# Detecting Abdominal Aorta Aneurysm using Bio-Computing Technology

A Dinesh Kumar
PG Scholar,
Department of Computer Science and Engineering,
Karpagam University,
Coimbatore, Tamil Nadu.

Hanah Ayisha V Hyder Ali
PG Scholar,
Department of Computer Science and Engineering,
Karpagam University,
Coimbatore, Tamil Nadu.

S Shahul Hammed
PG Scholar,
Department of Computer Science and Engineering,
Karpagam University,
Coimbatore, Tamil Nadu.

## ABSTRACT
Abdominal aortic aneurysm (AAA) is a localized dilatation of the abdominal aorta .It occurs when there is a increase in the normal diameter of the blood vessels by more than 50 percent. Approximately 90 percent of abdominal aortic aneurysms occur infrarenally , but they can also occur pararenally or suprarenally. This is because of some catastrophic outcome. Due to this, the blood flow is exaggerated so the blood hemodynamic interaction forces are affected. Therefore this will tends to wall rupture. To identify the AAA .it is important to identify the blood flow interaction and the wall shear stress. The blood and wall interaction is the wall shear stress. Computational fluid dynamics (CFD) is used to get the results for the mechanical conditions within the blood vessels with and without Aneurysms.CFD contains vast computations with Navier Stroke Equations so this will be very time consuming. So to make these CFD computations very efficient, Data mining (DM) techniques are to be used. And also DM techniques will be a best method to predict the shear stress at the AAA. This will estimate the wall shear stress. There is in need of thousands of CFD runs in a single computer for creating machine learning data so grid computing is used.

## Keywords
Computational fluid dynamics (CFD), data mining (DM), grid computing, hemodynamic parameters, predictive modeling;

## 1. INTRODUCTION
Cardiovascular disease is a class of diseases that involve the heart or blood vessels. Cardiovascular disease refers to any disease that affects the cardiovascular system. Cardiovascular diseases remain the biggest cause of deaths worldwide. Age is an important risk factor in developing cardiovascular diseases. It is estimated that 87 percent of people who die of coronary heart disease are 60 and older. Among these diseases the stenosis process is the most dangerous one. This stenonsis process will lead to stroke and aneurysm and now a days this is commonly present in patients. In this paper, we are giving importance to Abdominal Aorta Aneurysm (AAA) development that may cause rupture and fatal outcome.

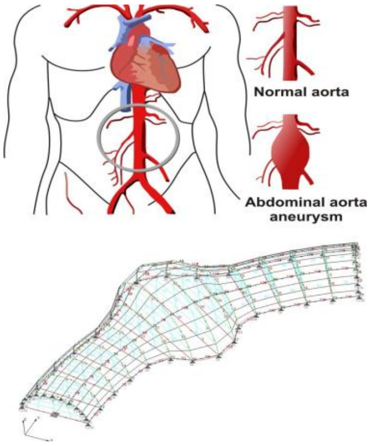


**Fig: 1. Difference between normal aorta and AAA (Taken from Reference)**

An aneurysm is an abnormal widening or ballooning of a portion of an artery due to weakness in the wall of the blood vessel. It is a localized, blood-filled balloon-like bulge in the wall of a blood vessel. Aneurysms can commonly occur in arteries at the base of the brain and an aortic aneurysm occurs in the main artery carrying blood from the left ventricle of the heart. When the size of an aneurysm increases, there is a significant risk of rupture, resulting in severe hemorrhage, other complications or death. Aneurysms can be hereditary or caused by disease, both of which cause the wall of the blood vessel to weaken. Aneurysms may be classified by type, location, and the affected vessel. Other factors may also influence the pathology and diagnosis of aneurysms. A consensus definition of an aneurysm was established in 1991 by the Society of Vascular Surgery and the International Society for Cardiovascular Surgery as a permanent localized dilatation of an artery having at least 50% increase in diameter compared with the expected normal diameter of the artery, or of the diameter of the segment proximal to the dilatation.

To predict the AAA many patient specific studies have been done, from this it's concluded that maximum stress within the vessel wall was more appropriate criterion than maximum diameter and also the shear stress on aneurysm wall is found. Shear stress is a frictional force produced by blood flow, affects biology and structure of the wall. There are many methods to predict this. The following paragraphs describes about the methods along with its disadvantages.

