; Infectious Medicine, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden; Medical Microbiology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden; Clinical Immunology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden; Åland Borrelia Group, Åland University of Applied Sciences, Mariehamn, Finland; Clinic of Infectious Diseases, University Hospital, Linköping, Sweden

Lyme borreliosis is a tick-borne disease caused by the bacterium *Borrelia burgdorferi*. The most frequent clinical manifestation is a rash called erythema migrans. Changes in antibody reactivity to *B. burgdorferi* 3 months after a tick bite are measured using enzyme-linked immunosorbent assays (ELISAs). One assay is based on native purified flagellum antigen (IgG), and the other assay is based on a recombinant antigen called C6 (IgG or IgM). Paired samples were taken at the time of a tick bite and 3 months later from 1,886 persons in Sweden and the Åland Islands, Finland. The seroconversion or relative change is defined by dividing the measurement units from the second sample by those from the first sample. The threshold for the minimum level of significant change was defined as the 2.5% level to represent the random error level. The thresholds were a 2.7-fold rise for the flagellar IgG assay and a 1.8-fold rise for the C6 assay. Of 1,886 persons, 102/101 (5.4%) had a significant rise in antibody reactivity in the flagellar assay or the C6 assay. Among 40 cases with a diagnosis of Lyme borreliosis, the sensitivities corresponding to a rise in antibodies were 33% and 50% for the flagellar antigen and the C6 antigen, respectively. Graphical methods to display the antibody response and to choose thresholds for a rise in relative antibody reactivity are shown and discussed. In conclusion, 5.4% of people with tick bites showed a rise in *Borrelia*-specific antibodies above the 2.5% threshold in either ELISA but only 40 (2.1%) developed clinical Lyme borreliosis.

*Borrelia burgdorferi sensu lato* is transmitted by hard ticks (*Ixodes ricinus* or *Ixodes persulcatus*) throughout Europe, especially in the temperate climate zone. Both the vector and the bacterium are prevalent in southern Scandinavia, and ticks frequently attach to humans and occasionally transmit *B. burgdorferi* (1–4). Lyme borreliosis (LB) is also common in parts of the United States. The term *B. burgdorferi sensu lato* denotes a species complex belonging to the genus *Borrelia* and is herein designated simply *B. burgdorferi* or *Borrelia*. The most frequent species infecting humans in Europe are *B. afzelii*, *B. garinii*, and *B. burgdorferi sensu stricto*. In the United States, the only prevalent species is *B. burgdorferi sensu stricto*. The infections may involve the nervous system, the skin, the joints, and the heart. The most frequent manifestation is a characteristic skin lesion called erythema migrans (EM).

This study was based on data from the tick-borne diseases (TBD) STING study (www.stingstudien.se). Ticks that had bitten humans in the Åland Islands, Finland, and different parts of Sweden were collected and analyzed at Linköping University (2–5). The focus of this study was to analyze the differences in the measurements of *B. burgdorferi*-specific antibodies in the paired samples and propose methods for detecting changes in antibody reactivity. Although the concept of antigen-induced stimulation and seroconversion has broad application in diagnostic medicine, epidemiology, and vaccine studies, methodological studies on how to evaluate the paired samples from each individual and establish decision thresholds have been scarce. The focus of interest is on detecting a rise in antibody reactivity indicative of antigen stimulation. The usual cutoff values based on interindividual variation with “positive” or “negative” classification are not optimal for comparisons of samples from the same individual. This large collection of systematically retrieved material, with paired samples collected 3 months apart from members of a tick-exposed population, provides a unique opportunity to analyze and set standards for changes in antibody reactivity.

**MATERIALS AND METHODS**

**Materials.** Serological samples and data collected from 1,886 persons in the TBD STING study bitten by one or more ticks from 2007 to 2009 were analyzed (2–5). A total of 394 of these participants were also described previously (3). The participants had blood samples taken at the time of the tick bite and 3 months later. Concerning the diagnosis of LB, if the participants had visited a health care provider during the 3-month study period, the medical record was obtained and scrutinized in detail by physicians with extensive clinical experience of LB and other tick-borne infections (6). Diagnostic criteria were according to European guidelines (7, 8). Samples were stored frozen at −70°C until they were analyzed.

