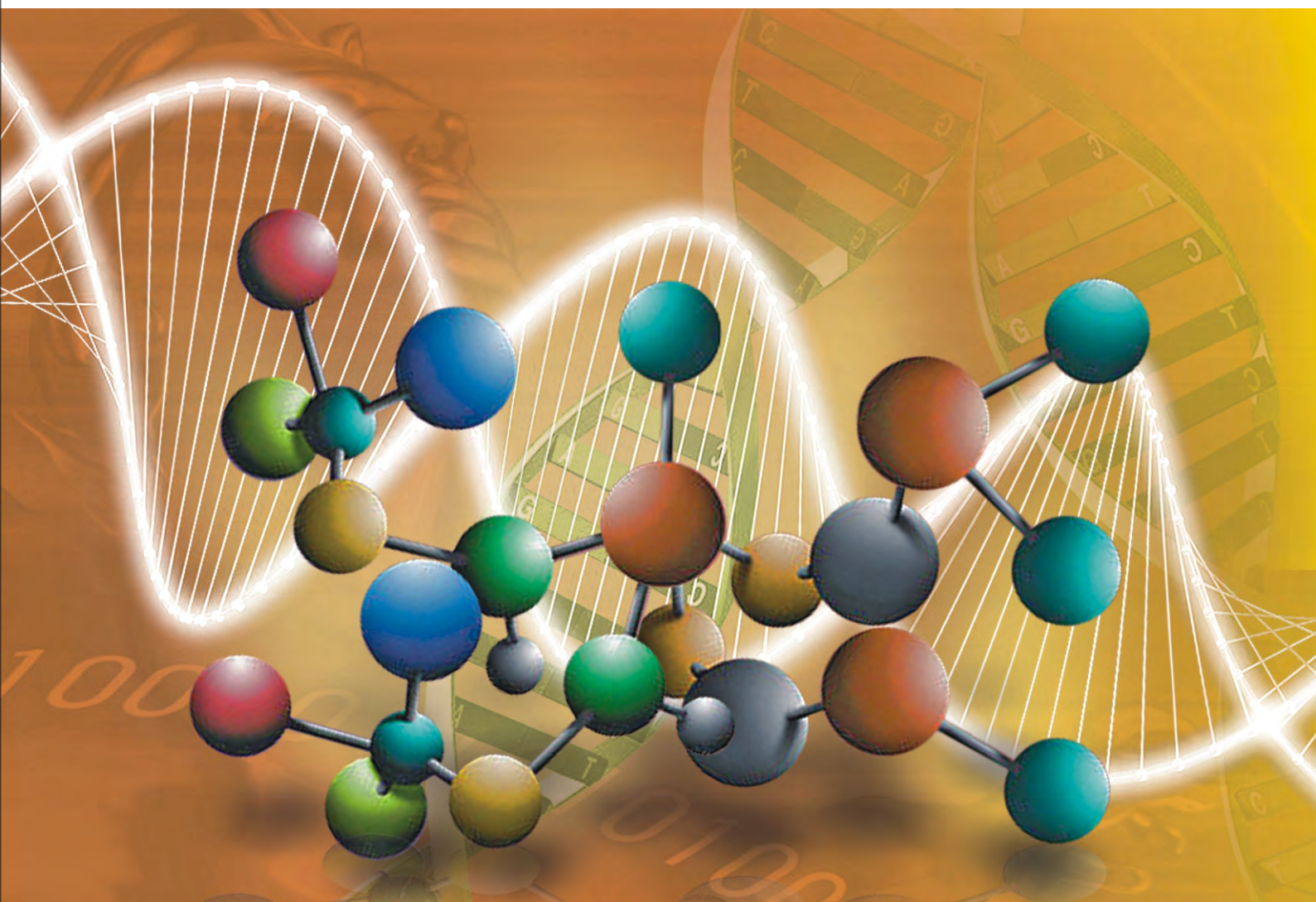
www.scirp.org/journal/jbise

Scientific Research

JBiSE

ISSN: 1937-6871
Volume 3 Number 6 June 2010

Journal of Biomedical Science and Engineering



ISSN: 1937-6871
9 771937 687008 06

Editor-in-Chief
Kuo-Chen Chou

Journal Editorial Board

ISSN 1937-6871 (Print) ISSN 1937-688X (Online)
http://www.scirp.org/journal/jbise

TABLE OF CONTENTS

Volume 3, Number 6, June 2010

Journal of Biomedical Science and Engineering (JBiSE)

Journal Information

SUBSCRIPTIONS

The *Journal of Biomedical Science and Engineering* (Online at Scientific Research Publishing, www.SciRP.org) is published monthly by Scientific Research Publishing, Inc., USA.

Subscription rates:

Print: $50 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jbise@scirp.org

J. Biomedical Science and Engineering, 2010, 3, 543-549 **JBiSE**
doi:10.4236/jbise.2010.36076 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Immune reaction characteristics and the mechanism of anergy induced by recombinant enterotoxin a of *Staphylococcus aureus*

Shang Wu[1,2*], Renli Zhang[1*], Dana Huang[1], Yijie Geng[1], Shitong Gao[1], Xiaoheng Li[1] Zhangli Hu[2]**

[1]Shenzhen Centre for Diseases Control and Prevention, Shenzhen, China; *They contributed equally to this work;
[2]College of Life Science, Shenzhen University, Shenzhen, China; **Corresponding Author.
Email: huzl@szu.edu.cn

Received 30 November 2009; revised 5 January 2010; accepted 15 January 2010.

ABSTRACT

To study immune reactions and the mechanism of anergy induced by recombinant enterotoxin A (rSEA) of *Staphylococcus aureus*. The gene encoding SEA was cloned from standard strain of *S. aureus* and high efficiently expressed in *E. coli*. After immunization with purified rSEA, mice were examined for production of specific antibody, subtype of IgG, cytokine mRNA levels such as IFN-γ, IL-2 secretion and T-cell surface PD-1 expression. Results showed that high levels of specific antibodies were produced in two weeks of primary immunization shot. During this time, humoral immune reactions prevailed (IgG_{2a}/ IgG_1 < 1). During the early phase, Th1 type cytokine mRNA is expressed at a higher level than Th2 type, indicating cellular immune reaction prevailed. Splenocyte IFN-γ secretion was significantly decreased after boosting immunization. The PD-1 expression was detected by a flow cytometry examination in the surface of T- lymphocytes which were induced by rSEA, and the expression of PD-1 molecules increased along with the number of boosting and the time after immunization.

Keywords: *Staphylococcus Aureus*; Enterotoxin A; Immune Respose

1. INTRODUCTION

Superantigens are distinguished from ordinary antigens by the ability that activates multiple T-cell clones. A very small amount can very effectively initiate the activation of immune system. Thus, superantigens have been considered to be widely applicable in tumor immune therapy. Their efficacy has also been considered superior to exogenous cell factor. With the development of relevant theory, superantigens have attracted much attention of studies [1,2]. Earlier studies applied superantigens independently in anticancer therapy. Later, they were used after modification with targeted monoclonal antibody, or as an enhancer of tumor vaccine, or gene vaccine or in combination with other therapy. In the United States, a superantigen Fab-SEA has been tested for anti-cancer ability in phase I clinical studies. In China, superantigens have been used to promoted leukocytes [3,4].

Most currently known exogenous superantigens are toxins from bacteria. For example, the antigen studied in the currently report, enterotoxin A of *Staphylococcus aureus*, is one of these toxins. Superantigens that produced by bacteria upon infection may cause shock, fever, dehydration, skin eruption, organ failure, even death. The pathogenesis is that toxin superantigens stimulate large number of T-cells to proliferate, induce the secretion of cytokines from T-cell and APC cell, which result in immune system disorder. Superantigens induced disease may occur through several ways: 1) Superantigen is the directly cause of the disease, such as toxic shock and food intoxication. 2) Superantigen enhances the effect of the other infectious factors. 3) Superantigen induces autoimmune reaction by activating large number of T and B cells [5]. Up to date, there is no proof on the relationship between superantigen and some disease. However, the effects of superantigens in food intoxication, toxic shock and some infectious disease are very clear [6]. An understanding of the mechanism by which bacterial superantigens activate T-cell and pathogenesis is theoretical bases for the application of superantigens in cancer immunotherapy. To this end, we studied the humoral and cellular immune reactions to recombinant SEA protein (rSEA), explored the relationship between inhibitory lymphocyte receptors and anergy induced by rSEA.

2. EXPERIMENTS

2.1. Materials and Methods

2.1.1. Materials

Staphylococcus aureus was provided by Shenzhen Cen-

tre for Diseases Control and Prevention. Two strains of *E. coli*, DH5α, BL21(DE3) and pET-28a were the collection in the author's laboratory. Cloning vector pGEM-T easy was purchased from TaKaRa.

Antibody FITC-CD3, PE-PD-1 and corresponding negative control were purchased from BD Company. ELISA kit was purchased from Shenzhen Biotech Limited. Trizol and reverse transcription kit were purchased from TaKaRa Bio Inc. 1640 media, Lympho-Spot TM serum-free media and mice IFN-γ ELISPOT kit were purchased from Dakewei Biotech Ltd. One hundred and eight BALB/c mice were purchased from Guangdong Center of Experimental Animal, maintained according to standard clean protocol. The mice used in the experiment were 15-20 g of body weight, 6-8 weeks of age.

2.1.2. Preparation of rSEA

Primers were derived from GenBank SEA sequence (AY827552): sea F: 5'GCC GCT AGC ATG AAA AAA ACA GCA TTT ACA TTA C 3' (underlined is *NheI* digestion site); sea R: 5'CGC CGT CGA CTT AAC TTG TAT ATA AAT ATA TAT CAA 3' (underlined is *Sal I* digestion site).

DNA template was from SEA producing standard strain. PCR product was prepared with sea F and sea R primers, separated with 1.5% agarose gel electrophoresis. After purification, the PCR product was inserted in pGEM-T easy vector. The recombinant plasmid was propagated in DH 5α, selected with ampicillin on LB plate.

An expression vector of SEA was constructed by subcloning of the gene in pET-28a plasmid with *E. coli* BL21 (DE3) host. A recombinant clone was cultured in shaking incubator and induced by IPTG (final concentration 1 mmol/L) for 6 hours. Bacteria were centrifuged, lysed with ultrasound. After centrifugation, rSEA was in the pellet. Subsequently, the pellet was resuspended in 8.0 mol/L urea, and purified by Ni^{2+} affinity chromatography. The elution solution was imidazole (500 mmol/L). The rSEA protein was refolded and further purified to a high purity.

2.1.3. IgG Level Determination by ELISA Assay

Twenty BALB/c mice were assigned to either experiment group or control group at random. Lyophilized rSEA was resuspended in PBS and diluted to targeted concentration, mixed with equal volume of adjuvant, and injected subcutaneously into mice. The first injection was 100 μg, the second and third injection was 50 μg each at 2 weeks interval. Tail blood was collected each week before and after injection for 8 continuous weeks.

rSEA protein was diluted in embedding buffer to 10 μg/mL, aliquoted to multi-well plate (100 mL/well). The plate was incubated at 4℃ overnight, and then washed with PBST (PBST containing 0.05% Tween-200) three times, blocked with blocking solution for 1 h at 37℃. Subsequently, serum from each experiment groups were diluted 1:1000 in blocking solution, added to the plate, incubated at 37℃ for 1 hour. The plated was then washed with PBST for 3 times. HRP conjugated Rabbit anti-mouse IgG (1:5000) was added to the plate and incubated for 1 hour at 37℃, washed with PBST three times. Finally, diaminobenzene substrate solution containing hydrogen peroxide was added to the plate and allowed to develop for 10 min in dark. The reaction was stopped by addition of 2 mol/L H_2SO_4. Absorbance was determined with spectrometry at 450 nm.

2.1.4. Subtype Antibody Level Determination with ELISA

Serum antibody subtypes were determined at 3, 5 and 7 weeks after immunization. The method was the same as described above, except the secondary antibody was replaced by HRP conjugated mice IgG_1 (1:1000) and mouse IgG_{2a} (100 μL/mL).

2.1.5. RT-PCR Analysis of Spleen Cytokine Expression

Sixty BALB/c mice were assigned to two groups and immunized with rSEA. Splenocytes were collected from 5 mice in each group at 0 h, 2 h, 12 h, 24 h, 2 w and 3 w post immunizations. Spleen total RNA was extracted with Trizol according to manufacturer's instruction. The mRNA levels of IL-2, IL-4 were determined with RT-PCR using specific primers. IL-2 primers were 5'- CTT CAA GCT CCA CTT CAA GCT-3' (forward) and 5'-CCA TCT CCT CAG AAA GTC CAC C-3' (reverse). The amplicon was 198 bp. IL-4 primers were 5'-CAT CGG CAT TTT GAA CGA GGT CA-3' (forward) and 5'-CTT ATC GAT GAA TCC AGG CAT CG-3' (reverse). The IL-4 amplicon was 203 bp. The expression of β-actin was used as internal control. The primers for β-actin were 5'-CAT CCG TAA AGA CCT CTA TGC CAA C-3' (forward) and 5'-ATG GAG CCA CCG ATC CAC A-3' (reverse). The β-actin amplicon was 238 bp. The thermal cycles of PCR were: pre denaturation at 94℃ for 5 min, 30 cycles of amplification (94℃ 30 s, 55℃ 30 s and 72℃ for 1 min). RT-PCR product was analyzed by 1% agarose gel electrophoresis [7].

