

# VIRAIID: WEB COMPUTER-AIDED SEROLOGICAL DIAGNOSIS OF ACUTE VIRAL HEPATITIS. AUTOMATIC CREATION OF A CLINICAL ALGORITHM

**R. Vignes, MD, PhD (1), SJ. Darmoni, MD, PhD (2), J. Lebbe, MD, PhD (1), T. Poynard, MD, PhD (3)**

(1) Classification, Evolution and Biosystematics, Pierre et Marie Curie (Paris VI) University, 4 Place Jussieu, 75005 Paris;

(2) Direction de l'Informatique et des Réseaux, Rouen University Hospital, 1 Rue de Germont, F76031 Rouen Cedex;

(3) Service de Hépato-Gastro-Entérologie, Groupe Hospitalier Pitié-Salpêtrière, 47 bld de l'hôpital, 75013 Paris.

## ABSTRACT

The serological diagnosis of acute viral hepatitis is complex for practitioners and students because of the number of markers and diagnoses, as well as its rapid evolution. The first step to generate the clinical algorithm is the description by the medical expert of 23 diagnoses of acute viral hepatitis using 12 serum markers. The medical expert estimates the relative frequency of each marker in each diagnosis, the prevalence of each diagnosis and the financial cost of each marker. From this description, the program MAKEY automatically generates a clinical algorithm. The program is a recursive procedure which step by step selects the best marker, i.e. first step IgM anti HAV. The clinical algorithm contains 96 nodes which gives the value, negative or positive of a marker, and 50 conclusions which contains at least one diagnosis, as well as a value which indicates whether the diagnosis is typical or not. This clinical algorithm has been available on the Web since November 1995 [address URL <http://www.chu-rouen.fr/dsii/html/makeyhep.html>]. In May 1997, 90 different machines accessed this clinical algorithm.

## Introduction

The serological diagnosis of acute viral hepatitis is complex for practitioners and students because of the number of markers and diagnoses, as well as its rapid evolution. In 1987, we developed an expert system on this subject for educational purposes using production rules (1). Currently, another knowledge representation now permits the creation of a clinical algorithm using the program MAKEY (2) developed by two individuals (JL & RV).

## 1. METHODS

### 1.1. Knowledge representation

#### 1.1.1. Knowledge-base

A knowledge-base includes a set of concept descriptions and background knowledge. The list of concept descriptions must be exhaustive, and the concepts must be mutually exclusive: an example belongs to only one concept, even if the example observation is not sufficient to identify this concept without uncertainty. The concept descriptions are given by experts. This information is considered certain, precise and true. In our application, all the descriptors are qualitative.

#### 1.1.2. Disjunction

When we describe biomedical concepts (i.e. acute viral hepatitis), the polymorphism is always present. For example, in Hepatitis A, the serum marker IgG anti EBNA is often present and rarely absent (i.e. Hepatitis A = [IgG anti EBNA = + OR IgG anti EBNA = -]). The disjunction indicate a

range of values and is neither due to the doubt the expert may have in describing concepts nor due to the imprecision of the descriptors. It reflects a variation within the taxon Hepatitis A.

### 1.1.3. Typicality

Typicality can be defined as the expert's subjective estimate of the relative likelihood of each alternative state. It is generally impossible to have more precise information than a pre-order among character states. This type of information is incorporating by giving a qualitative typicality value for each possible value in concept description. With such subjective estimates and without information about conditional likelihood, we cannot refer to probabilistic formalism.

## 1.2. MAKEY algorithm

An identification graph or a clinical alorithm is a rooted, directed acyclic inheritance graph. Each terminal leaf represents one or more values of the descriptor to be explained. In our medical application, the explained descriptor is the name of viral hepatitis diagnosis (so each leaf is a concept name). Each node represents an explanatory descriptor (serum markers). Each branch is one or more modalities (values) of the above descriptor. A path from the root to a terminal leaf describes a premise of an implication rule. For example: Ag HBs present and IgM anti HAV present and IgM anti HBc present => cohepatitis A B [typical].

To generate the clinical algorithm, a heuristic selects step by step the descriptor with the maximum discriminant power in the context. This selected descriptor makes a node with a branch for each possible modality. In each branch the process comes back to the selection of the best descriptor, except if a criteria to stop is matched. The path ends with a leaf which is the list of the compatible conclusions. The same conclusion can appear in more than one leaf in the graph, due to the disjunctions in concept descriptions. The quality of the graph depends on the method of computing the discriminant power of the descriptors. Different parameters can be modified to adapt the algorithm to a particular application and its domain theory: function to compare descriptions, function to aggregate typicality values, criteria to stop or to prune branches in the graph.

## 1.3. Statistics

Use of the Web page was measured with the Gestats program [<http://www.eit.com/software/getstats/getstats.html>] after excluding requests by our Hospital Information System.

## 2. RESULTS: Medical application

A knowledge-base describing serum profiles (descriptors) of acute viral hepatitis (concepts) was carried out by two individuals (SJD & TP). The knowledge-base concerns 23 diseases described with 12 serum markers. The medical experts estimate the relative frequency of each marker for each diagnosis, the prevalence of each diagnosis and the financial cost of each marker. For example, Hepatitis C = [IgM anti HAV = -] & [IgM anti HBc = -] & [Ag HBs = -] & [Sensitive markers of B virus (branched DNA, seric DNA, PCR) = -] & [Ac anti HCV = (+, Typical), (-, Atypical)] & [Sensitive markers of C virus (PCR) = +] & [Ag Delta = -] & [IgM anti Delta = -] & [IgG anti EBNA= (+, T), (-, AT)] & [IgM anti VCA = -] & [IgM anti CMV = -] & [Ac anti E = -].

The clinical algorithm contains 96 nodes which gives the value, negative or positive of a marker, and 50 conclusions which contains at least one diagnosis and a value which indicates whether the diagnosis is typical or not. For example, Hepatitis B is typical in one conclusion and very atypical in three.

Each version of the clinical algorithm has been reviewed by the experts. Since 1993, students and junior staff use the printed version of the algorithm in our department. Since November 1995, this clinical algorithm is available on the Web [address URL <http://www.chu-rouen.fr/dsii/html/makeyhep.html>]. In May 1997, users from 90 different machines accessed this

clinical algorithm originating from 15 different countries.

## **DISCUSSION**

Another team (3) has developed an expert system in the same field, viral hepatitis, using production rules but the domain covers only hepatitis A and B. Due to the rapid evolution of the number of markers we have scheduled a maintenance meeting each three months between the medical experts. They are easily updating the knowledge base and MAKEY automatically generates the Web version of the clinical algorithm which is always reviewed by the experts.

## **Conclusion**

Validated clinical algorithms on the Internet are a valuable tool for the health professional.

## **References**

1. DARMONI SJ., POYNARD T., MEMBRADO M., PARIGOT D., NAVEAU S., CHAPUT J.-C.: Elaboration d'un système expert pour l'aide à l'enseignement du diagnostic sérologique de l'hépatite virale aiguë. Press. Méd. 1987; **16**; 1903-1906.
2. LEBBE J, VIGNES R. Génération de graphes d'identification à partir de descriptions de concepts. In: Induction symbolique et numérique à partir de données, Kodratoff Y. et Diday E. Eds., CEPADUES, Toulouse , 1991, pp 193-239.
3. CHIZZALI-BONFADIN C, ADLASSNIG KP, KREIHSL M, HATVAN A, HORAK W: A WWW-accessible knowledge base for the interpretation of hepatitis serological tests. In: Mednet 96, European Congress of the Internet in Medicine, pp 72-73.