**Hepatitis C virus genotypes in eastern Algeria: a retrospective study**

**Biostatistician Review Report**

Samir Rouabhia1 , Mourad Sadelaoud2, Karima Chaabna-Mokrane3,

Wided Toumi4, Ludovico Abenavoli5

1- Associate Professor. Department of Internal Medicine, University Hospital Center Touhami Benfis, Batna 05000, Algeria

2,4- Medical Doctor. Sadelaoud Laboratory of Medical Biology, La verdure, Batna 05000, Algeria

3- M.P.H. Section of Cancer Information, International Agency for Research on Cancer, Lyon, France.

5- Assistant Professor. Department of Health Sciences, University “Magna Græcia”, Viale Europa, 88100, Catanzaro, Italy

Electronic Word Count (excluding abstract and references): 1550 words.

**Corresponding author**

Samir Rouabhia

Department of Internal Medicine,

University Hospital Center Touhami Benfis,

Batna 05000, Algeria.

Tel.: +213 7 71 39 82 90

Fax: +213 33 92 60 66

E-mail: rouabhiasamir@yahoo.fr

**Objective:** The aim of our study is to identify the frequency of different HCV genotypes in eastern Algeria and to assess the correlation between the HCV genotypes and the demographic profile.

This report details the statistical method used to describe and analyze the HCV genotype in the study population.

**Materials and Methods**

*Study population and data collection*

The study population and how data were collected are described in the section Material and methods of our article. A map is also presented to give a geographical view of the regional medical laboratory coverage where our data were from (Figure 1). This is a monocentric study; however, Sadelaoud Laboratory has been the only one in the Eastern Algerian region which carries out HCV genotyping. The patients who have been diagnosed with the HCV infection in Eastern Algerian local hospitals and have needed to assess the HCV genotype before the antiviral treatment, have been free to perform their test in another medical laboratory e.g. in the capital Algiers (North-central Algeria) or to send their blood sample to a foreign country (France). Because of logistic and road security reasons and the cost of testing blood samples in another laboratory in another city, we assumed that the proportion of HCV positive patients living in Eastern Algeria who did not perform their viral genotyping test at Sadelaoud Laboratory is low. Thus, the results obtained in this study may give a point of view of the frequency of HCV genotypes in Eastern Algeria.

The retrospective study presented in the current article includes 435 HCV infected patients tested between January 2010 and December 2012 at the regional medical laboratory in Batna.

*Data analysis*

To assess differences in patients’ characteristics of genotype, viral load, age, sex and region, Pearson's Chi-squared and Fisher's exact tests were used when the number of cases in each cell of the contingency tables was higher than 5 and lesser than 5, respectively. The Yates' continuity correction was used to make the chi-square approximation better. The null hypothesis was rejected if tests assign p-values equal to or smaller than 0.05. Age and genotype were categorized in two groups <60 years and ≥60 years; and in two groups “Genotype 1” and “Others” respectively. The patients studied lived in fifteen wilayas (provinces), the administrative regions which cover the eastern area of Algeria. These wilayas are classified in three eastern regions for this study: the central region covering five wilayas (Batna, Khenchela, M’Sila, Oum El Bouaghi and Tebessa); the northern region (Annaba, Bordj Bouararidj, Constantine, Guelma, Jijel, Mila, Setif) and the southern region (Biskra and El Oued). To assess genotype distribution according to age and region, binomial logistic regression was performed. The model was:

Y = logit(p) = β0 + β1.age + β2.region + ε.

With p: probability to be infected by the Genotype 1.

The interaction of the covariate region on the association between age and genotype was tested by the likelihood ratio test comparing the models with and without the interaction term. The interaction was significant if p <0.05.

The statistical analysis was performed using the R version 2.15.1 statistical software