2.1.6. Relative Quantification of Cytokine Expression

Gel image was scanned with UVI system, and bands were quantified with UV band software. Band intensities were determined using β-actin as internal control. The relative expression levels (Ling, R.Y., *et al.*,) of cytokines were determined by:

Relative expression level = (test gene band inten-

sity)/(β-actin band intensity)×100%.

2.1.7. IFN-γ Detection by ELISPOT Assay

Twenty four mice were assigned to either control group, or single immunization group or boost group (n = 8), injected intraperitoneally with rSEA (100 μg/animal) or PBS (control group). Mice were sacrificed by cervical dislocation 24 h after the last injection, and splenocytes were isolated with sterile procedure.

Lymphocyte preparation: spleen tissue was pressed against 200 micron mesh, filtered and spun. The second layer low density cells were collected and washed with 1640 medium. Cells were resuspended in Lympho-Spot TM serum-free medium. Cell concentration was adjusted to 2×10^6/mL and examined with trypan blue exclusion assay. Cell viability was greater than 95%.

IFN-γ detection with ELISPOT pre-embedded kit: 1) plates were seeded with splenocytes (1×10^5 cells/well), which were stimulated with 5 μg/mL rSEA antigen. Each sample was done with triplicates. ConA (5 μg/mL) or medium was added to the positive or negative control wells respectively. Cells were cultured at 37℃ for 36 h. After wash, biotinated anti mouse IFN-γ was added to the wells (100 μL/well), and the plate was incubated for 1 h at room temperature. 2) After washed, streptoavidin-HRP was added to each well (100 μL/well) and incubated for 1 h at room temperature. 3) Upon wash, HRP substrate AEC was added to the plate. Color was developed in dark. The plate was washed with water and dried. 4) Spots were counted. The unit was defined as spots/10^5 spleen cells. Negative control has less than 10 spots/10^5 splenocytes. Test wells that had greater than 2-fold the spot number of the negative control well was counted as positive.

2.1.8. PD-1 Expression Determination by Flow Cytometry

Sixty mice were assigned to either single immunization group or boost group or control group (n = 20) and injected with rSEA or PBS. Animals were sacrificed at 2 h or 24 h post last immunization. Splenocytes were isolated as described above. Cells were stained with FITC- conjugated anti-CD3 and PE-labeled PD-1. After fixation, cells were examined with flow cytometer. Cell concentration was adjusted to 2×10^6 cells/mL. 500 μL of the cell suspension was loaded to each of two flow tubes. FITC Anti-Mouse CD3 (5 μL), PE Anti-Mouse PD-1 (5 μL) or equal volume of respective control solution was added to the tubes. After mix by shaking, the tubes were set in dark for 45 min at room temperature. Cells were spun, washed with PBS once and resuspended in 200 μL PBS, fixed with 500 μL paraformaldehyde (4%) and examined with flow cytometry.

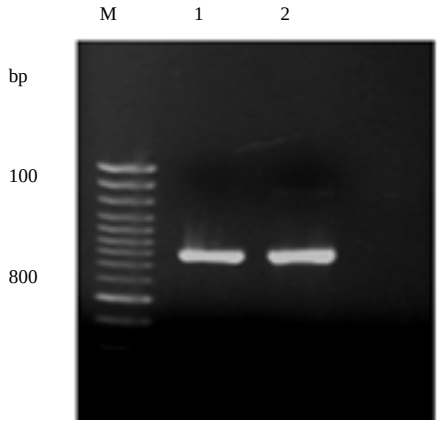


Figure 1. Amplification of SEA gene from *Staphylococcus aureus* genomic DNA. M was 100 bp DNA Ladder, lane 1, 2 was PCR product of SEA gene.

2.1.9. Statistical Analysis

Data were analyzed with SPSS 11.5 software. Results were presented as mean ±s. Comparisons between groups were analyzed with ANOVA. Statistical significances were inferred when $p < 0.05$ or $p < 0.01$.

3. RESULTS AND DISCUSSION

3.1. Cloning of SEA Gene

With the specific primers, we identified 2 strains, out of 10 wild type *S. aureus*, to be positive for producing enterotoxin A (**Figure 1**). The PCR product was inserted into pGEM-T easy vector and propagated in DH5α. The plasmid insert was sequenced, which confirmed that the insert sequence was identical to AY827552. The size was 786 bp encoding 261 residues.

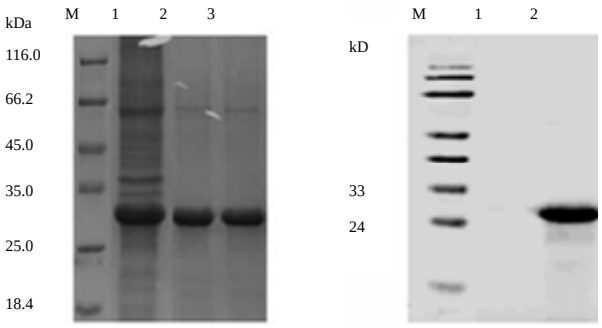
3.2. Expression, Purification and Refolding of rSEA

The recombinant plasmid pET-28-SEA was transformed into *E. coli* BL21 cells and the transformants were induced with IPTG for 6 h. Bacterial cells were disrupted with ultrasound. Upon centrifugation, the recombinant protein was identified in the precipitate, demonstrating that rSEA was expressed as inclusion body. During Ni^{2+} affinity chromatography, rSEA was eluted at 500 mmol/L imidazole. After purification and refolding, high purity rSEA was obtained (**Figure 2(a)**).

Immunobloting analysis showed that the refolded SEA was reactive with specific polyclonal antibody and the molecular weight was as expected 31 kD. This result demonstrated that the recombinant SEA has similar antigenicity as the natural one (**Figure 2(b)**).

3.3. Humoral Immunity Induced by rSEA

In order to study the process of humoral immunity induced by SEA, we examined specific IgG production in sera from mice immunized three times with conventional



(a) (b)

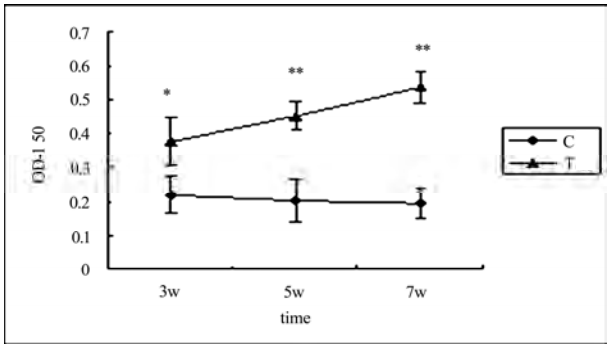
Figure 2. (a) SDS-PAGE analysis of recombinant SEA(rSEA) preparation by affinity chromatography. M was molecular weight markers, lane 1 was cell extract from pET28a-rSEA transformed *E.coli*, Lane 2,3 was recombinant SEA of affinity chromatography; (b) Western-blot analysis of purified rSEA protein. M was molecular weight markers, lane 1 was negative sera act as control sera, lane 2 was anti-rSEA sera act as first antibody.

protocol. Results showed that specific antibody level in the immunized mice (OD_{450} = 2.492 ± 0.082) was markedly higher than that in the control mice (OD_{450} = 0.054 ± 0.032) three weeks after the first immune shot ($P < 0.01$). There were no significant changes in antibody levels before and after the second (week 3, 4) and third shots (week 5, 6) in the immunized mice ($P > 0.05$). Two weeks after the last shot, the antibody level was slightly reduced, but not significantly ($P > 0.05$).

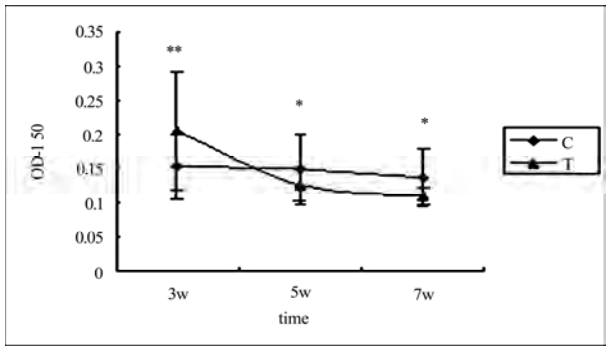
To evaluate the immunogenicity of rSEA as a candidate of cancer therapeutics, we assessed specific serum antibody in immunized mice, investigated the process of rSEA induced humoral immunization. The key steps of humoral immunization are the activation and proliferation of B-lymphocytes, which require the stimulation of secreted factor or exogenous antigen and the assistance of $CD4^+$ T cells. We performed ELISA analysis on the mouse sera collected at different time point, discovered that specific immunoreactions were strongly induced within two weeks of immunization. (The immunized animals produced significantly higher antibody levels than the control animals and) the high antibody levels were maintained for a long time. Compared with a single shot immunization, boost immunization did not increase rSEA specific antibody levels (Specific IgG induced by rSEA can be used to destroy cancer cells through the activation of complement and superantigen-dependent cell-mediated cytotoxicity) [7].

3.4. IgG Subtype Induced by rSEA

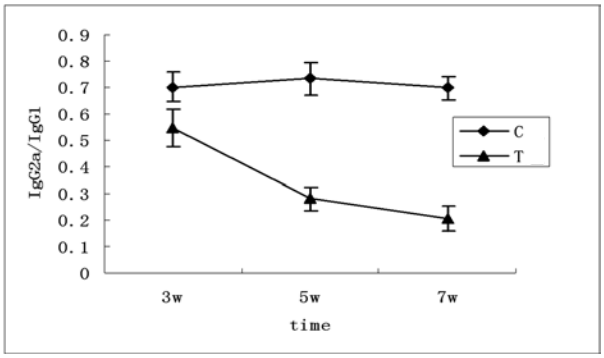
Changes in specific IgG subtype levels were shown in **Figure 3**. IgG_1 levels were significantly different between control group (C group) and test group (T group) three weeks after immunization ($P < 0.05$). The differences were even more pronounced 5 and 7 weeks after immunization ($P < 0.01$). IgG_{2a} levels were different between the two groups only at three weeks after immunization ($P < 0.05$). In the test group, specific IgG_1 levels continuously increased along with time and the number of immune shots (**Figure 3(a)**). In contrast, IgG_{2a} levels decreased during the same period (**Figure 3(b)**). The ratio of IgG_{2a}/IgG_1 did not change in the control mice at 3, 5 and 7 weeks (data not shown). However, this ratio showed a trend of decrease in the test group (**Figure 3(c)**), which



(a)



(b)



(c)

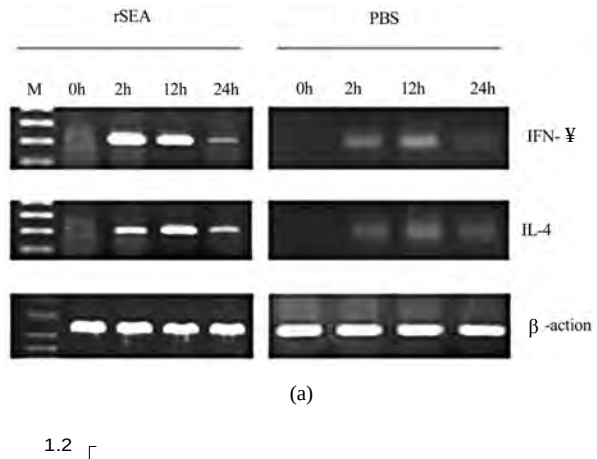
Figure 3. Specific IgG subtypes induced by rSEA from mice immunized. (a) IgG_1; (b) IgG_{2a}; (c) IgG_{2a}/IgG_1

was significantly different from that of control group at 5 and 7 weeks ($P < 0.05$).

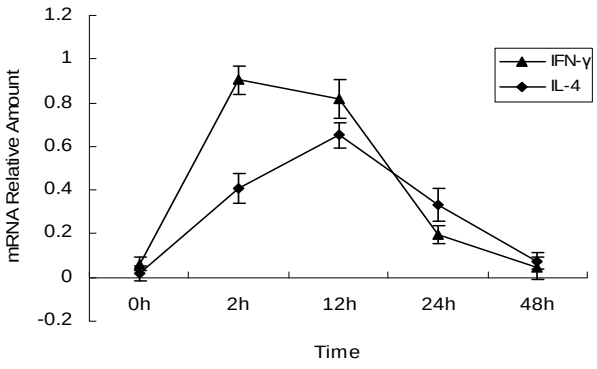
IgG_1 (Th2) and IgG_{2a} (Thl) are type markers of immunoreactions. The ratio of IgG_{2a}/IgG_1 indicates whether the humoral immunoreactions are dominated by Th1 or Th2. In the current study, we selected peripheral serum from mice two weeks after immunization, assessed levels of IgG_1 and IgG_{2a} with ELISA. Results showed that IgG_1 level was higher than IgG_{2a} level after initial immunization, and that IgG_1 exhibited a trend of increase along with time and the increase in the number of shots (**Figure 4(a)**), while IgG_{2a} level showed a trend of decrease (**Figure 4(b)**). Thus, the ratio of $IgG_{2a}/IgG_1$was always smaller than 1, and exhibited a trend of decline (**Figure 4(c)**). These data indicated that rSEA-induced immunity was dominated by Th2 reaction and the dominance tends to be enhanced by boost shots.

3.5. Effects of rSEA on Splenocyte Cytokine mRNA Levels

Messenger RNA levels of IFN-γ, IL-4 and β-actin in splenocytes were analyzed with RT-PCR. The PCR products were the expected sizes 198 bp, 203 bp and 238 bp (**Figure 4(a)**). Before immunization (0 h), IFN-γ, IL-4 mRNA was not detectable. After immunization, mRNA levels increased. Levels of mRNA in immunized groups were significantly higher than that in PBS group at 2, 12 and 24 h ($P < 0.01$). There was no significant difference between the two groups in the mRNA levels 24 h after immunization. At 2 and 12 h after immunization, IFN-γ mRNA levels were higher than IL-4 in the immunized grouping ($P < 0.01$). In contrast, at 24 h after immunization, IL-4 mRNA level was higher than IFN-γ mRNA ($P < 0.05$) (**Figure 4(b)**).



(a)



(b)

Figure 4. (a) RT-PCR results of cytokine in spleen of immunized mice with rSEA, PBS acted as control of immunization, and fragment of β-action was as a RT-PCR control; (b) Analysis of Cytokine mRNA expression in spleens of immunized mice at different time.

We investigated mRNA levels of two cytokines after immunization with semi-quantitative RT-PCR. Results showed that both cytokines were greatly increased after induction with rSEA. At 2 h post immunization, IL-2 mRNA was higher than IL-4 mRNA level and reduced soon after. In contrast, IL-4 mRNA was low at 2 h after immunization and gradually increased, peaked at 12 h and then gradually decreased. Therefore, during immune reaction, Th1 type cytokine proliferated earlier than Th2 type cytokine, but the level rapidly reduces to become lower than Th2 cytokine level. This kinetic process is consistent with the change in serum IgG subtype [8,9]. Microphages (MΦ) are very important immune cells with unique anti-cancer effect. They provide immune surveillance, antigen presentation and effector functions. Through antigen presentation, MΦ activates T-cell and enhances specific anti-cancer immunity. MΦ also non-specifically destroys cancer cell upon contact. Activated MΦ releases many cytokines and bioactive factors that regulate cancer immunity [10]. IFN-γ secreted by T cells and NK cells is the most potent MΦ activator. However, IFN-γ production during tumor genesis is insufficient, leading to insufficient MΦ activation. Superantigens have powerful immune activation ability that can induce the release of large amount of cytokines. This point has been proven by previous experiments [11,12]. To further study the relationship between rSEA and IFN-γ secretion, we analyzed splenocyte IFN-γ secretion in immunized mice with ELISPOT technique. Our results showed that splenocyte can secret IFN-γ at high frequency after primary immunization, (which is significantly different from the control splenocytes). This cellular immune reaction is rSEA specific, because splenocyte stimulated by other antigens produced little IFN-γ. Boost shot did not enhance antibody production ($P > 0.05$), suggesting rSEA as a superantigen may induce anergy after primary immunization [13].

3.6. Effects of rSEA on Splenocyte IFN-γ

To further explore immunoreactions to rSEA, we as sessed

changes in IFN-γ secretion by splenocytes with ELISPOT assay. As shown in **Figure 5**, little specific IFN-γ were produced without stimulators. In contrast, IFN-γ was produced in the three groups of mice stimulated with rSEA or ConA. Mice in single shot group and boosted group produced more IFN-γ than that in control group ($P < 0.01$). There were no differences in IFN-γ levels between the single shot group and the boost group ($P > 0.05$).

Regulation of immunity includes initiation and termination of immune reactions, dependent of internal and external signal. External signal is transmitted through surface receptors. There are stimulatory and inhibitory receptors that positively or negatively regulate cell activation. Inhibitory receptors of the B7 family include cytotoxic T lymphocyte associated antigen 4 (CTLA-4), programmed death-1 (PD-1) and B and T lymphocyte attenuator (BTLA) [14]. These three inhibitory receptors belong to the CD28 family. Upon binding with different members of B7 family, these receptors can inhibit activation of immune reactions. The mechanism has been widely investigated. Studies have shown that interaction between PD-1 and its complementary PD-L leads to the inhibition of T-cell proliferation and the attenuation of IL-2, -10 and IFN-γ secretion. This mechanism is very important for organ transplant, autoimmune disease and cancer immunity. An examination of splenocyte IFN-γ secretion with ELISPOT assay revealed that boost shot did not enhance cellular immunity, suggesting possibility of anergy after rSEA immunization.

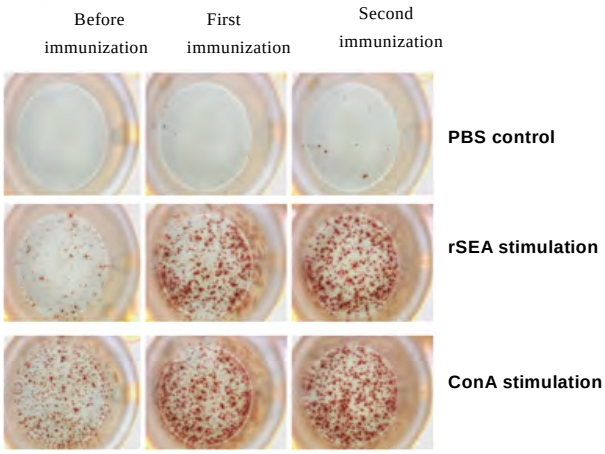


Figure 5. ELISPOT analysis of IFN-γ in mice spleen immunized with rSEA at different stage of immunization.

3.7. T-Cell PD-1 Expression Induced by rSEA

PD-1 expression in T-cell was examined 2 h and 24 h after a single shot immunization or boost immunization. Splenocytes were stained with FITC-anti-CD3 and PE-anti-PD-1, fixed and examined with flow cytometer. Results showed that T-cell surface PD-1 expression in the rSEA immunized mice were significantly higher than that in the control mice ($P < 0.01$). Significant differences were observed between 2 h and 24 h in the expression of PD-1 in the single shot and boost shot immunization. Also, significant differences were also observed in PD-1 expression between single shot and boost shot ($P < 0.91$) at 2 h and 24 h (**Figure 6**).

We examined the expression of PD-1 in immunized splenocytes with flow cytometry to interrogate its relationship with immunized time and the number of shots. Results showed rSEA immunized mice had higher PD-1 expression and boost shots further enhanced PD-1 expression. PD-1 levels in boosted group were higher than that in the non-boosted group. Also PD-1 levels at 24 h after immunization were higher than that at 2 h. This reaction was rSEA-specific, since the PD-1 level in control group did not change over time. These results explained unresponsiveness of IFN-γ level to boost shots, suggesting that inhibitory receptor PD-1 attenuated T-cell proliferation and IFN-γ secretion [15].

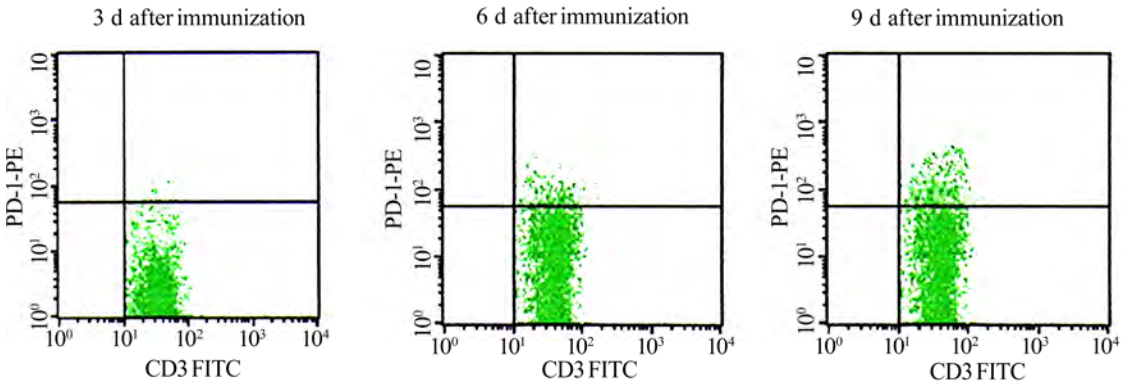


Figure 6. Expression of PD-1 molecular on spleen T cells of mice immunized with rSEA, FCM showed different expression of PD-1 molecular on 3d ,6d and 9 d.after immunized.

4. CONCLUSIONS

Superantigens are powerful activator of T-cell. Distinct from ordinary antigens, superantigens can directly attach to the outer groove of MHC-II and Vβ domain of TCR on T-cell surface, without the processing by antigen presentation cells. Even a trace amount of superantigen can activate 5%-20% of T-cells. Therefore, superantigens have been tried in cancer immune therapy to promote internal anticancer immunity. Good results have been obtained from these trials. However, it has also been found that superantigen may induce apoptosis and inability after activating T-cells, leading to the attenuation of response to additional stimulation. This property of tolerance induction limits the efficacy of superantigens. Therefore, we have used rSEA to immunized mice, studied the characteristics of rSEA induced specific humoral immunity and cellular immunity and the mechanism of anergy. These studies could provide foundations for further studies of superantigens as tumor suppressors.

rSEA has superantigen properties and that it can induce powerful humoral and cellular immune responses. However, boosting immunization with rSEA caused anergy through PD-1 mediated inhibition.

5. ACKNOWLEDGEMENTS

The authors acknowledge research funding from the National Natural Science Foundation of China (Grant No. 30770340,30470281), the national major program of Science and technology for water pollution control and restoration in china (Grant No. 2009ZX07423-003) and Shenzhen Grant Plan for Science and Technology.

REFERENCES

[1] Lu, S.Y., Sui, Y.F., Li, Z.S., *et al*. (2004) Superantigen-SEA gene modified tumor vaccine for hepatocellular carcinoma: An *in vitro* study. *World Journal of Gastroenterology*, **10(1)**, 53-57.

[2] Proft, T. and Fraser, J.D. (2003) Bacterial superantigens. *Clinical and Experimental Immunology*, **133(3)**, 299-306.

[3] Balaban, N. and Rasooly, A. (2000) Staphylococcal enterotoxins. *International Journal of Food Microbiology*, **61(1)**, 1-10.

[4] Cheng, J.D., Babb, J.S. and Langer, C. (2004) Individualized patient dosing in phase I clinical trials: The role of escalation with overdose control in PNU-214936. *Journal of Clinical Oncology*, **22(4)**, 602-609.

[5] Han, W.N., Cao, Y.X. and Li, Qi. (2004) The different influence of SEA on the direnfiafion and proliferation of Tcl and Tc2 subsets. *Chinese Journal of Microbiology and Immunology*, **24(4)**, 300-303.

[6] Petersson, K., Forsberg, G. and Walse, B. (2004) Interplay between superantigens and immunoreceptors. *Scandinavian Journal of Immunology*, **59(4)**, 345-355

[7] Melanie, J.R., Nisebita, S. and Avery, A. (2006) Differential regulation of cytokine production by CD1d-restricted NKT cells in response to superantigen staphylococcal enterotoxin B exposure. *Infection and Immunity*, **74(1)**, 282-288

[8] Mondal, T.K., Bhatta, D., Biswas, S. and Pal, P. (2002) Superantigen-induced apoptotic death of tumor cells is mediated by cytotoxic lymphocytes, cytokines, and nitric oxide. *Biochemical and Biophysical Research Communications*, **290(4)**, 1336-1342.

[9] Zhang, M.H., Cao, X.T. and Chen, G.Y. (1999) Presentation of tumor antigen to T lymphocytes by macrophage like bone marrow stromal cells. *Chinese Journal of Microbiology and Immunology*, **19(5)**, 412-416.

[10] Ling, R.Y., Ding, J.B. and Wen, H. (2003) A semi-quantitative RT-PCR analysis for IL-2/IL-4 mRNA expression in heart transplantation in rats. *Journal of Xinjiang Medical University*, **26(5)**, 427-429.

[11] Frauwirth, K.A., Alegre, M.L. and Thompson, C.B. (2000) Induction of T cell anergy in the absence of CTLA-4/B7 interaction. *Journal of Immunology*, **164**, 2987-2993.

[12] Honstettre, A., Mège, J.L., Gérard, L., Aubaniac, J.M. and Drancourt, M. (2003) Relationship of relapsing hip prosthesis infection by *Staphylococcus aureus* with gamma interferon deficiency. *Journal of Clinical Microbiology*, **41(11)**, 5344-5346.

[13] Plaza, R., Rodriguez, J.L. and Juarez, C. (2007) Staphylococcal enterotoxin B in vivo modulates both gamma interferon receptor expression and ligand-induced activation of signal transducer and activator of transcription 1 in T cells. *Infection and Immunity*, **75(1)**, 306-313.

[14] Xu, G.L., Zhu, X.H., Guo, B. and Wu, Y.Z. (2004) Involvement of CTLA-4 in T-cell anergy induced by staphylococcal enterotoxin A in vitro. *Molecular Immunology*, **41(1)**, 1-8.

[15] Keir, M.E., Francisco, L.M. and Sharpe, A.H. (2007) PD-1 and its ligands in T-cell immunity. *Current Opinion in Immunology*, **19(3)**, 309-314.

J. Biomedical Science and Engineering, 2010, 3, 550-555 **JBiSE**
doi:10.4236/jbise.2010.36077 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Gene silencing of E-selectin block recruitment of endothelial progenitor cell to vascular endothelium under flow

Sunil Sharma, Masayuki Yoshida

Life Science and Bioethics Research Center, Tokyo Medical and Dental University, Tokyo, Japan.
Email: masavasc@tmd.ac.jp

Received 16 February 2010; revised 20 March 2010; accepted 25 March 2010.

ABSTRACT

Short interfering RNA (siRNA) is a powerful technique that can suppress gene expression in a variety of cells including mammalian cells. Endothelial progenitor cells (EPCs) are bone marrow—derived haematopoietic progenitor cells that have been implicated in vasculogenesis. We demonstrated for the first time that gene silencing of endothelial E-selectin using siRNA transfection in human umbilical vein endothelial cells (HUVECs) causes inhibition of EPC adhesion under flow conditions. Fluorescence immunobinding assay analysis showed that significant reduction of E-selectin surface expression in HUVECs (activated with IL-1β (10 U/mL) for 4 h) transfected with siRNA against E-selectin, but not in HUVECs transfected with *LacZ* siRNA (control). An EPC adhesion assay under flow conditions (shear stress = 1.0 dyne/cm^2) then demonstrated that HUVECs transfected with E-selectin siRNA supported significantly less adhesion of EPCs than those HUVECs treated with control siRNA and no siRNA after activation by IL-1β ($p < 0.05$). Our experiments have shown the importance of E-selectin in EPC adhesion to HUVECs and the potential utility of gene silencing of E-selectin in EPC recruitment.

Keywords: Endothelial Progenitor Cell; E-Selectin; Rnai

1. INTRODUCTION

The vascular endothelium is a vital border zone between the circulatory system and the tissues it supplies. It is a monolayer of endothelial cells (EC) that has pivotal roles in coagulation, inflammation, vasodilatation and vasoconstriction through substances such as nitric oxide and prostaglandins [1].

Recently identified vascular progenitor cells from bone marrow (BM) and non-BM origin have been shown to contribute to neovessel formation [2]. Upon inflammatory conditions, the vascular endothelium was activated to capture progenitor cells in regions where endothelium regeneration is needed.

Though ischemia is believed to be the physiological stimulus for EPC mobilisation from the bone marrow, the mechanisms of EPC recruitment are less well understood.

Little is understood concerning the molecular mechanisms of EPC adhesion to endothelial cells. In contrast, extensive studies have been conducted regarding leukocyte-endothelial interactions. Previous experiments have demonstrated that P-selectin and E-selectin are involved in leukocyte rolling, whilst intercellular adhesion molecule-1 (ICAM-1) is associated with leukocyte firm adhesion [3]. The molecular players involved in leukocyte transmigration have been found to be molecules such as platelet–endothelial-cell adhesion molecule-1 (PECAM-1), junctional adhesion molecules (JAMs), vascular endothelial-cadherin (VE-cadherin) and CD99 [4].

Chavakis *et al* had implicated roles for β2-integrins in the homing and revascularisation of EPCs [5]. Vajcokzy *et al* had carried out *in vivo* experiments with mice to demonstrate that E-selectin is involved in adhesion of embryonic EPCs [6].

We recently reported an important role of endothelial E-selectin in mediating endothelial progenitor cell (EPC) recruitment [7].

In this study, we used a flow chamber system to mimic the *in vivo* circulation. The major advantage of this set-up is that conditions can be much more tightly controlled, so that the effect of individual molecules can be better ascertained than with an *in vivo* procedure.

RNA inhibition has allowed us to silence genes for a variety of purposes. Short interfering RNA (siRNA) is a useful technique to define the function of a given molecule [8].

In this study we look at the role of E-selectin in EPC adhesion to endothelial cells using double stranded short interfering RNA (siRNA), in conjunction with the *in vitro* flow assay system. Knowledge of this study will allow us to better understand the process of EPC re-

cruitment to endothelial cells, which will give greater insight into the process of vasculogenesis and its clinical implications.

2. MATERIALS AND METHOD

2.1. Cell Culture and Reagents

Human umbilical vein endothelial cells (HUVECs) were cultured on 0.1% gelatin-coated tissue culture dishes as described previously [9,10]. The following antibodies were used in this study; H18/7 [11] (anti-human E-selectin mAb), Hu5/3 [12] (anti-human ICAM-1 mAb). Recombinant human IL-1β was a gift from Biogen Inc (Cambridge, MA). The HL60 cell lines were obtained from the American Type Culture Collection and cultured in RPMI-1640 containing 10% FCS. For use in the flow-chamber apparatus, HUVECs were plated onto 22 mm fibronectin-coated glass coverslips as has been described previously [13,14].

Preparation of endothelial progenitor cells was previously described in detail [7].

In brief, 25 mL of the blood was collected from human volunteers, to which 15 mL Histopaque®-1077 solution (Sigma-Aldrich Japan, KK) was added. After centrifugation for 20 minutes at 1600 rpm, the upper layer was collected and diluted in Dulbecco's phosphate buffered saline, then add 2 mM EDTA (DPBSE, Sigma) and centrifuged at 2000 rpm for 5 minutes. The residual pellet was resuspended in 10 mL ammonium chloride solution (Stem Cell Technologies Inc) to lyse red blood cells and centrifuged at 1000 rpm for 10 minutes. The pellet was resuspended in Endothelial Cell Basal Medium (EBM-2, Clonetics®) and cultured in each well of a C-6 plate (coated with fibronectin). Media exchange of the EPC culture was carried out after 4 days, and EPCs were used for flow assay after 7 days. All experimental procedures were approved by the Experimental Research Review Committee of the Tokyo Medical and Dental University (No. 0090014).

2.2. Flow Assay

The parallel-plate flow chamber that was used in this study has been described previously in more detail [15,16]. In summary, the chamber was composed of 2 aluminium steel plates separated by a 200 μm thick silastic gasket, and the flow channel was formed by removal of a 5 × 20 mm rectangular section from the gasket. Defined levels of flow were applied to the HUVEC monolayer by drawing the flow media (see above) through the channel with a syringe pump (model 44 Harvard Apparatus). A plastic heating plate (Tokai Hit Co) was attached to the stage of an inverted microscope (IX50, Olympus) to maintain the temperature at 37℃. Using this set-up, the channel flow could be modelled as a two-dimensional fully developed laminar flow with a simple parabolic velocity profile.

HUVEC monolayers on coverslips were stimulated for 4 h using IL-1β (10 U/mL), mixed with RPMI and 1% FBS, then positioned in the flow chamber. The monolayers were perfused for 5 minutes with perfusion medium (flow media), then examined carefully to establish the monolayer as confluent. HL60 cells or EPCs were then diluted in the flow media to 1×10^6 cells/mL. The cells were then drawn through the chamber at controlled flow rates to achieve wall shear stresses of 1.0 and 2.0 dyne/cm^2 for 10 minutes.

2.3. Fluorescent Immunobinding Assay (FIA)

HUVEC monolayers in 96-well plates were incubated on ice with the indicated primary antibody in RPMI/1% FCS at 10 μg/mL for 45 minutes. These wells were then washed three times with RPMI/1% FCS, after which they were incubated with a FITC-conjugated goat anti-mouse polyclonal F(ab')$_2$ antibody (Caltag Laboratories) diluted 1:100 in DPBS on ice. After 45 minutes, the wells were washed with DPBS/20% FCS twice, followed by teo washes with DPBS only. Cells were lysed with 0.01% NaOH in 0.1% SDS. The fluorescence was then measured using a CytoFluor II (Perspective Biosystems) fluorescent plate reader set at 485 (excitation)/535 (emission).

2.4. siRNA Transfection of E-Selectin

Short interfering RNAs (siRNAs) were designed to target the coding sequence of human E-selectin cDNA. The target sequences were directed as described previously[8]. Briefly, the sequences were targeted to the single-strand region consistent with the predicted secondary RNA structure and sequences of the form (AA/CA)N_{19} with GC contents of less than 70% were isolated from this region. Nineteen RNA nucleotides followed by TT/TG were selected, then chemically synthesised, and finally gel-purified.

In order to anneal the single stock strands of RNAi (JbioS) 20 μL 50 μM sense, 20 μL 50 μM antisense and 10 μL 10 mM $MgCl_2$ in DEPC-PBS was mixed together, giving a dsRNA concentration of 20 μM. This was then heated at 95℃ for 2 minutes, 70℃ for 1 minute, 20℃ for 30 minutes, and then 4℃ before being stored at –80℃. Solutions of 120 μL Optimem® (Invitrogen), 6 μL double-stranded siRNA for E-selectin (siE-01) (or control using *LacZ* (scE-01) or no siRNA (*i.e.* Lipofectin alone)), 120 μL Optimem® (Invitrogen) and 5 μL Lipofectin® (Invitrogen) were prepared, and left for 30 minutes then combined and again left for 30 minutes. Optimem® was then added to make up a total volume of 1.2 mL. This solution was then added to HUVEC monolayers prewashed with Optimem®, then incubated for 4 h at 37℃,

after which the HUVEC media was changed to EBM-2. The cells were used 24 h after transfection.

2.5. Western Blot

Western blot analysis was performed using lysates prepared from HUVEC, as described previously [7]. An equal amount of protein (10 µg) from each condition was subjected to 5-20% SDS-PAGE. I mmunoreactive proteins were detected using an enhanced chemiluminescence (ECL) kit (Amersham Bioscience).

2.6. Statistical Analysis

Results are presented as mean ± standard deviation as indicated. Two-tailed *Student's t tests* were performed using Microsoft Excel to analyse data. Probability values represent the results of these *t tests* with a value of $p < 0.05$ considered statistically significant.

3. RESULTS

3.1. Activation of HUVEC with IL-1β

First, we tried an adhesion assay using monocytic HL60 cells over an activated HUVEC monolayer to validate our assay system. As shown in **Figure 1**, we found a time-dependent increase of HL60 rolling and adhesion when stimulated with IL-1β (10 U/mL), with a peak activation after 4 h stimulation. 2 h IL-1β stimulation led to a dramatic increase in adhesion (13.3 ± 9.92 cells/HPF, $p < 0.0005$) and rolling (4.8 ± 3.34, $p < 0.0005$) of HL60 cells when compared to those HUVECs without activation (adhesion, 1.8 ± 1.78; rolling, 0 ± 0). 4 h IL-1β stimulation led to a further increase in adhesion (25.8 ± 12.1 cells/HPF, $p < 0.00001$) and rolling (3 ± 2.83, $p < 0.03$) of HL60 cells. IL-1β stimulation for 6 h caused an increase in adhesion of HL60 cells greater than that with 2 h IL-1β stimulation but not as significant as that with 4 h IL-1β stimulation, when compared to the control sample (6 h, 16.5 ± 4.90 cells vs 0 h, 1.8 ± 1.78 cells, $p = 5.62 \times 10^{-8}$). We then checked the surface expression of adhesion molecules after IL-1β activation using fluorescent immunobinding assay (FIA), as shown in **Figure 2**. To check the validity of the FIA analysis, intensity of negative controls (no primary antibody) and positive controls (W6/32, anti-HLA Class I antibody) were examined (negative control: 0 h, 4.8 ± 3.19 RFU vs 4 h, 5.6 ± 3.29 RFU, $p = 0.71$; positive control: 0 h,17 ± 2 RFU vs 4h, 21.2 ± 2.59 RFU, $p = 0.02$). E-selectin expression (detected using H18/7 Ab) was significantly upregulated after 4 h IL-1β stimulation compared to no stimulation (4 h, 25.2 ± 1.79 RFU vs 0 h, 9 ± 2.92 RFU, $p = 5.52 \times 10^{-6}$). ICAM-1 expression (detected using Hu5/3 Ab) was also significantly upregulated after 4 h IL-1β stimulation (4 h, 23.4 ± 3.05 RFU vs 0 h, 9.2 ± 0.84 RFU, $p = 8.23 \times 10^{-6}$).