First, the Shear stress can be obtained by the Computational Fluid dynamics (CFD). This approach has been employed to study wall-shear stress distribution on idealized models of blood vessels. CFD uses numerical methods and algorithms to solve and analyze problems that involve fluid flows. CFD involves the Navier Stroke Equations and this CFD involves many computations. So this method is very time consuming. Since there are uncertainties in the patients it is difficult to characterize geometric variability using a small number of recorded parameters.

Second, Statistical assessment to identify the relationship between flow patterns and geometric attributes. The main concept is to construct the probabilistic models for the input parameters uncertainties that give a reliable output of interest very quickly, without classical CFD calculations. An example of this idea is reported by Kolachalama who used Bayesian–Gaussian process emulator to generate a relationship between geometric parameters and maximal wall shear stress (MWSS), and to identify geometries having maximum and minimum of the wall shear stress (WSS).

Third, Statistical Analysis as the Monte Carlo simulation technique. In this computer runs for the generated random input values and the resultant data is post processed to estimate the output statistics. CPU requirement for a large sample size, this approach becomes computationally prohibitive, particularly when high-fidelity models are used.

Fourth, it uses data Mining (DM) system that can be used to avoid CFD simulations. The DM model uses Discovery Bus software where the input is the aneurysm shape, determined by several independent geometric parameters, while the output parameters are the MWSS over the aneurysm for peak systolic flow, and MWSS over full heart cycle. Also, the shear stress averaged over the aneurysm region and over the heart cycle can show which aneurysms are likely to rupture so that the risk can be properly quantified.

The aim of this paper is to introduce Data Mining Algorithm for detecting AAA for multiple CFD runs. Regression model used for this purpose. A back propagation and Ant Colony algorithm for splitting the training and test set.

## 2. MATERIALS AND METHODS

The basic task in our approach was first to feed the DM system with necessary data acquired from the CFD simulations, and then to perform machine learning. We ran a number of various aneurysm models in order to "teach" the DM system for producing the most accurate predictions. As in any CFD modeling task, there are a number of issues to be considered: geometry, boundary conditions, initial conditions, mesh generation, etc. We are using the CFD finite element (FE) models, which differ only in geometry and rely on random values of ten parameters.

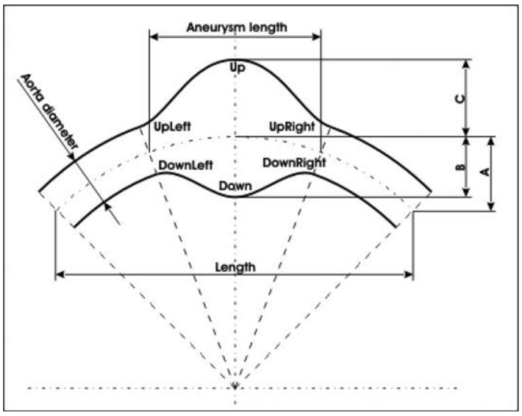


**Fig. 2.1. Parametric model of the aneurysm geometry. (Taken from Reference)**

**Table I. Ranges of AAA Geometrical Parameters**

| **Parameter** | **Range** |
|---|---|
| Aneurysm Length (cm) | 6 to 15 |
| A (cm) | 0 to 3 |
| B (cm) | 1.4 to 3 |
| C (cm) | 1.4 to 3 |
| Curvature parameters (UpLeft, ...) | 0.15 to 0.4 |

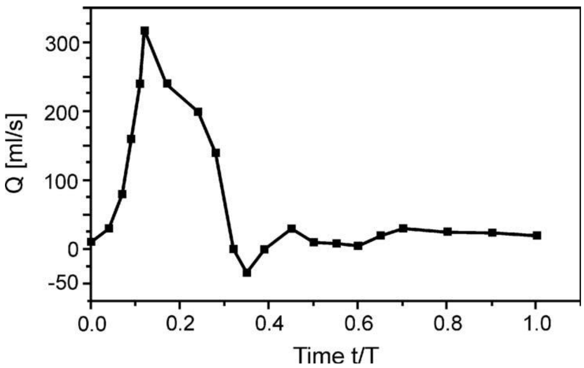


**Fig. 2.2. Typical inflow waveform at the artery entry. (Taken from Reference)**

### 2.1 CFD Modeel Of An Aneurysm

In this we have to consider the AAA only, among various kinds of artery aneurisms. The shape of the AAA is defined by two splines with half-circle extrusions between them, as shown in Fig. 2. We divide the geometric parameters into three groups.

1) *aneurysm* length, *A, B, C*—variable parameters;

2) *upleft, up, upright, downleft, down, downright*—quantifying the curvature in the Bezier description fashion, variable parameters;

3) *length, aorta* diameter—parameters considered are constant, and these are taken from the literature for typical AAA, thus this will give ten independent parameters having a ranges as shown in Table I.

The next step in the prebuilding models for the CFD analysis. This includes the specification of boundary conditions. The identical boundary conditions are prescribed for all combinations of the variable input parameter values. We used an equivalent length at the aneurysm outlet to model the resistance to the blood flow. Other relevant quantities with fixed values are blood density $\rho$ = 1.05 g/cm3, kinematic Viscosity $\nu$ = 0.035 cm2 /s, length = 24 cm, and aorta diameter $D$ = 2 cm.

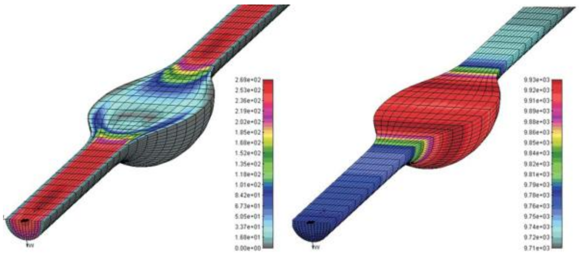

**Fig. 2.3. Velocity and pressure field for symmetric AAA on the straight artery (Taken from Reference).**

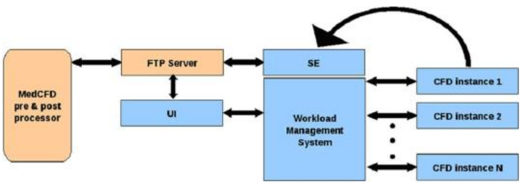


**Fig. 2.4. Grid computing scheme. CFD module generates finite element models for each CFD run. All data are then transferred to Storage Element (SE) from User Interface (UI). Then, each CFD instance reaches destination worker node anywhere in the grid, determined by the Workload Management System (WMS). This procedure significantly saves time, enabling a large number of CFD runs to be executed concurrently. (Taken from Reference)**

The fundamental equations for flow of a viscous incompressible fluid (such as blood) are the Navier–Stokes equations [12]. These Navier–Stokes equations are used to compute the CFD. For the 3-D model of blood flow, eight-node finite element is used *with* linear interpolation of velocities from all nodes, while the pressure, taken to be constant over the element, is eliminated by a penalty parameter [13]. The incremental iterative form of the FE equilibrium equations is

$$\left(\frac{1}{\Delta t}\mathbf{M}_{\mathbf{v}} + {}^{t+\Delta t}\mathbf{K}_{\mathbf{vv}}^{(i-1)} + {}^{t+\Delta t}\mathbf{K}_{\mu\mathbf{v}}^{(i-1)} + {}^{t+\Delta t}\hat{\mathbf{K}}_{\mu\mathbf{v}}^{(i-1)} \right.$$
$$\left. + {}^{t+\Delta t}\mathbf{J}_{\mathbf{vv}}^{(i-1)} + \mathbf{K}_{\lambda\mathbf{v}}\right)\Delta\mathbf{v}^{(i)} = {}^{t+\Delta t}\hat{\mathbf{F}}_{\mathbf{v}}^{(i-1)}$$

Where the matrices and vectors are defined in a standard FE manner (see the references). As an illustration of the CFD solution, the velocity field (left panel) and pressure distribution (right panel) for the peak systole $t/T$=0.16 of an AAA with $D/d$=2/75 ($D$ is *aneurism* diameter), $d$ = 12.7 mm, are shown in Fig. 4. After initializing input parameters (see Fig. 5) and creating appropriate FE mesh, a specialized CFD FE analysis module performs hemodynamic simulation and lists result values of various hemodynamic quantities at specified mesh points. Among these quantities, the DM system only considers the wall shear stress values. A total number of models (and therefore FE meshes) constructed using random parameter values and automatic mesh generator was 6000. Since a huge number of CFD finite element analyses had to be executed (for 6000 different geometries based on ten variable parameters) it will take more time (takes even days). So Grid Computing is used to reduce the computation time. The infrastructure used in this work has been developed under EGEE project with *gLite* middleware [16]. Since each CFD run takes around 20 min on a typical personal computer, simple computational time estimation gives around 80 days to run on a single CPU. Equivalent run on a grid platform took 5 h only, while the infrastructure utilization peak during that run was around 600 CPUs at a time.

