Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2015, 7(12):197-205

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

A new algorithm for deduction of time to detect Alzheimer's disease

A. Manimaran[1*], V. M. Chandrasekaran[1], Keshav Gupta[2], Megha Kanwar[2] and P. M. Karthick[3]

[1]School of Advance Sciences, VIT University, Vellore, India
[2]School of Information Technology, VIT University, Vellore, India
[3]Centre for Nanotechnology Research, VIT University, Vellore, India

---

ABSTRACT

Alzheimer's disease basically leads to memory loss. People suffering from this disease tend to forget things that may had happened in recent past. For example they may not recognize their family members and neighbors or tends to forget about having regular activities. Many researchers proposed various algorithms to identify the disease at the early stage in a short period of time. In this paper, we suggest a new algorithm to reduce the time period for the screening of AD by means of a two-stage hybrid intelligent approach based on multi-neuropsychological rating scales analysis and with the knowledge of genetic algorithm.

Keywords: Alzheimer's disease, Rough sets, Genetic algorithm, Chromosome and multi-neuropsychological rating scales analysis

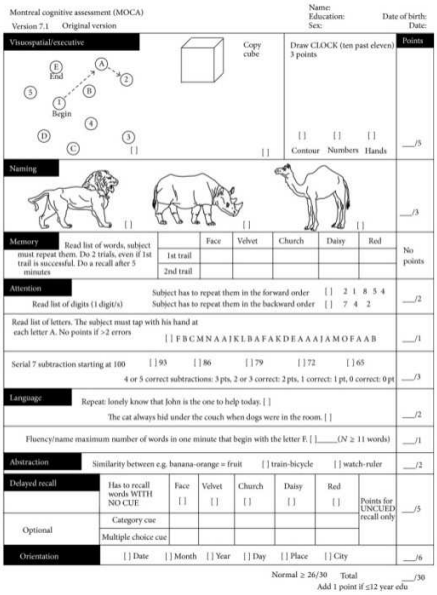
---

INTRODUCTION

Alzheimer's disease (AD) is a degenerative senile dementia characterized by memory loss and cognitive functions disorders, and it is also one of the main types of senile dementia. As AD has a slow onset and no highly specific diagnostic indicators at the early stage of the disease, it is particularly challenging for primary clinicians to identify transition points (from the asymptomatic phase to the symptomatic predementia phase to dementia onset) for individual patients. It is, nevertheless, important to identify these transition points between different stages, because studies have proved that targeted therapies may help slow down the progress of the disease and improve quality of life for patients and their families.[1,2,3,4,5,6]

EXPERIMENTAL SECTION

Due to the lack of advanced medical facilities (advanced imaging and cerebrospinal fluid measures), the screening of AD usually depends on the use of neuropsychological rating scales in primary clinics. Various neuropsychological rating scales, which are considered as a reliable and valid standardized testing tool, have been designed for cognitive abilities screening, and many of them have yielded good results as decision-making tools, such as minimental state examination (MMSE) , clinical dementia rating (CDR) , Montreal Cognitive Assessment (MoCA) , Geriatric Depression Scale (GDS) , and Activity of Daily Living Scale (ADL) . and Figure1 are two most commonly used rating scales (the MMSE and the MoCA) in clinical practice.

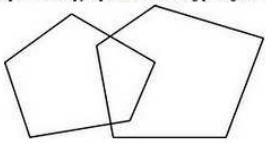


However, each neuropsychological rating scale has its emphasis and limitation. A previous study has shown some scales do not perform well in one or more cognitive domains. Multiple neuropsychological rating scales can cover more comprehensive cognitive domains. Therefore, multiple scales should be used together in order to get patients' comprehensive cognitive status, which can help doctors to make correct diagnosis. However, this will bring two challenges:

1. Neuropsychological testing requires highly trained assessors, while most primary clinicians are not qualified to conduct a full mental status examination or interpret a battery of scales' score; it is difficult for them to offer exact judgments about the examinee's cognitive state .
2. Neuropsychological testing is quiet time consuming; the elders cooperate well only for short periods with the limitation of vitality and cognition , so long-time testing will bring negative impact on the quality of neuropsychological testing. Thus, we can conclude that the screening of AD in primary clinics should be based on the criteria that can get maximum accuracy in a convenient way within limited time.

Table 1

Minimental state examination (MMSE).

