Reviewer's report

Title: Utilization of serology for the diagnosis of suspected Lyme borreliosis in Denmark: Survey of patients seen in general practice

Version: 3 Date: 24 September 2010

Reviewer: Pascal Meylan

Reviewer's report:

I believe that the controversy between the authors and reviewer 2 is rooted in the misunderstanding that this study is a study of serology utilization, not of serology performance, in the diagnosis of Lyme borreliosis. On one hand, the authors are clear about that study goal (see title or the background section). On the other hand, they express more than once opinions suggesting that the single tier approach used in Denmark (though not the International Standard) is performing quite well, based on blood donor data (see page 3 laboratory methods). In fact, based on personal experience, I have major doubts for instance on the specificity of flagellin based IgM ELISAs, based on IB data in particular. But if one looks at the present study data, is it likely that regarding an infection that leaves long lasting IgG blood levels (unless treated early) IgM seroprevalence would be 3 fold that of IgG? Unless the majority of infections are caught and treated at an early time (which may be true for patients seeking care, but not for a population), leaving a majority of patients with transient IgM responses only, the alternate hypothesis that flagellin IgM specificity in patient population is much lower than in blood donor sounds a lot more likely.

I note also that a good deal of the discussion ("seropositivity rates and the clinical variables) is devoted to discussing how the OR of having specific IgG or IgM are related to the clinical presentation, providing a kind of post hoc validation of the tests.

As an example, see page 9 second paragraph. I think that it is close to be abusive to try to deduce a test specificity from its use in this patient population (suspected Lyme arthitis), assuming that at that stage all patients should be IgM negative.

In the end, the limits of the present paper re test performance, as acknowledged by the authors, lie in the lack of gold standard (short of detailed clinical and serological characterization, or PCR or culture documentation of infection), which can only be performed in (non representative) subpopulations. Of note also, there is no mention of the fact that IgM result interpretation should be very cautious in those patients with more than 4-8 weeks of illness. Still, I find a lot of value in this paper that, in quite original manner, present the use and limits of current Lyme testing in large population.

Level of interest: An article of importance in its field
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests