Reviewer's report

Title: Molecular Epidemiology of Hepatitis B Virus Infection in Switzerland: A Retrospective Cohort Study

Version: 1 Date: 29 June 2015

Reviewer: Carla Osiowy

Reviewer's report:

Molecular epidemiology of hepatitis B virus infection in Switzerland: A retrospective cohort study

Hirzel, C., et al. BMC Infectious Diseases (Submitted: 2015-03-13)

Major Compulsory Revisions:

1. Page 5 (line 111): Please provide more detail on how “region of origin” was determined. For example, do medical charts include the patient’s birthplace or country of emigration? How is a person of “Swiss origin” determined? This classification has bearing on the statistical analysis and final conclusions, and so requires clarification.

2. Page 6 (line 140): Although the geno2pheno tool may estimate subgenotype designation, subgenotypes can only be definitively determined by full genome phylogenetic analysis. This should be discussed or included as a limitation, and kept in mind throughout the Results and Discussion sections.

3. Pages 6 and 7 (Serologic Analysis): To what extent do serology results stated reflect testing in the authors’ reference laboratory on the same samples used for HBV genotype analysis? If the majority of serology results are from charts that “were performed elsewhere” (line 150), it reduces the confidence of interpretation, as the methodology is not known, and a previous result may not be coincident with the genotyping result stated or may not indicate coinfection with HBV, etc.

4. Page 10 and Fig. 3 (Phylogenetics): It is curious that four sequences from patients of Asian origin were determined to cluster with subgenotype D4, as this has not really been described before (except a report of C/D recombinants from Tibet – J Med Virol 2014 86:1307). As mentioned above, subgenotyping cannot be accurately determined with a 413 nt region of the HBV genome, and so should be included as a limitation. However, I would recommend re-analyzing the data by including all HBV/D subgenotype GenBank reference sequences (D1-D6) to fully evaluate the study genotype D sequences (reference sequences for all subgenotypes can be found in Kramvis, A. (2014) Intervirology 57:141-150). Furthermore, bootstrap analysis was described as being carried out but no bootstrap values are shown in Fig. 3.

Minor Essential Revisions:
1. Page 5 (line 113): Presumably all patients consented to being tested for HBV genotyping through their medical doctor who requested the reference test for patient management purposes – is this correct? Although written informed consent was not directly provided to the study authors, it is still important to state that patient consent was granted through the public health/medical system.

2. Page 6 (line 125): The primers HBV P1F_f and HBV S6_r are not found (by name) in the Schildgen, et al. reference (#11). Please correct.

3. Page 7 (line 153): How were the 136 sequences included in the phylogenetic analysis chosen? Are these sequences representative of all populations by region of origin? For example, 43 patients of Swiss origin were genotype A, but only 10 sequences are included in the phylogenetic analysis.


5. Supplementary Table 2 and Figure 3: An omission or error in either or both was observed: HK (HBV/C1) and MA (HBV/D4) are part of Fig. 3 but not listed in Table S2.

6. Table 1: Check that the p-value is lined up with the appropriate data for the statistic. The values for sex and genotype (?) are difficult to interpret.

Discretionary Revisions:

1. Abstract (line 67-68): “HBV genotypes of patients living in Switzerland but sharing the same original region of origin were closely related”. It may be better to say, “HBV genotypes of patients living in Switzerland but sharing the same original region of origin were consistent with their place of birth (or region of origin)”, as “closely related” may be misconstrued as a sequence or phylogenetic relationship.

2. Page 9 (lines 206-207): It might be helpful to describe the countries included in “Other”, and their proportion within the Other/Unknown category. Particularly, as the prevalence of genotype G (including A/G) is relatively high (4/77; 5.2%) within this group compared to all other regions of origin.

3. Page 10 (HBeAg positivity): Patients of Swiss origin were most likely to be HBeAg positive (38.1%) yet had the oldest median age (45). As multivariable analysis described an association of older age with a lower probability of being HBeAg positive, perhaps this observation could be discussed.

**Level of interest**: An article whose findings are important to those with closely related research interests

**Quality of written English**: Acceptable
**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.