Orientation	Year Month Day Date Time	__/5
	Country Town District Hospital Ward	__/5
Registration	Examiner names 3 objects (e.g. apple, table, and penny). Patient asked to repeat (1 point for each correct answer). THEN patient to learn the 3 names repeating until correct.	__/3
Attention and calculation	Subtract 7 from 100 and then repeat from result. Continue 5 times: 100 93 86 79 65 Alternative: spell "WORLD" backwards-"DLROW"	__/5
Recall	Ask for names of 3 objects learned earlier.	__/3
Language	Name a pencil and watch	__/2
	Repeat "No fits, ands, or buts"	__/1
	Give a 3-stage command. Score 1 for each stage. E.g., "Place index finger of right hand on your nose and then on your left ear"	__/3
	Ask patient to read and obey a written command on a piece of paper stating "Close your eyes"	__/1
	Ask patient to write a sentence. Score if it is sensible and has a subject and a verb	__/1
Copying	Ask the patient to copy a pair of intersecting pentagons: 	__/1
	Total	__/30

Solution of above mentioned problem

To solve the above-mentioned challenges, identifying the items with the best ability to distinguishing AD (called critical items for short) from a battery of commonly used rating scales may help improve the efficiency of cognitive abilities screening. Then, a well-performance decision-making model, while the previously selected items can be taken as its input, may help primary clinicians improve diagnostic accuracy in routine clinical practice. So in this paper, we suggest dealing with the screening of AD by means of a two-stage hybrid intelligent approach based on multi-neuropsychological rating scales analysis: in Stage 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Stage 2, use a Bayesian network to develop a diagnosis assisting model of AD based on the selected items. This hybrid intelligent technique takes the advantage of attributes reduction of rough set theory requiring no prior knowledge and the uncertain reasoning ability of Bayesian network to build a relatively convenient and accurate decision-making model for primary clinicians.

Theoretical knowledge required

We introduce rough set theory first and then discuss genetic algorithm and Bayesian network, so as to set up a necessary context for describing our approach.

1. Rough Sets

i. Basic Idea behind this approach

Rough set theory forms the very basis of critical decision making in terms of mathematical approach [13]. Decision tree analysis and other classical approaches can be linked back to this concept. It is based on bound value tables. Lower bound values means that the object's certainty of belonging to a particular target class is sure [14]. Upper bound values denote those points in the whole set which cannot be classified as belonging to the target set with a sure certainty [15]. Boundary- line elements decide whether an object can be classified as a part of the set or its complement or none. Basically thus we see that rough sets measures the uncertainty in relations not precisely by the membership values of properties but with the boundary value limits of a function [3]. If the boundary region is a null region then the set is said to be crisp set otherwise any not null values in boundary defines that the set is a rough set. Formal definitions of approximations and the boundary region are as follows:

- *R-lower approximation* of *X*

$$R_*(x) = \bigcup_{x \in U} \{R(x) : R(x) \subseteq X\}$$

- *R-upper approximation* of *X*

$$R^*(x) = \bigcup_{x \in U} \{R(x) : R(x) \cap X \neq \varnothing\}$$

- *R-boundary region* of *X*

$$RN_R(X) = R^*(X) - R_*(X)$$

Rough sets memberships functions can be defined as:

$$\mu_X^B(x) = \frac{|X \cap B(x)|}{|B(x)|}, \text{ where}$$

$$\mu_X^B(x) \in [0,1].$$

The values related to the member function $\mu_X(x)$ can be interpreted as a conditional assumption with some certainty for which *x* belongs to *X*

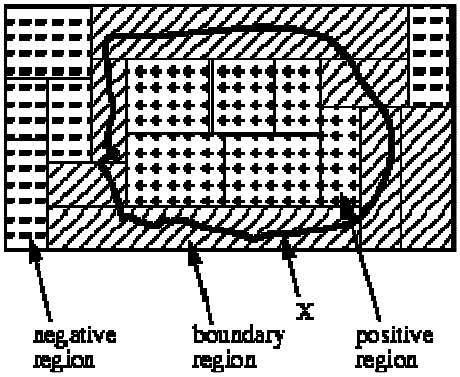


Fig. 1 – Rough set representation

ii. Applications of the Rough Set theory

Rough set have wide applications in data mining and knowledge discovery methods. Rule induction and feature selection problems have been mainly deal using the concepts of rough set attribute reductions (semantics-preserving reduction patterns). The kind of perfection in results and computation reduction provided by using the rough set approach is really feasible, as mining data from large repositories can be very much time and resource consuming. One of the most useful features of rough sets is the ability to reduce or remove the unwanted attributes [4]. So by using this method we can apply the probabilistic neural network to classify the given image segments or image feature datasets. This framework of task is efficient by computational measures and can be trusted for accuracy and reliability. Neural Networks can be trained and tested for the same purpose.