3.2. Adhesion of EPC to Activated HUVEC under Flow

Based on the preliminary experiment using HL-60 cells

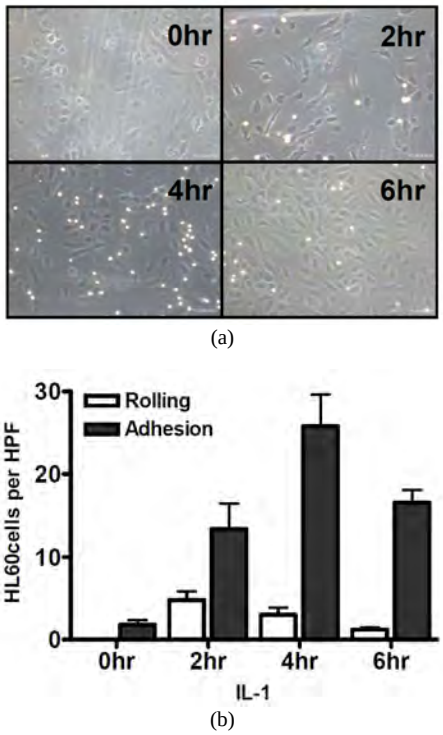


Figure 1. IL-1β induces a time-dependent rolling and adhesion of HL60 cells to HUVEC. (a) Representative micrographs taken from video recording during the flow assay. HL60 cells (white particles) were perfused over a HUVEC monolayer (grey background) that had been subjected to activation by IL-1β (50 U/mL) for four different time periods; (b) The number of rolling and adherent HL60 cells were measured as described in Methods. (* = $p < 0.05$ vs 0 h IL-1β); data based on 10 fields of observation for each condition.

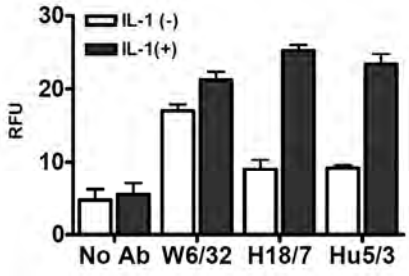


Figure 2. Cell surface expression of E-selectin and ICAM-1 in HUVEC after IL-1β stimulation. The expression intensity of each of the adhesion molecules was determined using a fluorescent immunobinding assay as described in Methods. (* = $p < 0.05$ vs corresponding 0 h IL-1β); data based on 5 repeats for each condition.

(**Figure 1**), we used a condition of IL-1β (10 U/mL) for 4 h in our following experiments. When EPCs, cultured ex vivo for 7 days, were perfused over activated HUVEC monolayer, adhesion of EPCs was dramatically increased compared to the control HUVEC monolayer (4 h, 4.4 ± 1.50 cells vs 0 h, 1.3 ± 0.9 cells, $p = 8.97 \times 10^{-5}$), as shown in **Figure 3**. There was however no rolling of EPCs observed both in the control and 4 h IL-1β stimulation.

3.3. Gene Silencing of E-Selectin in HUVEC

Having demonstrated that 4 h of IL-1β stimulation was optimal for adhesion, we looked at the effects of silencing E-selectin, using siRNA transfection, on the rolling and adhesion of EPCs. siRNA transfection on HUVEC monolayers was carried out 24 h prior to the adhesion assay under flow conditions (the time point was determined to optimize the effect of siRNAs in HUVECs, as previously described [8]). We utilized FIA assay to demonstrate the level of E-selectin in these cells. As demonstrated in **Figure 4**, positive controls (W6/32, anti-HLA Class I antibody) exhibited no significant difference between siRNA treatment groups, enhancing the validity of these results. However, there was a significant reduction of E-selectin surface expression in the E-selectin siRNA—treated HUVECs when compared to the *LacZ* siRNA—treated HUVECs after 4 h IL-1β stimulation (E-selectin, 27.3 ± 5.40 RFU vs *LacZ*, 45.5 ± 7.37 RFU, p = 0.013) or no siRNA group (E-selectin, 27.3 ± 5.40 RFU vs no siRNA, 52.5 ± 10.4 RFU, p = 0.010).

3.4. Effect of Gene Silencing of E-Selectin on EPC Adhesion to HUVEC under Flow

Figure 5 showed results of the flow assay using IL-1-activated HUVECs after siRNA transfection. There was a significant decrease in EPC adhesion in HUVECs transfected with E-selectin siRNA compared to those treated with *LacZ* siRNA (controls) (E-selectin siRNA, 2.7 ± 0.64 cells vs *LacZ* siRNA, 5.1 ± 1.22 cells, $p = 5.74 \times 10^{-5}$). In comparison, there was no significant difference in rolling between those HUVECs transfected with E-selectin siRNA and the control HUVECs (E-selectin siRNA, 0.2 ± 0.4 cells vs *LacZ* siRNA, 0.5 ± 0.67 cells, p = 0.26). There was also a significant decrease in EPC adhesion in HUVECs treated with E-selectin siRNA compared to those treated with no siRNA (Lipofectin® only) (E-selectin siRNA 2.7 ± 0.64 cells vs no siRNA, 7 ± 1.26 cells, $p = 3.74 \times 10^{-8}$).

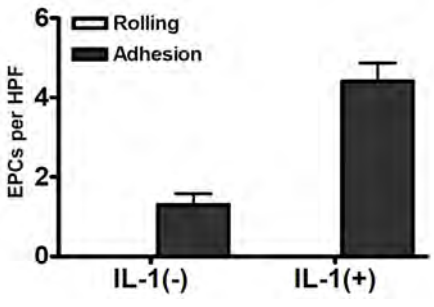


Figure 3. The number of rolling and adherent EPC to activated HUVEC were measured as described in Methods. (* = p < 0.05 vs 0 h IL-1β); data based on 10 fields of observation for each condition. EPCs were perfused over a HUVEC monolayer (grey background) that had been subjected to activation by IL-1β (50 U/mL) for two different time periods (0 h IL-1β and 4 h IL-1β).

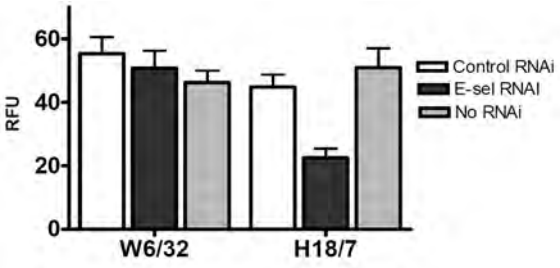


Figure 4. Cell surface expression of E-selectin (H18/7) and positive control (W6/32) in HUVEC after siRNA silencing. The expression intensity of each of the adhesion molecules was determined using a fluorescent immunobinding assay as described in Methods. (* = p < 0.05 vs corresponding 0hr IL-1β); data based on 5 repeats for each condition.

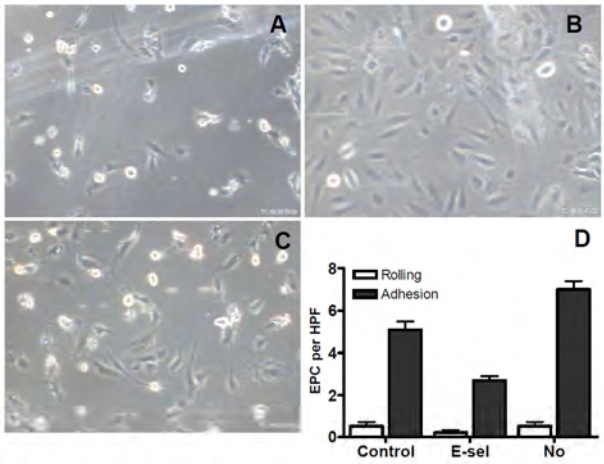


Figure 5. The number of rolling and adherent EPC to activated HUVEC after siRNA transduction. Graph showing the effect of E-selectin siRNA transfection on rolling and adhesion of EPCs compared to control siRNA transfection and no siRNA treatment. (* = p < 0.05 vs control); data based on 10 fields of observation for each condition.

4. DISCUSSION

In this project, we have demonstrated that 1) 4 h of in-

cubation with IL-1β induces adhesion of HL60 cells to HUVEC under flow; 2) 4 h of incubation with IL-1β induces expression of E-selectin and ICAM-1 in HUVEC; 3) 4 h of incubation with IL-1β induces adhesion of EPCs to HUVEC under flow; 4) siRNA against E-selectin reduces endogenous expression of E-selectin in HUVEC; 5) gene silencing of E-selectin resulted in reduction of EPC adhesion to HUVEC. These lines of findings point a predominant role for this molecule in neovascularisation. In previous work from our laboratory using siRNA transfection carried out by Nishiwaki *et al* (2003), we successfully silenced endogenous E-selectin expression in vascular endothelium using siRNA transfection, and demonstrated that this inhibition caused a significant reduction in leukocyte adhesion [8]. Current results compliment this work and emphasise the efficacy of siRNA in the field of vascular biology.

We have also reinforced the work of Nishiwaki *et al* (using antibodies to downregulate E-selectin) and Vajkoczy *et al* (using an *in vivo* method to silence E-selectin expression) in demonstrating that E-selectin is important in EPC adhesion by using a different method (siRNA transfection) [6,7].

Using siRNA transfection to inhibit E-selectin expression, as demonstrated by our experiments, may have implications for its role in angiogenesis. Bischoff has reported that E-selectin is involved in angiogenesis through a series of experiments [17]. One of these experiments showed that antibodies directed against E-selectin or sialylated fucosylated oligosaccharides (structures to which E-selectin binds) inhibit the formation of capillary-like tubes *in vitro* [18]. The presence of E-selectin in human infantile haemangioma tumours (endothelial cell tumour of capillary blood vessels), as well as in tissues with on going growth of microvessels (such as human placenta and neonatal foreskin) has been shown *in vivo* [19]. Yu *et al* had also shown that endothelial progenitor cells are present in infantile haemangioma; 11 of 12 proliferating haemangioma specimens contained EPCs (determined by flow cytometry using CD133 and CD34 as markers for EPCs) [20]. This raises the possibility of using our method of silencing E-selectin expression in treating tumour angiogenesis, by preventing localised EPC adhesion to the endothelium, thus inhibiting angiogenesis in the tumour, reducing its growth. As described previously, we found that there was no significant reduction in rolling after E-selectin down-regulation. This could have been because of the low level of EPC rolling in general, as demonstrated in **Figure 3**. On the other hand, the binding characteristics of E-selectin (E-selectin has a greater role in adhesion than rolling) may have some effects. This speculation has been supported by the results from Mazo, *et al.*'s experiments [21]. They demonstrated that blocking E-selectin with antibodies resulted in a reduction of haematopoietic progenitor cells (HPC) rolling of 32% compared to a reduction in HPC rolling of 58% when P-selectin was inhibited (measured using intravital microscopy). Therefore the other selectins (P- and L-selectins) are probably more important for HPC rolling.

A limitation of this experiment is that we have only demonstrated that gene silencing of E-selectin expression reduces EPC adhesion *in vitro*, but not *in vivo*. Therefore future work could be aimed at seeing if siRNA transfection has similar effects *in vivo*. It would also be interesting to try similar experiments using siRNA transfection against other adhesion molecules such as ICAM-1 or P-selectin, and look at the effects on EPC recruitment.

In conclusion, we have demonstrated that E-selectin—short interfering RNA inhibits endothelial progenitor cell recruitment to vascular endothelium under flow conditions.

5. ACKNOWLEDGEMENTS

Grateful thanks to Dr Naokazu Nakamura, Michiyo Deushi, Noriko Nitta and Mariko Tani for their experimental help and advice, Professor Peter Sugden for his help and advice throughout the course of the project.

REFERENCES

[1] Cines, D.B., Pollak, E.S., Buck, C.A., Loscalzo, J, Zimmerman, G.A., McEver, R.P., Pober, J.S., Wick, T.M., Konkle, B.A., Schwartz, B.S., Barnathan, E.S., McCrae, K.R., Hug, B.A., Schmidt, A.M. and Stern, D.M. (1998) Endothelial cells in physiology and in the pathophysiology of vascular disorders. *Blood*, **91(10)**, 3527-3561.

[2] Shi, Q., Rafii, S., Wu, M.H., Wijelath, E.S., Yu, C., Ishida, A., Fujita, Y., Kothari, S., Mohle, R., Sauvage, L.R., Moore, M.A., Storb, R.F. and Hammond, W.P. (1998) Evidence for circulating bone marrow-derived endothelial cells. *Blood*, **92(2)**, 362-367.