**Ethical permission.** This study was approved by the Regional Ethical Review Board, Linköping University (M132-06), and by the local Ethics Committee of the Åland Health Care (2008-05-23).
Data from 1,886 cases were available. Among the patients represented by the cases, 668 (35%) were men and the median age was 63 years (range, 19 to 92 years). Forty (2.1%) people with tick bites were diagnosed with LB (39 certain cases and 1 possible case) in the interval between the collection of the first sample and the collection of the second sample. Of these, 34 were diagnosed with EM, 2 with unspecified LB, 1 with borrelial lymphocytoma, and 2 with neuroborreliosis, the latter 2 being both antibody index positive and positive for pleocytosis in the spinal fluid. The patient in one case had facial palsy, but no lumbar puncture was performed, so this case did not have a certain diagnosis of LB. This patient (male, 60 years old) had high levels of antibody in both assays, with a nonsignificant level of relative change (RC). All 40 cases were used in the subsequent analysis. All 40 patients diagnosed with LB received routine antibiotic treatment.

**Assays.** The samples were analyzed with native purified IgG flagellar antigen (IDEIA; Oxoid, United Kingdom) and C6 (B. burgdorferi Lyme enzyme-linked immunosorbtent assay [ELISA]; Immunetics, Boston, MA, USA). The C6 assay is based on a recombinant invariable fragment of the VlsE antigen. The C6 assay does not distinguish between IgG and IgM antibodies. Both assays have been described in previous publications from Scandinavian studies (9–19). The C6 assay was implemented on a Bio-Rad Coda microplate processor, and the IDEIA ELISA was performed manually. Paired samples were analyzed in the same ELISA plate.

**Methods.** Calibrated units were calculated as the optical density (OD) of the sample divided by that of the calibrator for both the flagellar IgG and C6 assays. For both assays, the RC was defined as the unit value for the second sample taken 3 months after the tick bite divided by the unit value for first sample taken at the time of the tick bite as follows: unit value = OD of sample/OD of calibrator; RC = unit value for sample 2/unit value for sample 1.

Thus, the calibrator was normalized at a value of 1 and the RC was close to a value of 1 for the cases in which there was no change in antibody reactivity at 3 months. In this manner, it was possible to analyze and exploit the full range of measurements, and the units used in this study were not calculated as specified by the manufacturer of the assays. Using the “normal” decision threshold with a simple negative or positive interpretation may also detect false “seroconversion” due to simple random variation in two nearly identical measurements close to the cutoff value. The setting of thresholds for changes in antibody reactivity was developed in this study.

**Statistics.** All statistical analyses and graphics were performed with the freely available R statistical software (20). Implementation of the statistical methods was part of the study, and details of the methods are explained in the relevant sections of Results. Statistical significance was defined as observations exceeding the 2.5% threshold. Further details are provided in the supplemental material.

**RESULTS**

**Background seroreactivity.** Using the routinely specified cutoff values, 212 (11%) and 732 (39%) samples gave positive results in the flagellar IgG and C6 assays, respectively, indicating high background immunity. Looking at the distribution of units (Fig. 1), there is an obvious difference between the two assays. The native flagellar antigen gives a left tail with a decreasing frequency of cases with high values (Fig. 1A). The C6 assay shows a marked bimodal distribution with a cluster of high unit values. However, the cutoff value for the C6 assay is placed within the distribution of the low-reactivity samples, which explains the high background seropositivity (Fig. 1B).

**Change in antibody reactivity 3 months after a tick bite.** Scatter plots with sample 1 data on the x axis and sample 2 data on the y axis show many points scattered along the diagonal where the x values and the y values are equal (Fig. 2). This shows that most persons do not develop a rise (or fall) in antibody reactivity after a tick bite and that, in general, antibody measurements in the same individual are reproducible, even when samples are measured 3 months apart (Fig. 2). The cases with a rise in reactivity in the second sample are scattered in the upper left part of Fig. 2. Those are the main targets of interest in this study. To show the statistical distribution, upper and lower percentiles are shown together with the 40 patients who developed LB between sample 1 and sample 2 collected 3 months later. There are some cases which showed a fall in antibody reactivity (gray color), and these data represent either random measurement variations or a faster downregulation of the antibody response in a few individuals. Downregulation could have been the case, especially with the C6 assay, for a few individuals who had very high values in the first sample (gray circles in the lower right part of Fig. 2). This was probably true, as these outliers were not seen among the low-measurement values and thus were present only in individuals with some reactivity in the first sample. Note that none of the 40 LB cases had a fall in antibody reactivity and that the reactivity data for many remained the same.

The changes in Borrelia-specific antibody reactivity in the same individual in the two assays were compared using the RC value (defined as the value from the second sample divided by the value from the first sample) (Fig. 3). The pattern shows that most of the reactive samples tended to show increased antibody reactivity in both assays. Data corresponding to samples with low reactivity in the second sample were removed from Fig. 3, except for the four cases diagnosed with LB. These lower measurement values were considered to be “truly” nonreactive without an increase in reactivity and were thus filtered out. The values of 0.1 and 1.5 for the flagellar and C6 assays, respectively, were chosen by visual inspection of Fig. 2 to rule out larger RCs in very low measurement values (“low-level noise”) and at the same time to include some LB cases with a lower degree of reactivity. A total of 331 cases, including three LB cases, were nonreactive in both assays by this definition (see Fig. S1 in the supplemental material).