RESULTS AND DISCUSSION

Genetic algorithm

The genetic algorithm (GA) is an optimized algorithm based on the Darwinian principle of natural selection. It can be used with other data mining techniques for optimization and performance amelioration. The genetic algorithm process starts with the randomly generated and encoded initial population, which includes several hundreds or thousands of potential solutions to the problem. Each encoded individual in the population is called chromosome and each bit in the chromosome is called gene and has a value [7]. The next step is called genetic operators. The most widespread genetic operators include selection, crossover, and mutation. Each chromosome in the population is evaluated by user-defined fitness function. The higher a chromosome's fitness value is, the more likely it is to produce offspring. In this way the overall fitness of the population is guaranteed to increase and those with weak fitness will be eliminated gradually. Crossover forms new chromosomes for the population by exchanging a fixed part between two chromosomes. The chromosomes most often used for crossover are those destined to be eliminated from the population. Mutation can be applied by randomly flipping bits (or attribute values) within a single chromosome to avoid the local optima. New offspring is reevaluated by fitness function to search the solution. The whole process is repeated until reaching the pre-specified number of generations or the desired level of fitness.

Bayesian network

Bayesian network is an acyclic directed graph for representing probabilistic relationships among a set of random variables. Trained Bayesian networks can be used for classification. There are two key elements of a Bayesian network:

1. A directed acyclic graph (DAG) encoding the dependence relationships among a set of variables and
2. A probability table associating each node to its immediate parent nodes. Each node in the directed acyclic graph represents actual attributes given in the data. Each arc represents a probabilistic dependence. If an arc is drawn from a node X to a node Y, then X is a parent of Y, and Y is a descendant of X. Bayesian network has one conditional probability table (CPT) for each node. The CPT for a node Yspecifies the conditional distribution $P(Y \mid \text{Parents}(Y))$, where Parents($Y$) are the parents of Y. A node within the network can be selected as an "output" node, representing a class label attribute. There may be more than one output node. Given a set of variables, the network can be used to compute the probabilities of the presence of various classes, rather than return a single class label. There have been some works with applications using Bayesian network in diagnosis of AD. [8,9,10,11,12]

Overview
In this section, we present the formation process of AD diagnosis assisting model with the proposed hybrid intelligent method, which consists of two steps: in Step 1, use a genetic algorithm-rough sets (GA-RS) model to identify critical items, and in Step 2, use a Bayesian network to build a diagnosis assisting model of AD based on selected items.

Attributes Reduction Based on Genetic Algorithm and Rough Set Theory
GA-RS is used to identify critical items from a battery of rating scales. Each step of the algorithm is described as follows.

Chromosome representation
Because genetic algorithm cannot deal with data in solution space directly, we must represent them as binary strings of length *M* which is the number of the condition attributes by encoding. Binary encoding is simple and easy to operate. Each binary string is called a chromosome, in which "1" means that the corresponding attribute is selected and "0" means not. Attributes in *Core* should take "1", and remain the same in the whole process of evolution, since genetic search starts from the *Core*.

Fitness function
Fitness function is a user-defined function which is used to measure each chromosome's optimization calculation in the groups. The fitness value of each chromosome represents suitability for the environment. In this paper, we expect the "best" chromosome could have the minimal length and the strongest classification performance as the algorithm proceeds. So the fitness function is defined as follows:

$$F(x) = \beta f(x) + p(x) = \beta\left(1 - card(x)card(C)\right) + card\left(POSX(D)\right)card\left(POSC(D)\right)$$

where $card(x)$ is the number of "1" in chromosome, which means the number of condition attributes contained by chromosome; $card(C)$ is the length of chromosome, which is the total number of condition attributes; $f(x) = 1 - card(x)/card(C)$ indicates the chromosome *x* that is not included in the proportion of condition attributes. $p(x)$ indicates the distinction ability of attribute *x*.