## 2.2 DM Approach

After obtaining data from a number of FE runs machine learning process can be done. For this purpose Back Propagation (BP) and Ant Colony Optimization (ACO) Algorithm is used. The back propagation algorithm looks for the minimum of the error function in weight space using the method of gradient descent. The combination of weights which minimizes the error function is considered to be a solution of the learning problem. Since this method requires computation of the gradient of the error function at each iteration step, we must guarantee the continuity and differentiability of the error function. We have used the regression model for finding $\tau$ systolic (maximum shear stress for the systolic phase) and $\tau$ max (maximum shear stress value for the entire cycle domain). Fig. 6 shows these steps

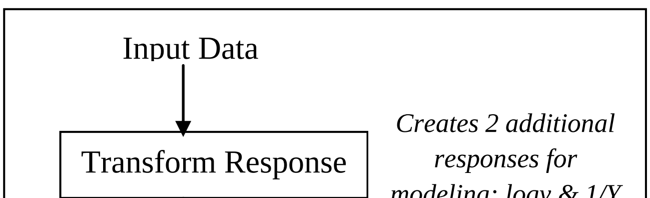

**Fig. 2.5. Overview of the main stages in building and testing**

The first step is to create additional transformations to the dependent variables, in our case $\tau$ systolic and $\tau$ max. Sometimes, it is easier to create a regression if the dependent variable $Y$ is first transformed as $\log(Y)$ or $1/Y$. Therefore, in addition to building models for $Y$, the Bus automatically builds models for $\log(Y)$ and $1/Y$. The second step is to sort the dependent variable in the ascending order and take every tenth out for testing. This regression model will take 10% for testing and creates models on 90% of the dataset. The third step is to perform variable selection. As a result, it produces up to five different subsets of features, including a solution where all variables are selected. The fourth step is to build regression models and to cross validate them. All regression models are at the end cross-validated (tenfold) and the following statistics is calculated.

1) Tenfold cross-validated RMSE—root mean square error after tenfold cross validation.

2) Tenfold cross-validated $Q2$—coefficient of determination after tenfold cross validation.

3) Tenfold cross-validated RelSE—relative squared error after tenfold cross validation.

4) Size of the training dataset ($N$tr )

The last step is to apply all regression models to the test set and following statistics are calculated.

1) RMSE—root means square error.

2) $R2$—coefficient of determination.

3) RelSE—relative squared error.

4) Size of the test dataset ($N$test).

From these values we have to produce a performance metrics in which the error rate should be a minimum. RMSE can have a error of 0.001 and $R^2$ as 0.9.

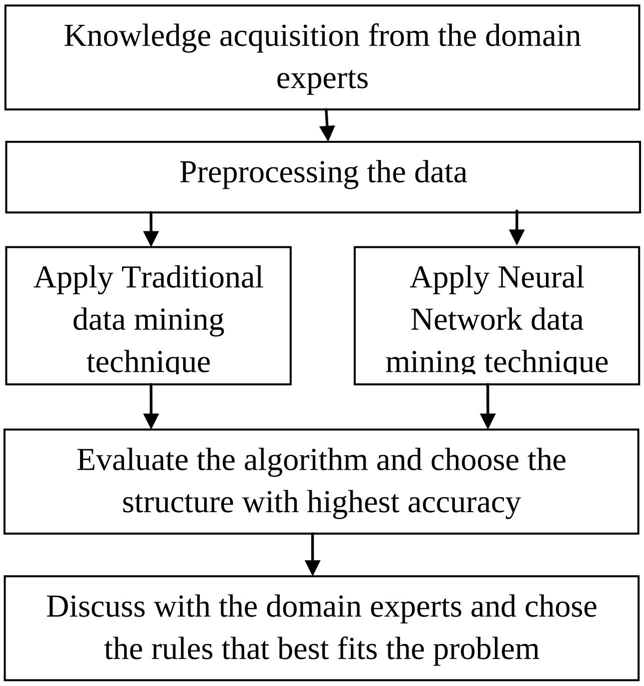


**Fig 2.6. Algorithm 1: Framework for Back propagation Algorithm**

```
Training set = all training cases;
    WHILE (No. of cases in the
    Training set >
    max_uncovered_cases)
        i=0;
        REPEAT
        i=i+1;
    Anti incrementally constructs a
    classification rule;
    Prune the just constructed rule;
    Update the pheromone of the trail
    followed by Anti;
    UNTIL (i ≥ No_of_Ants) or (Anti
    constructed the same rule as the
    previous No_Rules_Converg-1
    Ants)
        Select the best rule among
        all constructed rules;
        Remove the cases correctly
        covered by the selected
        rule from the training set;
END WHILE
```