A receiver operating characteristic (ROC) curve (Fig. 4) was constructed to assess if RC values determined using two samples would be more accurate for diagnosis of erythema migrans than the use of the data from single sample 2 alone. RC was significantly more specific up to a sensitivity of 60%. Concerning RC, the C6 assay appears to have been slightly more specific than the flagellar IgG assay but only at around a sensitivity of 40%.

**How to determine a threshold value for a significant rise in antibody reactivity.** Different methods of establishing a threshold for RC were considered.
(i) Use of the upper percentiles of RC as done for Fig. 2. As the distribution is skewed due to the rises in the antibody reactivity that we actually want to detect, this approach may be considered too conservative.

(ii) Use of the ROC curve (Fig. 4) based on LB cases (mainly EM). However, the ROC determined using the clinical diagnosis for classification excludes the subclinical antibody stimulation after the tick bite, which we also want to characterize. Thus, possible increases in antibody reactivity are considered controls.

(iii) Possible use of the mirror-symmetrical value of 2.5% from the lower percentile (Fig. 2, gray points), assuming that this represents random error and is a symmetrical phenomenon. The lower percentiles for RC are thus assumed to represent measurement errors and variations in samples collected from the same individual on two occasions 3 months apart. We also assume that the natural logarithmic transformed measurements have similar RC values for all reactivity levels along the diagonal line of identity (Fig. 2). This is not strictly the case due to low- and high-level limits in the measurement range, and outliers appear mostly in the middle parts of Fig. 2A and B and in the lower part of panel 2A.

Therefore, the 2.5% lower-percentile value was chosen as a threshold to represent the random error of measurement (see also Fig. S2 in the supplemental material). The thresholds were 2.7 for the flagellar IgG assay and 1.8 for the C6 assay. Using these thresh-

FIG 2 Results determined for the first sample collected when the tick was detected and for the second collected 3 months later for the flagellar IgG (A) and C6 (B) assays. The optical density reading of each sample was divided by the cutoff value. To indicate a statistically significant rise in seroreactivity, the values above 95% of the RC in seroreactivity are indicated with colored circles. Data from patients diagnosed with EM or other manifestations of LB are shown with larger solid points. All very low measurements below 0.1 and 1.5 for the second sample for the flagellar IgG and C6 assays, respectively, are recorded as representing no change. The scales are logarithmically transformed (base ln).

FIG 3 The RC over 3 months in seroreactivity for the flagellar IgG and C6 assays plotted to show coreactivity. A total of 40 patients with LB and 1,518 tick-bitten cases of patients without LB are plotted. Data from 328 non-LB cases with very low absolute measurements below 0.1 and 1.5 in the second sample for the flagellar IgG and C6 assays, respectively, have been removed.

FIG 4 ROC curve of data from the 40 cases of patients with LB compared to people with tick bites without LB as controls. For the RC determinations, the data from the 331 cases of patients with very low values have been removed. The C6 curve based on RC is shown with 95% bootstrapped confidence intervals. The sensitivities and specificities corresponding to RC thresholds of 2.7 for the flagellar IgG assay and 1.8 for the C6 assay are shown as black and red dots, respectively. The second serum samples alone have a low discriminatory power.
olds, of 1,886 people with tick bite, 102 and 101 (−5.4%), respectively, were classified as having a significant rise in antibody reactivity for each assay separately. A total of 59 (3.1%) had a rise in that value in the flagellar assay, 58 (3.1%) had a rise in that value in the C6 assay, and 43 (2.3%) had a rise in that value in both assays. Of the 40 cases diagnosed with LB, two (5%) had a titer rise in the flagellar IgG assay only, 9 (23%) had a rise in the C6 assay only, and 22 (11%) had a rise in both assays. Thus, the rates of titer rises in the remaining 1,846 cases without LB were 89/1,846 (4.9%) and 81/1,846 (4.4%) for the two assays, respectively (see also Fig. S1 in the supplemental material).