Selection method
Select chromosomes based on their fitness values from the current population to produce offspring for the new population. Tournament selection is used, which means the higher the fitness value is, the higher probability of that chromosome is selected for reproduction. This step is repeated until the number of chromosomes selected is equal to the number of the population.

Crossover and mutation
One-point crossover method is used to reproduce with a probability of P_c. In mutation process, we first select a chromosome to be mutated with probability P_m and then replace a single gene of the chromosome from "1" to "0" or from "0" to "1" randomly.

Elitist Strategy
We take the elite strategy (7) to preserve the best individual of the fitness function value. Copy the individual of highest fitness value in the current generation to the next generation, unaltered.

The detail of the whole algorithm is as follows.
Input. Decision table $IS = \langle O, A, V, f\rangle$; O is a nonempty finite set of objects. A is a nonempty finite set of attributes: $A = C \cup D$, C is the set of condition attributes, and D is the set of decision attributes. $V = \bigcup_{a \in A} V_a$ where V_a is the set of values of attribute $a \in A$. $f : O \times A \to V$ is an information function so that, for any $a \in A$ and $x \in O$ and $f(x, a) \in V_a$.

Output. There is an attributes reduction *R* of decision table.

Step 1:
Calculate the dependency $\gamma_c(D)$ between decision attributes set D and condition attributes set C by formula (3).

Step 2:
Let $Core(C) = \phi$, to get rid of each attribute $c \in C$ one by one, if $\gamma_{C-c} \neq \gamma_C$, $Core(C) = Core(C) \cup \{c\}$ which means the core is *Core(C)*; if $\gamma_{Core}(D) = \gamma_C(D)$, then the core is minimal attributes reduction and if not, go to Step3

Step 3:
Generate m binary strings with length n randomly, which can be seen as the initial population. n is the number of the condition attributes. "1" means that the corresponding attribute is present, and "0" indicates not. For attributes in core, corresponding position is "1" and for others, corresponding position is "1" or "0" randomly.

Step 4:
Calculate the fitness value for each individual by formula (6) and select individuals by tournament selection.

Step 5:
Perform crossover operation according to the crossover probability P_c, using single-point crossover mode.

Step 6:
Perform mutation operation according to the mutation probability P_m. We basically bit mutation strategy while the corresponding bit of attributes in the *Core* does not change.

Step 7:
Select the individuals with the best fitness values to be offspring of the current generation. This strategy is to guarantee the best chromosome could carry over to the next generation.

Step 8:
Repeat the genetic operation until either one of the following conditions is satisfied: (1) the maximum number of generations is achieved or (2) the fitness value of the best individual for the present generation no longer changes during several successive generations.

Step 9:
Convert the best individual to condition attribute and get the final result.

Parametric settings of genetic algorithm are as follows: population scale N = 1000, crossover ratio P_c = 0.5, mutation ratio P_m = 0.03, and the largest number of iterations is 500

The fitness function employed in this paper controls the chromosomes that evolve in the direction of the minimum reduction while keeping the classification performance: the higher the card(x) is, the smaller the $f(x)$ is; the larger $p(x)$, the more dependence between the condition attribute C and decision attribute D. This algorithm ensures the two requirements, so the result is the optimal solution of the problem.

In our approach, attributes reduction mentioned above is not the final goal but an intermediate process and core technology of AD diagnosis assisting for clinician in primary clinic. An uncertainty inference model for AD should be built after attributes reduction.

Bayesian Network Model for AD Diagnosis
Based on the above step, we attempt to construct the structural model for AD diagnosis. These selected items can be represented as input variables of the model. Since there is strong diagnostic uncertainty earlier in the disease process, an uncertainty inference model must be built. A popular modeling tool for complex uncertain domains is a Bayesian network.

Data collection and experiments

The experimental data set is composed of 500 consecutive historical cases collected by the neurology department of a certain top hospital in China from 2009 to 2014. Each case is a series of scale scores belonging to one subject, and each subject has only one case. All neuropsychological tests were conducted by trained neuropsychologists and administered on the same day. The mean age of subjects is 74.4 (range, 51–92); 59.5 of the subjects' percent were female. These 500 historical cases have the following characteristics.