**Algorithm 2: Ant Colony Optimization (ACO) Algorithm.**

## 3. RESULTS AND DISCUSSION

From 85 models for property $\tau$ systolic and the best model was found for $\tau$ −1 systolic. It is a feed forward neural network model that uses all ten input variables. From 82 models for property $\tau$max and the best model was found for $\tau$−1max. It is a feed forward neural network model that uses all ten input variables.

The models show a very low error statistics on the test and training sets. In order to better quantify the DM method, we also analyzed an AAA model based on the patient-specific clinical data (see Fig. 8). Relevant values of $\tau$ systolic and $\tau$max were obtained using the following three approaches.

1) ***CFD model based on patient's AAA cross-sectional data.*** The cross-sectional AAA images have been converted into FE mesh and analyzed using velocity boundary conditions measured with a standard Doppler device. The final refined FE mesh used in CFD analysis consisted of 53 335 eight node Brick elements.

2) ***FE model based on ten proposed geometric parameters.*** The real AAA shape was converted empirically into the CFD model fully described by ten geometric parameters.

This FE mesh consisted of 6273 eight node brick elements. The values of geometric parameters are *A*: 1.9, *B*: 1.5, *C*: 2.3, *aneurysm length*: 10.1, *upleft*:0.19, *up*: 0.37, *upright*: 0.31, *downleft:* 0.3, *down*: 0.22, and *downright*: 0.29.

***3) Model based on DM approach.*** The best model among 85 models produced by BP algorithm (feed forward neural network that uses all ten input variables) has been chosen in order to obtain the best approximation for $\tau$ systolic and $\tau$ max. As expected, the most realistic model from the first approach consumes the most computing time. The values of $\tau$max and $\tau$ systolic were taken as reference values for the parametric and DM models.

It can be observed that all relative errors are below 10%, with significant computing time reduction in favor of the DM model.

Some methods and results are taken from the references as a review.

## 4. CONCLUSION

This paper will provide a better performance for detecting the aneurysm. In normal method the computation takes more than 80 days to run in a single computer. So by introducing Grid Computing it has reduced to 5 hours. The CFD calculations were performed using the FE method and the shear stress was evaluated. A grid platform was employed to make the process of machine learning faster. Several building techniques for the regression problem were used. From these results, it can be stated that a new direction for reducing computational time for patient-specific AAA modeling is suggested. Also, this approach of coupling the computer modeling and DM methods (Back propagation (BP) and Ant Colony Optimization Algorithm) can further facilitate the development of predictive diagnostic system for clinical practice.

For future enhancement efficient data mining algorithm can be used to provide improved accuracy and performance.