DISCUSSION

Relative change (also called “fold rise”) determinations and plotting of the pre- and post-antibody responses on a logarithmic scale (Fig. 2) have been used in vaccine studies (21). These types of logarithmic plots have the advantage of displaying both the original measurements and the RC as a straight line parallel to the diagonal no-response (x = y) line. In addition, as shown in this study, a threshold for a rise in antibody reactivity may be set using the percentiles, which appears biologically plausible as judged from Fig. 2. Thresholds may be set in various ways, but the use of the percentile does not involve the assumption that underlying (parametric) distributions and a symmetric standard deviation would be inappropriate. More-complex models could be used, but they involve more complexity and assumptions about the shape of the underlying distributions (22, 23). In our case, we would have two assays with different distributional profiles needing separate modeling. A second low-reactivity-level threshold is also chosen to filter out large RCs when the absolute measurements are small. The choice of thresholds in this study was based on the distribution of measurements. Percentiles were chosen because of the skewed distribution of RC, as a regression line would assume symmetrical standard deviations from the straight line. Thus, the maximal analytical capacity of the assays to detect change is used. In this case, there is no independent standard on which to base the thresholds for a titer rise as we may not know in which of the 1,886 tick-bitten persons a true immune stimulation did occur. Vaccine studies are different, as such studies represent donors and consecutive routine specimens has been reported previously (13, 29). The flagellar IgG assay and the C6 assays were chosen because they were routinely used by the laboratories in the region. A separate IgM assay was not chosen, as this was considered to give a high nonspecific background, and has not been used for clinical routine diagnosis in the local diagnostic laboratories. Immunoblotting was not performed, as this has not been used routinely in the region. A Western blot or immunoblot confirmation is not a requirement in Sweden. In a future project, it would be interesting to investigate changes in patterns of reactivity to multiple different antigens. Methods to model changes in reactivity patterns are complex and need to be developed. Perhaps a multiplex assay giving quantitative results and using a score based on generalized linear modeling should be considered (30).

In our study, 40 (2.1%) study subjects developed LB, which is higher (95% confidence interval [CI], 0.01% to 0.28%) than the rate determined in a previous report on a subset of the same study population, where none of the 394 study subjects developed LB (3). However, the message, i.e., that the risk of LB after a tick bite remains low and that testing for Borrelia-specific antibodies is not relevant after a tick bite, remains the same. Also, in patients with EM, the diagnosis is clinical, as the sensitivity of serology for determinations of both increased antibody levels in a single sample and a rise in antibody levels in paired samples is low.

The ROC curve shows that paired samples, if available, represent a better diagnostic marker than a single first sample alone. However, in the normal clinical situations, the samples are rarely taken before the development of disease. In conclusion, 5.4% of people with tick bites developed a rise in the level of Borrelia-specific antibodies above the 2.5% threshold in one of the two ELISAs, but only 40 (2.1%) developed clinical LB. For diagnostic purposes, a 2-fold to 3-fold or greater rise in antibody reactivity may be considered a sign of Borrelia-specific immune stimulation.

indicative of a true (subclinical) exposure to B. burgdorferi or of a non-Borrelia cross-reacting stimulus. However, as Borrelia IgG appears quite specific in other populations, such as blood donor controls or patients suspected of Lyme arthritis, the majority of significant increases in seroreactivity should be indicative of endemic exposure to B. burgdorferi resulting in a (subclinical) infection. In Europe, asymptomatic infections in exposed populations are considered quite frequent (24–28). This is in contrast to the results seen with the recombinant C6 antigen, which had a bimodal distribution, meaning that there was a “positive” group of cases with strong binding of the antigen and a separate “negative” group with low-level reactivity in the test system. The native flagellar antigen is a more complex and larger molecule from cultured Borrelia bacteria which probably has several binding sites with some biological variability. The C6 assay has a high background seroreactivity level of 39%, and, due to the bimodal distributions, it would not help much to raise the cutoff value. It is very interesting that the native flagellar antigen has much lower reactivity in general but has the same ability to detect a rise in antibody reactivity. The flagellar antigen data have larger variability around the no-change diagonal in Fig. 2A, and a higher threshold was needed at a 2.7-fold rise than at a 1.8-fold rise in the C6 assay. Nevertheless, the flagellar antigen had higher specificity in this endemic population in analysis of the single first sample. Concerning the C6 assay, a high rate of seropositivity in both blood donors and consecutive routine specimens has been reported previously (13, 29). The flagellar IgG assay and the C6 assays were chosen because they were routinely used by the laboratories in the region. A separate IgM assay was not chosen, as this was considered to give a high nonspecific background, and has not been used for clinical routine diagnosis in the local diagnostic laboratories.

Immunoblotting was not performed, as this has not been used routinely in the region. A Western blot or immunoblot confirmation is not a requirement in Sweden. In a future project, it would be interesting to investigate changes in patterns of reactivity to multiple different antigens. Methods to model changes in reactivity patterns are complex and need to be developed. Perhaps a multiplex assay giving quantitative results and using a score based on generalized linear modeling should be considered (30).
REFERENCES