1. All these 11 neuropsychological rating scales are selected from a large number of scales by leading experts in neurology, including the MMSE, the MoCA, the CDR, the GDS, the ADL, the Word-List Learning, the figure copying, the new word discriminating, the trail making test, the similarity, and the perception. All these neuropsychological rating scales are commonly used instruments for screening cognitive or noncognitive impairment in the clinical diagnosis.

2. Each scale consists of a series of items. In total, there are 101 testing items in these scales. Some items are straightforward Q & A pattern, for instance, "What is the date?" Some others need the subject to do some actions, "Please read this and do what it says. (Show subjects the following words on the stimulus form: Close your eyes.)" Each of the tests scores points if it is answered correctly.

To ensure the correctness of diagnosis of each case, an expert panel group composed of three neuropsychologists was set up, and the diagnosis of each case was determined by the panel. The diagnosis of experts not only depended on an objective neuropsychological testing, but also on the history-taking from the patient and a knowledgeable informant. Their diagnosis was regarded as the gold standard. In the current study, the diagnosis of cases could be divided into three types: patients with AD, patients meeting criteria for mild cognitive impairment [16] (called MCI for short, which is regarded as the predementia stage of AD), and the elderly subjects with normal cognition, in which, the number of each type is 33.5%, 37.7%, and 28.8%, respectively.

Parts of cases are given in Table 3. In the table, each column is one testing item of scales, for instance, Time Orientation, Place Orientation and Repetition belong to the MMSE while Visuospatial Skills belongs to the MoCA. They are regarded as the condition attributes. The last column, Result, is the decision attribute (the diagnosis of each patient).

Verification Of results

To verify the feasibility and validity of the proposed approach, the performance of proposed approach can be measured by the following evaluations: (1) reduction ratio on testing duration and reduction ratio on quantity of items; (2) comparison with multiple classifiers; (3) comparison of classification accuracy before and after reduction; (4) the performance of classification compared with two existing cognitive screening scales.

We applied two evaluation methods to prove the reliability of our experimental results, one was 10-fold cross-validation, and the other was 0.632 bootstrap. Obtain computing results by averaging after executing 10 times. Recall rate, precision rate, and accuracy were selected as the performance evaluation metrics.

Reduction results

After attributes reduction, 10 items were selected finally.

Some items represent one test, such as Figure Copy, Figure Short Memory, Figure Delay Memory, Word Delay Recall, and Reading Comprehension, while others are a set of several tests, which cannot be separated (because these tests significantly correlate with one another and must be performed together), such as Visuo-spatial Execution, IADL, Naming, Attention, and Word AVG.

Reduction Ratio

We used reduction ratio including the reduction ratio of testing duration and the reduction ratio of quantity of items as measurable metrics. Assume that the number of condition attributes before and after reduction is m and n, respectively. The reduction ratio r is defined:

$$r = (m - n)\, m \times 100\%$$

Before reduction, the number of items is 34, while only 10 items left after reduction using the proposed method, so we can conclude that the reduction ratio is 70.59%. The experimental results indicate that using GA-RS to select subset can reduce items dramatically.

Similarity, the reduction ratio of testing duration can also be calculated using formula . In clinical practice, the duration of finishing these scales varies a lot, which depends on the subject's state of cognitive impairment. According to the performance time for the MMSE and the MoCA is 13.4 minutes and 14.8 minutes on average, respectively. Based on the past experience, a skilled clinician administers the scale for more than one hour to complete all the 11 scales mentioned above. By using the proposed model, clinicians do not have to finish all the scales but only need to complete the selected testing items. Hence the test duration is reduced greatly and ranges from 12 to 15 minutes with a mean time of 13.5 minutes and a standard deviation of 2.3 minutes.

Classification performance before and after reduction

In order to evaluate the validity of attributes reduction, we used Bayesian network algorithm to compute the classification performance before and after attributes reduction, respectively, and to check whether or not the classification performance had changed.

Each subject had been given the probability of each classification. The highest probability was regarded as the diagnosis of the model.

From the variance of recall rate and precision rate after attributes reduction we found that the recall rate and precision rate of each group decreased a little, but less than 3.05%. We analyzed the result data using Wilcoxon Signed-Rank Test. The calculated *P* value was 0.853 and larger than 0.05, so the null hypothesis was true, which means that there was no significant statistical difference between these two methods. In conclusion, the comparative experimental results indicated that the proposed method could find the shortest or minimal reducts while keeping high-quality classification performance.