[3] Kakkar, A.K. and Lefer, D.J. (2004) Leukocyte and endothelial adhesion molecule studies in knockout mice. *Current Opinion in Pharmacology*, **4(2)**, 154-158.

[4] Muller, W.A. (2003) Leukocyte—endothelial-cell interactions in leukocyte transmigration and the inflammatory response. *Trends in Immunology*, **24(6)**, 326-333.

[5] Chavakis, E., Aicher, A., Heeschen, C., Sasaki, K., Kaiser, R., El Makhfi, N., Urbich, C., Peters, T., Scharffetter-Kochanek, K., Zeiher, A.M., Chavakis, T. and Dimmeler, S. (2005) Role of beta2-integrins for homing and neovascularization capacity of endothelial progenitor cells. *Journal of Experimental Medicine*, **201(1)**, 63-72.

[6] Vajkoczy, P., Blum, S., Lamparter, M., Mailhammer, R., Erber, R., Engelhardt, B., Vestweber, D. and Hatzopoulos, A.K. (2003) Multistep nature of microvascular recruitment of ex vivo-expanded embryonic endothelial progenitor cells during tumor angiogenesis. *Journal of Ex-*

perimental Medicine, **197(12)**, 1755-1765.

[7] Nishiwaki, Y., Yoshida, M., Masuda, H. and Isobe, M. (2004) Recruitment of bone marrow-derived endothelial progenitor cells to vascular endothelium involves E-selectin dependent mechanism. *Cardiovascular Pathology*, **13(3)**, 171.

[8] Nishiwaki, Y., Yokota, T., Hiraoka, M., Miyagishi, M., Taira, K., Isobe, M., Mizusawa, H. and Yoshida, M. (2003) Introduction of short interfering RNA to silence endogenous E-selectin in vascular endothelium leads to successful inhibition of leukocyte adhesion. *Biochemical and Biophysical Research Communications*, **310(4)**, 1062-1066.

[9] Yoshida, M. and Gimbrone, M.A. Jr. (1997) Novel roles for E-selectin in endothelial-leukocyte adhesion. *Annals of the New York Academic Sciences*, **811**, 493-497.

[10] Yoshida, M., Szente, B.E., Kiely, J.M., Rosenzweig, A. and Gimbrone, M. Jr. (1998) Phosphorylation of the cytoplasmic domain of E-selectin is regulated during leukocyte-endothelial adhesion. *The Journal of Immunology*, **161(2)**, 933-941.

[11] Bevilacqua, M.P., Pober, J.S., Mendrick, D.L., Cotran, R.S. and Gimbrone, M.A. (1987) Identification of an inducible endothelial-leukocyte adhesion molecule. *Proceedings of the National Academy Sciences of USA*, **84(24)**, 9238-9242.

[12] Luscinskas, F.W., Cybulsky, M.I., Kiely, J.M., Peckins, C.S., Davis, V.M. and Gimbrone, M.A. (1991) Cytokine-activated human endothelial monolayers support enhanced neutrophil transmigration via a mechanism involving both endothelial-leukocyte adhesion molecule-1 and intercellular adhesion molecule-1. *The Journal of Immunology*, **146(5)**, 1617-1625.

[13] Gerszten, R.E., Garcia-Zepeda, E.A., Lim, Y.C., Yoshida, M., Ding, H.A., Gimbrone, M. Jr., Luster, A.D., Luscinskas, F.W. and Rosenzweig, A. (1999) MCP-1 and IL-8 trigger firm adhesion of monocytes to vascular endothelium under flow conditions. *Nature*, **398(6729)**, 718-723.

[14] Kiely, J.M., Luscinkas, F.W. and Gimbrone, M. Jr. (1999) Leukocyte-endothelial monolayer adhesion assay (static conditions). *Methods in Molecular Biology*, **96**, 131-136.

[15] Gerszten, R.E., Luscinskas, F.W., Ding, H.T., Dichek, D.A., Stoolman, L.M., Gimbrone, M.A. Jr. and Rosenzweig, A. (1996) Adhesion of memory lymphocytes to vascular cell adhesion molecule-1—transduced human vascular endothelial cells under simulated physiological flow conditions in vitro. *Circulation Research*, **79(6)**, 1205-1215.

[16] Luscinskas, F.W., Kansas, G.S., Ding, H., Pizcueta, P., Schleiffenbaum, B.E., Tedder, T.F. and Gimbrone, M. Jr. (1994) Monocyte rolling, arrest and spreading on IL-4-activated vascular endothelium under flow is mediated via sequential action of L-selectin, β1-integrins, and β2-integrins. *Journal of Cell Biology*, **125(6)**, 1417-1427.

[17] Bischoff, J. (1997) Cell adhesion and angiogenesis. *Journal of Clinical Investigation*, **99(3)**, 373-376.

[18] Nguyen, M., Strubel, N.A. and Bischoff, J. (1993) A role for sialyl Lewis-X/A glycoconjugates in capillary morphogenesis. *Nature*, **365(6443)**, 267-269.

[19] Kraling, B.M., Razon, M.J., Boon, L.M., Zurakowski, D., Seachord, C., Darveau, R.P., Mulliken, J.B., Corless, C.L. and Bischoff, J. (1996) E-selectin is present in proliferating endothelial cells in human hemangiomas. *American Journal of Pathology*, **148(4)**, 1181-1191.

[20] Yu, Y., Flint, A.F., Mulliken, J.B., Wu, J.K. and Bischoff, J. (2004) Endothelial progenitor cells in infantile hemangioma. *Blood*, **103(4)**, 1373-1375.

[21] Mazo, I.B., Gutierrez-Ramos, J.C., Frenette, P.S., Hynes, R.O., Wagner, D.D. and von Andrian, U.H. (1998) Hematopoietic progenitor cell rolling in bone marrow microvessels: parallel contributions by endothelial selectins and vascular cell adhesion molecule 1. *Journal of Experimental Medicine*, **188(3)**, 465-474.

J. Biomedical Science and Engineering, 2010, 3, 556-567 **JBiSE**
doi:10.4236/jbise.2010.36078 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine

Yuedong Song, Pietro Liò

Computer Laboratory, University of Cambridge, Cambridge, United Kingdom.
Email: ys340@cam.ac.uk; Pietro.Lio@cl.cam.ac.uk

Received 4 March 2010; revised 15 April 2010; accepted 18 April 2010.

ABSTRACT

The electroencephalogram (EEG) signal plays a key role in the diagnosis of epilepsy. Substantial data is generated by the EEG recordings of ambulatory recording systems, and detection of epileptic activity requires a time-consuming analysis of the complete length of the EEG time series data by a neurology expert. A variety of automatic epilepsy detection systems have been developed during the last ten years. In this paper, we investigate the potential of a recently-proposed statistical measure parameter regarded as Sample Entropy (*SampEn*), as a method of feature extraction to the task of classifying three different kinds of EEG signals (normal, interictal and ictal) and detecting epileptic seizures. It is known that the value of the *SampEn* falls suddenly during an epileptic seizure and this fact is utilized in the proposed diagnosis system. Two different kinds of classification models, back-propagation neural network (BPNN) and the recently-developed extreme learning machine (ELM) are tested in this study. Results show that the proposed automatic epilepsy detection system which uses sample entropy (*SampEn*) as the only input feature, together with extreme learning machine (ELM) classification model, not only achieves high classification accuracy (95.67%) but also very fast speed.

Keywords: Epileptic Seizure; Electroencephalogram (EEG); Sample Entropy (*SampEn*); Backpropagation Neural Network (BPNN); Extreme Learning Machine (ELM); Detection

1. INTRODUCTION

Epilepsy, the second most common serious neurological disorder in human beings after stroke, is a chronic condition of the nervous system and it is characterized by recurrent unprovoked seizures. Approximately one in every 100 individuals worldwide are suffering from epilepsy [1]. Electroencephalography (EEG) is an important clinical tool, monitoring, diagnosing and managing neurological disorders related to epilepsy. In comparison with other methods such as Electrocorticogram (ECOG), EEG is a clean and safe technique for monitoring the brain activity.

In spite of available dietary, drug and surgical treatment options, currently nearly one out of three epilepsy patients cannot be treated. They are completely subject to the sudden and unforeseen seizures which have a great effect on their daily life, with temporary impairments of perception, speech, motor control, memory and/or consciousness. Many new therapies are being investigated and among them the most promising are implantable devices that deliver direct electrical stimulation to affected areas of the brain. These treatments will greatly depend on robust algorithms for seizure detection to perform effectively. Because the onset of the seizures cannot be predicted in a short period, a continuous recording of the EEG is required to detect epilepsy. However, analysis by visual inspection of long recordings of EEG, in order to find traces of epilepsy, is tedious, time-consuming and high-cost. Therefore, automated detection of epilepsy has been a goal of many researchers for a long time. With the advent of technology, the digital EEG data can be input to an automated seizure detection system, allowing physicians to treat more patients in a given time because the time taken to review the EEG data is greatly reduced by automation.

In recent years, there has been an increasing interest in the application of pattern recognition (PR) methods for automatic epileptic seizure detection. Several methods have been developed for handling EEG signals classification, and among these methods, Multi-layer Perceptron Neural Network (MLPNN) [2-7] and Support Vector Machine (SVM) [8-10] are two widely-used classification paradigms. Most of the automatic epileptic seizure detection system is built by time-frequency domain based feature extraction followed by a variety of classification models. It has been found that the classification

performance of these automatic detection systems totally depends on the feature extraction of the EEG time series [6,11,12].

As an effective tool for detection and characterization of signals, deterministic chaos plays a key role. Many chaos-producing mechanisms have been created and applied for recognizing the behaviour of the dynamics of the system. The physiological time-series signals are considered chaotic. Recently studies on the basis of measuring entropies have been employed for biomedical studies [13]. The randomness of non-linear time series data is well embodied by calculating entropies of the time series data and it can supply recognizable variation for normal and abnormal physiological signals. Entropy is a measure of uncertainty. The level of chaos can be measured by applying entropy of the system. Higher entropy stands for higher uncertainty and a more chaotic system. X. L. Li, *et al.* [14] investigated permutation entropy as a tool to predict the absence seizures of genetic absence epilepsy rats by applying EEG recordings. H. Ocak [15] presented a new scheme on the basis of approximate entropy and discrete wavelet transform to detect epileptic seizure from EEG time series data that was recorded from normal subjects and epileptic patients. K. S. Pravin, *et al.* [16] had shown some initial investigations on wavelet entropy for epileptic seizure detection.

In this study, we proposed a new method for epileptic seizure detection by using feature extraction based on sample entropy (*SampEn*) followed by two non-linear classification models, namely, back-propagation neural network (BPNN) and extreme learning machine (ELM) which is a recently-proposed classification model [17]. The proposed scheme was tested using clinical electroencephalogram (EEG) signals obtained from five healthy subjects and five epileptic patients during both interictal and ictal periods. The results showed that the proposed scheme (*SampEn* + ELM) was capable of detecting epileptic seizures not only with a high accuracy but also with a very fast speed, which demonstrates its potential for real-time implementation in an automated epilepsy

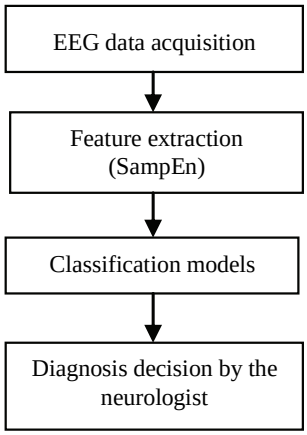
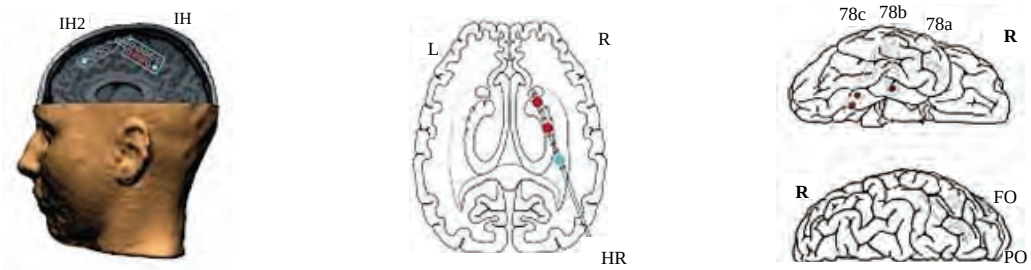


Figure 1. Schematics of the proposed diagnostic expert system.

detection and diagnosis support systems. Up to now, to the best of our knowledge, there is no study in the literature related to the assessment of classification performance using sample entropy based feature extraction followed by ELM classification model when applied specifically to the normal/interictal/ictal discrimination problem. **Figure 1** shows the schematics of the proposed diagnosis expert system.