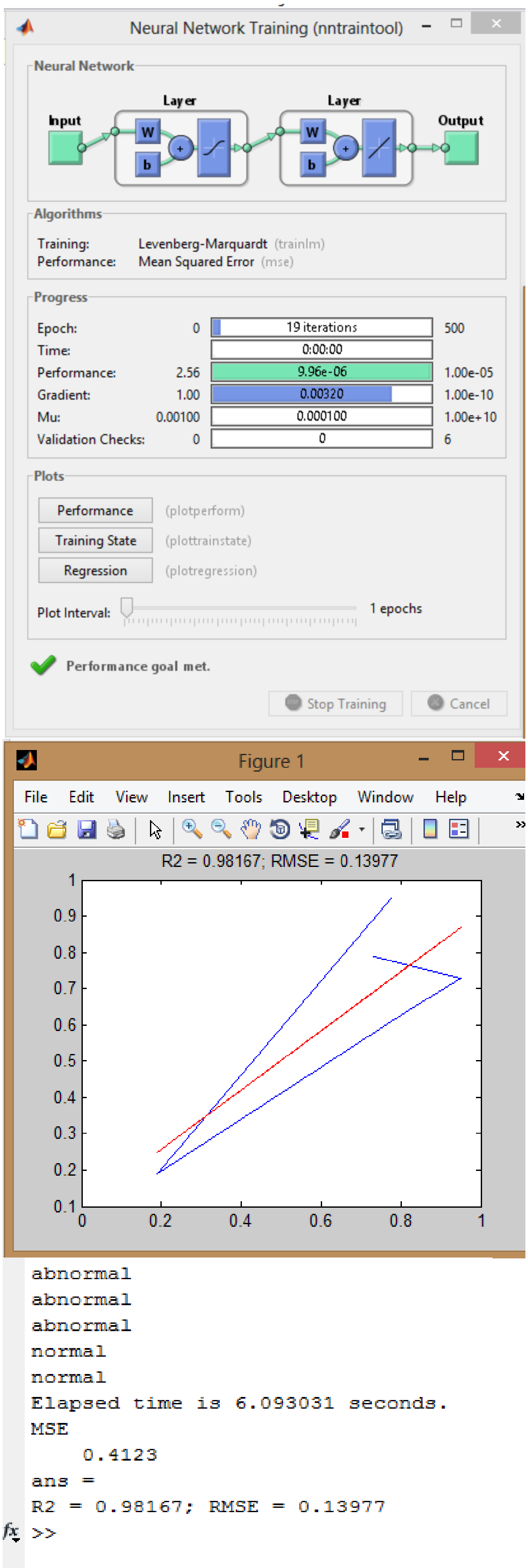


**Fig. 2.7. Experimented Stimulated Output in MATLAB.**

## 5. REFERENCES

[1] J. Cartmell, S. Enoch, D. Krstajic, and D. E. Leahy, "Automated QSPR through competitive workflow," J. Comput. Aided Mol. Des., vol. 19, no. 11, pp. 821–833, Nov. 2005.

[2] J. Cartmell, D. Krstajic, and D. E. Leahy, "Competitive workflow: Novel software architecture for automating drug design," Current Opinion Drug Discovery Develop., vol. 10, no. 3, pp. 347–352, May 2007.

[3] P. B. Dobrin, "Distribution Of Lamellar Deformation Implications For Properties Of The Arterial Media," Hypertension, Vol. 33, Pp. 806–810, 1999.

[4] P. Z. Douglas, P. Libby, R. O. Bonow, And E. Braunwald, Braunwald's Heart Disease: A Textbook Of Cardiovascular Medicine.

[5] N. Filipovic, S. Mijailovic, A. Tsuda, And M. Kojic, "An Implicit Algorithm Within The Arbitrary Lagrangian-Eulerian Formulation For Solving Incompressible Fluid Flow With Large Boundary Motions," Comput. Methods Appl. Mech. Eng., Vol. 195, Pp. 6347–6361, 2006.

[6] N. Filipovic, M. Kojic, M. Ivanovic, B. Stojanovic, L. Otasevic, And V. Rankovic, Medcfd, Specialized Cfd Software For Simulation Of Blood Flow Through Arteries. Kragujevac, Serbia: Univ. Kragujevac, 2006.

[7] N. Filipovic, M.Kojic, M. Ivanovic, Anda. Tsuda, "Blood Flowand Thrombosis Modeling In Large Arteries Parallel Blood Flow Simulation (Pbfs)," Presented At The Bioinfogrid Symp., Milan, Italy, Dec. 10–13, 2007.

[8] M. F. Fillinger, S. P. Marra, M. L. Raghavan, And F. E. Kennedy, "Prediction Of Rupture Risk In Abdominal Aortic Aneurysm During Observation:Wall Stress Versus Diameter," J. Vascular Surg., Vol. 37, No. 4, Pp. 724–732,2003.

[9] S. Glagov, C. Zarins, D. P. Giddens, And D.N.Ku,"Hemodynamicsand Atherosclerosis Insights And Perspectives Gained From Studies Of Human Arteries," Archives Pathology Laboratory Med., Vol. 112, Pp. 1018–1031, 1988.

[10] Glite User Guide. (2009). [Online]. Available:Http://Glite.Web.Cern.Ch/Glite/Documentation/Default.Asp.

[11] M. A. Hall, "In Correlation-Based Feature Selection Of Discrete And Numeric Class Machine Learning," In Proc. Int. Conf. Mach. Learning, San Francisco, Ca, 2000, Pp. 359–366.

[12] M. Koji´C, N. Filipovi´C, B. Stojanovi´C, And N. Koji´C, Computer Modeling In Bioengineering. Hoboken, Nj: Wiley, 2008