Conclusion and Future Work

The increasing aging population has led to a high increase in the prevalence of AD. Due to the fact that targeted care and therapies may slow down the progression of disease, the identification of different stages of AD is very important. In this paper, we proposed a computer-aided diagnosis method for AD based on analyzing the practical scores of rating scales. We especially identified the most discriminative items based on rough set theory and genetic algorithm. The selected items cover multiple cognitive domains and can be administered generally within 15 minutes. So it is user-friendly and is quickly administered, it may be appropriate use in primary clinics where assessment time is often limited. By comparing the classification performance, the result showed that the approach can effectively reduce the representation space of the attributes whilst hardly decreasing classification precision. The data also indicated that it has satisfactory reliability for both MCI and AD comparing with other existed cognitive screening scales.

Without doubt, opportunities for future research are abundant. First, we plan to further evaluate the built model with a perspective study in a real clinical setting. Second, more rating scales for specific dementias are going to be involved in the training set data and more comprehensive model for senile dementia will be built in the future work. Based on above work, a “three-level medical service network” for AD is going to be built in the near future and different computer-aided diagnosis tools for each level hospital will be developed; for example, the simple cognitive screening tool helps clinicians in primary clinics to judge whether patients suffer from cognitive impairment; the advanced cognitive assessment tool helps clinicians in second class hospitals to estimate the severity of cognitive impairment; the comprehensive assisted diagnosis tool is designed for clinicians in top hospitals to differentiate the types of dementia. The setup of such network will improve diagnosis accuracy of AD greatly and reduce the burden on public health care resource.

REFERENCES

[1] Porto, Cláudia S., et al., Brazilian version of the Mattis dementia rating scale: diagnosis of mild dementia in Alzheimer's disease, Arquivos de neuro - psiquiatria, **2003**, 61(2B), 339-345.
[2] Sperling, Reisa A., et al., *Alzheimer's & dementia*, **2011**, 7(3), 280-292.
[3] M C Khann, Guy M., et al., *Alzheimer's & Dementia*,**2011**, 7(3), 263-269.

[4] Rhodes - Kropf, Jennifer, Palliative care for patients with Alzheimer's dementia: advance care planning across transition points, Choices in Palliative Care, Springer US, **2007**, 144-156.
[5] Ahmad, Fadzil, et al., *Journal of medical systems*, **2013**, 37(2), 1-8.
[6] Daliri, Mohammad Reza, *Journal of medical systems*, **2012**, 36(2), 1001-1005.
[7] Goldberg D. E, Genetic Algorithms in Search, Optimization, and Machine Learning, Menlo Park, Calif, USA, Addison-Wesley, **1989**.
[8] Shimodaira, Hisashi, A new genetic algorithm using large mutation rates and population-elitist selection (GALME), Tools with Artificial Intelligence, 1996., Proceedings Eighth IEEE International Conference on IEEE, **1996**.
[9] Pearl, Judea, Probabilistic reasoning in intelligent systems: Networks of plausible reasoning, **1988**.
[10] Hall, Charles B., et al., *Computational Statistics & Data Analysis*, **2003**, 42(1), 91-109.
[11] Béalisle, Patrick, et al., *Canadian Journal of Statistics*, **2002**, 30(1), 37-54.
[12] Prince, Martin J. " *American journal of epidemiology*, **1996**, 143(3), 301-308.
[13] Peters, G. and Lampart, M., A partitive rough clustering algorithm, Proceedings of the Fifth International Conference on Rough Sets and Current Trends in Computing (RSCTC'06), Lecture Notes in Artificial Intelligence, LNAI-4259, (Kobe, Japan,**2006**), Springer,657-666.
[14] Deogun, J.S., Raghavan, V.V. and Sever, H. Exploiting upper approximations in the rough set methodology. In U.M. Fayyad and R. Uthurusamy(Eds.), Proceedings of First International Conference on Knowledge Discovery and Data Mining, AAAI Press, (Canada, **1995**), 69-74.
[15] Pawlak, Z. and Skowron, A. *Information Sciences*, **2007**, 177(1), 3-27.
[16] Gauthier, Serge, et al. *The Lancet* 367.9518, **2006**, 1262-1270.