2. DATA ANALYZED

In this study, a publicly-available database introduced in [18] has been used. The EEG data considered in this work is composed of three different sets, each containing 100 single-channel EEG data of 23.6 s duration. The first data set includes surface EEG recordings that were acquired from five healthy volunteers using a standardized electrode placement scheme. The subjects were awake and relaxed with their eyes open (normal periods). The data set for the last two sets was taken from five epileptic patients experiencing pre-surgical diagnosis. The second data set was composed of intracranial EEG recordings during seizure-free intervals (interictal periods) from within the epileptogenic zone of the brain. The EEG signals in the third data set were recorded during

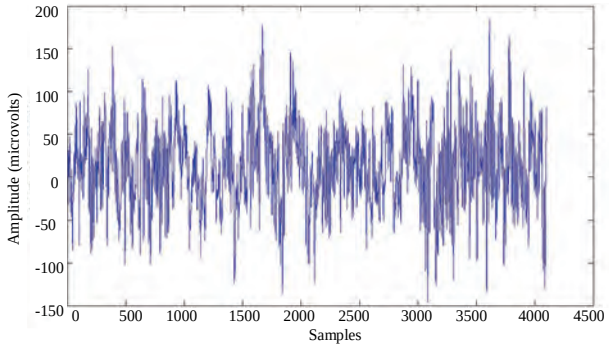


(a) (b) (c)

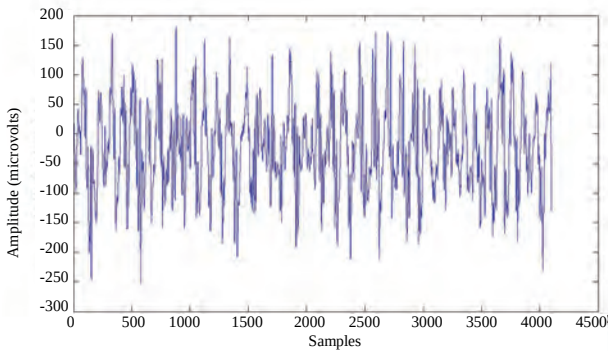
Figure 2. Intracranial electrode placements.

Table 1. Summary of the clinical data.

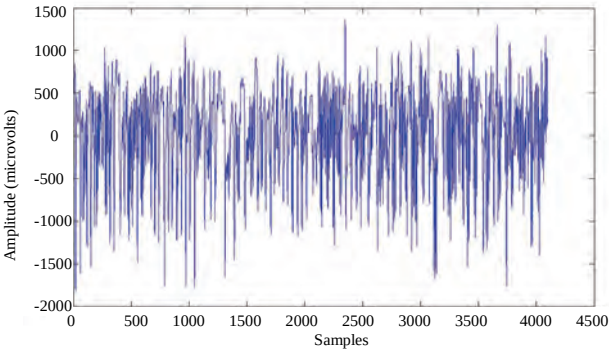
	Data set 1	Data set 2	Data set 3
Subjects	Five healthy subjects	Five epileptic patients	Five epileptic patients
Electrode type	Surface	Intracranial	Intracranial
Electrode placement	International 10-20 system	Within epileptogenic zone	Within epileptogenic zone
Patient's state	Awake and eyes open (Normal)	Seizure-free (Interictal)	Seizure activity (Ictal)
Number of epochs	100	100	100
Epoch duration (s)	23.6	23.6	23.6



(a)



(b)



(c)

Figure 3. Sample EEG recordings. (a) Normal EEG; (b) Interictal EEG; (c) Ictal EEG.

seizure activity (ictal periods) using depth electrodes placed within the epileptogenic zone of the brain. All EEG signals were recorded with the same 128-channel amplifier. The data were digitized at 173.6 samples per second at 12-bit resolution. Band pass filter was set to 0.53-60 Hz. The total number of EEG signals is 300 (100 normal signals, 100 interictal signals and 100 ictal signals). Each data set has 4096 sampling points. **Figure 2** describes the electrode placement for recording of EEG signals. A summary of the data set is given in **Table 1**. **Figure 3** describes example of EEG signals of each of the three data sets.

3. THEORY AND MODEL

3.1. Sample Entropy

Entropy is a concept handling predictability and randomness, with higher values of entropy always related to less system order and more randomness. In recent years, a variety of estimators have been proposed to quantify the entropy of time series. These methods can be roughly divided into two categories, embedding entropy and spectral entropy [19]. Embedding entropy supplies information regarding how EEG time series signals change with time, by comparing each time series signal with a lagged form of itself [20]. In [13], a new family of statistics called Sample Entropy (*SampEn*) was introduced and characterized. This measure is embedding entropy quantifying the complexity in time series data without the weaknesses that widely utilized non-linear approaches have. The *SampEn* is less sensitive to noise and can be applied for short-length time series data [13]. Additionally, it is resistant to short strong transient interferences (outliers) such as spikes. These characteristics make Sample Entropy an appealing tool for nonlinear analysis of physiological signals.

In spite of its advantages over other non-linear estimators, the *SampEn* is not widely used. In [20], sample entropy was used to analyze the electroencephalogram background activity of Alzheimer's disease patients for testing the hypothesis that the regularity of their EEGs is higher than that of age-matched controls. M. Aboy [21] conducted a characterization study of *SampEn* for supplying additional insights about the interpretation of this complexity metric in the context of biomedical signal

analysis. Moreover, this entropy measure has been used to evaluate the signal complexity of the cyclic behaviour of heart rate variability (HRV) in obstructive sleep apnea syndrome [22]. In this study, *SampEn* is investigated for the first time as a feature extracted in the automatic detection of epilepsy.

For calculating the *SampEn*, the embedding dimension (m) and vector comparison distance (r) must be specified. It is common to set the embedding dimension parameter m to be m = 1, 2 or 3 and to set the vector comparison distance r to be some percentage of the standard deviation of the time series so as not to depend on the absolute amplitude of the signal [13]. *SampEn*(m,r,N) is the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability. Thus, a larger value often corresponds to more irregularity or complexity in the time series data. In the proposed automated epileptic seizure detection system, the value of the *SampEn* is determined as shown in the following steps:

1) Given N data points from a time series $\{x(n)\}$ = $x(1)$, $x(2)$,...,$x(N)$, take m vectors $X_m(1),...,X_m(N-m+1)$ defined as $X_m(i)=[x(i),x(i+1),...,x(i+m-1)]$, for $1\le i\le N-m+1$. These vectors stand for m consecutive x values, starting at the ith sample.

2) Let r denote the noise filter level which is defined as

$$r = g \times Std \quad for \quad g = 0.1, 0.2, ..., 0.5 \tag{1}$$

where *Std* represents the standard deviation of the data sequence X.

3) The distance between vectors $X_m(i)$ and $X_m(j)$, $d[X_m(i), X_m(j)]$, is defined as the maximum absolute difference between their scalar components:

$$d[X_m(i), X_m(j)] = \max_{k=0,...,m-1}(|x(i+k)-x(j+k)|). \tag{2}$$

4) For a given $X_m(i)$, count the number of j ($1 \le j \le N-m, j \ne i$), such that $d[X_m(i), X_m(j)] \le r$. This number is represented as B_i, Then, for $1 \le i \le N-m$,

$$B_i^m(r) = \frac{1}{N-m-1} B_i \tag{3}$$

Here, note that only the first $N-m$ vectors of length m are considered in order to ensure that for $1 \le i \le N-m$, the vector $X_{m+1}(i)$ is also defined.

5) Define $B^m(r)$ as

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r). \tag{4}$$

6) We increment the dimension to m + 1 and compute A_i as the number of $X_{m+1}(i)$ within r of $X_{m+1}(j)$, where j ranges from 1 to $N-m$ ($j \ne i$). We then define $A_i^m(r)$ as

$$A_i^m(r) = \frac{1}{N-m-1} A_i. \tag{5}$$

7) We define $A^m(r)$ as

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r). \tag{6}$$

Thus, $B^m(r)$ represents the probability that two sequences will match for m points, whereas $A^m(r)$ represents the probability that two sequences will match for m + 1 points.

The sample entropy is defined by

$$SampEn(m,r) = \lim_{N\to\infty}\{-\ln[\frac{A^m(r)}{B^m(r)}]\} \tag{7}$$

Since the time series length is finite, *SampEn* is estimated as

$$SampEn(m,r,N) = \ln[\frac{B^m(r)}{A^m(r)}] \tag{8}$$

3.2. An Example of the Computation Procedure of *SampEn*

Assume that the sequence X_N is composed of 50 sampling points (*i.e.*, N = 50).

X_N = {51,52,53,54,55, 51,52,53,54,55,51,...,55}

The total sequence is periodic of 5. Let us choose m=5 and r = 2, so we have:

$X_5(1) = \{51, 52, 53, 54, 55\}$

$X_5(2) = \{52, 53, 54, 55, 51\}$

$X_5(3) = \{53, 54, 55, 51, 52\}$

$X_5(4) = \{54, 55, 51, 52, 53\}$

$X_5(5) = \{55, 51, 52, 53, 54\}$

and so on. Firstly we want to find the number of $X_5(i)$ which is similar to $X_5(1)$. Because we have chosen r = 2 as the threshold parameter, which means each of the five elements of $X_5(i)$ has to be within ±2 units of the corresponding element of $X_5(1)$. For instance, $X_5(2)$ is not similar to $X_5(1)$ because the last elements in these two sequences (51,55) differ by more than two units. The conditions of similarity to $X_5(1)$ are satisfied only by $X_5(6)$, $X_5(11)$, $X_5(16)$, $X_5(21)$,..., $X_5(41)$ (excludes $X_5(1)$ since $j \ne i$ and also excludes $X_5(46)$ since only the first 45 elements of the $X_5(i)$ are considered in terms of the definition of *SampEn*). Thus, we get B_1 = 8. Because the whole number of $X_5(i)$ is $N-m-1=50-5-1=44$, so we have

$$B_1^5(2) = \frac{B_1}{N-m-1} = \frac{8}{44}.$$

The above steps are repeated for determining the number of $X_5(i)$ which are similar to $X_5(2)$, $X_5(3)$ and so on. By using the same inference, $X_5(2)$ is similar to $X_5(7)$, $X_5(12)$, $X_5(17)$,..., $X_5(42)$. Thus, we also get $B_2 = 8$. Generally, in this example, we have $B_i = 8$ for $1 \le i \le N-m$. Therefore, $B_i^5(2)$ is 8/44 and we can get the mean value of all 45 of B_i^5:

$$B^5(2) = \frac{45 \cdot \frac{8}{44}}{45} = \frac{8}{44} \approx 0.1818$$

In order to get *SampEn*(5,2,*N*), the above-mentioned computation procedure needs to be repeated for m = 6. Doing so, we get

$$X_6(1) = \{51,52,53,54,55,51\}$$

$$X_6(2) = \{52,53,54,55,51,52\}$$

$$X_6(3) = \{53,54,55,51,52,53\}$$

and so on. By the same reasoning steps as that with m = 5, we find that:

For $1 \le i \le N-m$, $\mathrm{A}_i = \begin{cases} 7 \text{ if i=5 modulo 5} \\ 8 \text{ otherwise} \end{cases}$

Therefore, $A_i^6(2)$ is either 7/43 or 8/43, depending on A_i, and the mean value of all 44 elements of the $A_i^6(2)$ is:

$$A^6(2) = \frac{8 \cdot \frac{7}{43} + 36 \cdot \frac{8}{43}}{44} = \frac{8}{44} \approx 0.1818$$

Finally, we compute the value of *SampEn* as follows:

$$SampEn(5,2,N) = \ln[\frac{B^5(2)}{A^6(2)}] = \ln 1 = 0$$

This is the smallest value of *SampEn*, which indicates that the original time series data is highly regular and predictable.

3.3. Levenberg-Marquardt Algorithm

Artificial neural network training is often regarded as a nonlinear least-squares problem and the Levenberg-Marquardt algorithm is a least-squares estimation algorithm utilizing the maximum neighbourhood idea, and it appears to be the fastest method for training feed- forward neural networks. Let $E(w)$ be an objective error function composed of n individual error terms $e_j^2(w)$ as follows:

$$E(w) = \sum_{j=1}^{n} e_j^2(w) = \| f(w) \|^2 \quad (9)$$

where $e_j^2(w) = (y_{dj} - y_j)^2$ and y_{dj} is the desired value of output neuron j, y_j is the actual output of the neuron.

The objective of the Levenberg-Marquardt algorithm is to calculate the weight vector $\boldsymbol{w}$ so that $E(w)$ is minimized. By utilizing the LM algorithm, a novel weight vector w_{p+1} can be obtained from the previous weight vector w_p as follows:

$$w_{p+1} = w_p + \delta W_P \quad (10)$$

where s is defined as

$$\delta w_p = -(J_p^T f(w_p))(J_p^T J_p + \gamma I)^{-1} \quad (11)$$

In Equation (10), J_p is the Jacobian of f assessed at w_p, γ is the Marquardt parameter, and I is the identity matrix. The Levenberg-Marquardt algorithm can be described as follows: 1) Calculate $E(w_p)$. 2) Begin with a small value of γ ($\gamma = 0.01$). 3) Solve equation (11) for δw_p and calculate $E(w_p + \delta w_p)$. 4) If $E(w_p + \delta w_p)$, $\ge E(w_p)$ increment γ by a factor of 10 and go to step 3. 5) If $E(w_p + \delta w_p) < E(w_p)$, decrement γ by a factor of 10, update $w_p : w_p \leftarrow w_p + \gamma w_p$ and go to step 3.

3.4. Extreme Learning Machine (ELM)

The general trend in current study of automatic epileptic seizure detection has focused on high accuracy but has not considered the time taken to train the classification models, which should be an important factor of developing an EEG-based detection device for epileptic seizures because the online device will need to update its training during use. Therefore some classification models with high classification accuracy may not be satisfactory when considering the trade-off between the classification accuracy and the time for training the classification models. In our study, in addition to exploring the potential of a nonlinear feature of the EEG signal called sample entropy for electroencephalogram time series classification and epileptic seizure detection, we also investigate the use of a novel paradigm of learning machine called Extreme Learning Machine (ELM) [17], in order to obtain a balance between high classification accuracy and short training time. In recent years, Extreme Learning Machine has been increasingly popular in classification tasks due to its high generalization ability and fast learning speed. In [23], a classification system is built using ELM to classify protein sequences with ten classes of super-families obtained from a domain database, and its performance is compared with that of Back-propagation Neural Networks. The results show ELM greatly outperforms BPNN in terms of both training time and classification accuracy. R. Zhang, *et al.* [24] developed an ELM for multi-category classification in three Cancer Microarray Gene Expression datasets,

and the results reveal that ELM can not only obtain high classification accuracy but also avoid problems such as over-fitting, local minima, and improper learning rate. In addition to the field of Bioinformatics, Extreme Learning Machine has also been successfully applied to Biosignal Processing. N. Y. Liang, *et al.* [25] proposed an ELM-based classification scheme to classify five mental tasks from different subjects using EEG signals available from a Brain Computer Interfaces (BCIs) database. Performance of ELM is compared with a Back-propagation Neural Network (BPNN) and also Support Vector Machine (SVMs). Experimental results show that ELM needs an order of magnitude less training time compared with SVMs and two orders of magnitude less training time compared with BPNN, and the classification accuracy of ELM is similar to that of SVMs and BPNN. In [26], a classification scheme based on ELM and Principle Component Analysis (PCA) was developed for arrhythmia classification, and finally achieved 97.5% in average accuracy, 97.44% in average sensitivity, 98.46% in average specificity, and 2.423 s in learning time. The idea behind ELM is presented as follows:

Suppose learning N arbitrary different instances (x_i, t_i), where $x_i = [x_{i1}, x_{i2}, ..., x_{in}]^T \in R^n$, and $t_i = [t_{i1}, t_{i2}, ..., t_{im}]^T \in R^m$, standard Single-layer Feedforward Networks with N hidden neurons and activation function $g(x)$ are mathematically modelled as a linear system

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = o_j, \ j = 1, ..., N \tag{12}$$

where $w_i = [w_{i1}, w_{i2}, ..., w_{in}]^T$ denotes the weight vector connecting the ith hidden neuron and the input neuron, $\beta_i = [\beta_{i1}, \beta_{i2}, ..., \beta_{im}]^T$ denotes the weight vector connecting the i-th hidden neuron and output neurons, and b_i represents the threshold of the i-th hidden neuron. $w_i * w_j$ represents the inner product of w_i and x_j. If the Single-layer Feedforward Network with N hidden neurons with activation function $g(x)$ is able to approximate N distinct instances (x_i, t_i) with zero error means that

$$H\beta = T \tag{13}$$

where

$$H(w_1, ..., w_{Nh}, b_1, ..., b_{Nh}, x_1, ..., x_N) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{Nh} \cdot x_1 + b_{Nh}) \\ \vdots & & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_{Nh} \cdot x_N + b_{Nh}) \end{bmatrix}$$

$$w = \begin{bmatrix} w_1^T \\ \cdot \\ w_{Nh}^T \end{bmatrix}_{Nh \times m} \qquad T = \begin{bmatrix} t_1^T \\ \cdot \\ t_N^T \end{bmatrix}_{N \times m}$$

H is the hidden layer output matrix of the SLFN. Hence for fixed arbitrary input weights w_i and the hidden layer bias s, training a Single-layer Feed-forward Network equals to discovering a least-squares solution $\hat{\beta}$ of the linear system $H\beta = T$, $\hat{\beta} = H^{\dagger}T$ is the best weights, where $H^{\dagger}$ is the Moore-Penrose generalized inverse. In terms of [17], Extreme Learning Machine utilizes such Moore-Penrose inverse approach for obtaining good generalization performance with extremely fast learning speed. Unlike some conventional methods, for example Backpropagation algorithm, Extreme Learning Machine is able to avoid problems in tuning control parameters (learning epochs, learning rate, and so on) and keeping to local minima.

The procedure of ELM for single-layer feedforward networks is expressed as follows:

1) Choose arbitrary value for input weights w_i and biases b_i of hidden neurons.

2) Calculate hidden layer output matrix H.

3) Obtain the optimal $\hat{\beta}$ using $\hat{\beta} = H^{\dagger}T$.

Figure 4 shows the structure of ELM.

4. EXPERIMENTS AND RESULTS

4.1. Performance Evaluation Parameters

All the simulations were based on a 2.27 GHz 2-core CPU with 2 GB memory. In order to compare the performance of ELM classifiers, we also implemented a backpropagation neural network (BPNN) based on a Levenberg-Marquardt back-propagation (LMBP) learning algorithm which is thought of as the fastest method for training moderate-sized feed-forward neural networks according to [27]. For the BPNN and ELM, all of the input values were normalized in the range of [–1,1]. The performance of the BPNN and ELM algorithms was

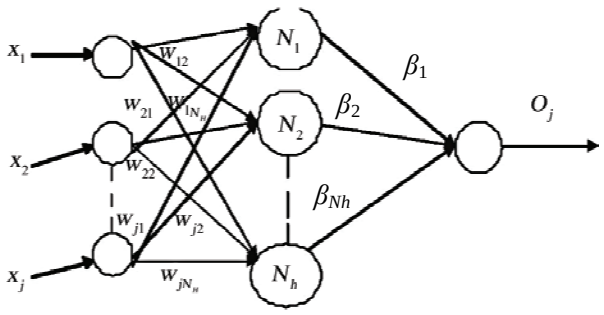


Figure 4. The structure of ELM.

evaluated by the following measures:

1) Learning time: A measure of the time spent in training classification models;

2) Sensitivity (seizure free epileptogenic zone segments): Number of correct classified seizure free epileptogenic zone segments/Number of total seizure free epileptogenic zone segments;

3) Sensitivity (epileptic seizure segments): Number of correct classified epileptic seizure segments/Number of total epileptic seizure segments;

4) Specificity: Number of correct classified healthy segments/Number of total healthy segments;

5) Total classification accuracy: Number of correct classified segments/Number of total segments;

4.2. Training and Testing: 10-Fold Cross-Validation

There are a variety of methods of how to divide the EEG dataset into training and testing datasets. To reduce the bias of training and testing data, a 10-fold cross-validation technique is used. 10-fold cross-validation is a method to improve over the holdout method. This technique will be implemented during the training periods, for estimating how well the classification models that learn from the training data will operate on future data not seen during the testing period. Generally, with 10-fold cross-validation, the data set is divided into 10 subsets, and the holdout approach is re-iterated 10 times. Each time, one of the 10 subsets is utilized as the testing dataset and the other 9 subsets are put together for forming a training dataset. Then the average error across all 10 trials is calculated. According to [28], the result obtained from one 10-fold cross validation may not be dependable. In order to get low mean square error and bias, the 10-fold cross-validation procedure is performed 10 times. All the simulation results were averaged over ten repetitions of 10-fold cross validation.

4.3. Experiment Results and Discussion

Although the pattern length parameter m, the threshold r and the number of sampling points of the time series data play an important role in determining the outcome of *SampEn*, there are no guidelines to set the values of these parameters. In essence, the accuracy and confidence of the entropy estimate improve when the number of matches of length m and m + 1 increases. The number of matches can be increased by choosing small m and large r. However, if r is too large, some fluctuations of the signal are not detected, and if r is too small, noise has effect on the *SampEn* measure [29]. In this study, *SampEn* values are calculated for selected combinations of m, r, and N. The values of m, r, and N that are employed in the experiments are described as follows:

1) m = 1, 2, 3;

2) r = 10%-50% of standard deviation of the EEG data sequence in increases of 10%;

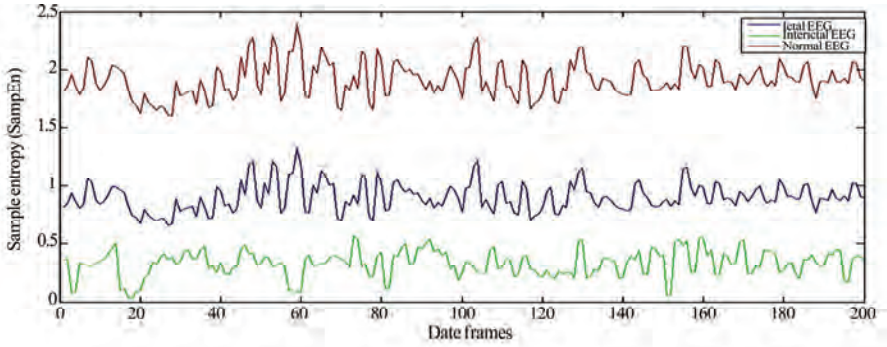
3) N = 256, 512, 1024, 2048, 4096.

Values of *SampEn* are calculated for all normal (healthy segments), interictal (seizure free epileptogenic zone segments) and ictal (epileptic seizure segments) EEG signals, and are fed to two classification models. Using rectangular-window with different sizes, data frames with different sizes (256, 512, 1024, 2048, 4096) are formed and the values of *SampEn* are computed for each data frame. **Figure 5** demonstrates the sample plots of the *SampEn* having clear distinction among normal, interictal and ictal EEG signals. The values of *SampEn* demonstrated in **Figure 5(a)** and **Figure 5(b)** are computed with N = 2048 and 1024, respectively. From **Figure 5**, one can see that the value of *SampEn* is small for ictal EEG signals (between 0.5 and 1.5) compared to normal EEG signals (larger than 1.5). The value of *SampEn* of interictal EEG signals (less than 0.5) is smaller than that of ictal EEG signals. **Figure 6** demonstrates the sample plots of the values of *SampEn* according to N = 1024 and 512 which have a partial overlap among normal, interictal and ictal EEG signals. From these figures, we find that the capability of the *SampEn* for classifying normal, interictal and ictal EEG signals totally depends on the parameter values of m, r, and N.

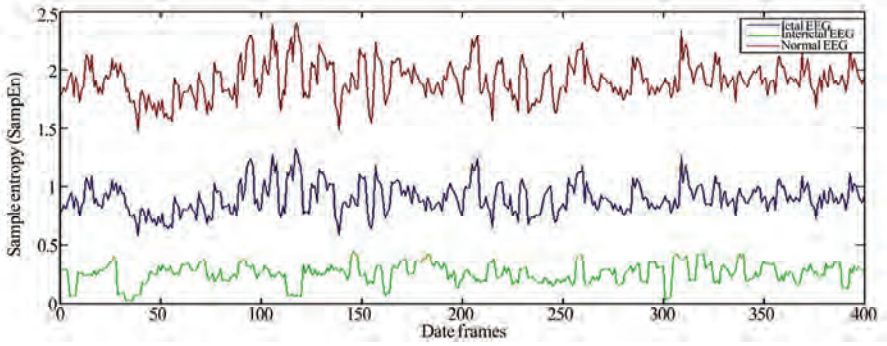
From these figures, it can be noted that utilizing a simple linear discriminator may not achieve good results since *SampEn* demonstrates clear distinction among the normal, interictal and ictal EEG signals only for several particular parameter combinations of m, r and N. For example, a simple linear discriminator would be inefficient for the *SampEn* values, as demonstrated in **Figure 6**, because a clear partial overlapping among the normal, interictal and ictal EEG signals can be seen.

Figures 7-12 demonstrate the whole classification accuracy achieved by neural network and extreme learning machine by employing *SampEn* as the input feature.

It can be observed from **Figures 7-12** that BPNN shows good average accuracy in the range of 94.25%-95.33%, only for several combinations of m, r, and N (for example, m = 2, r = 0.3*Std, N = 2048; m = 3, r = 0.1*Std, N = 1024 and m = 2, r = 0.2*Std, N = 2048). The BPNN achieves the best average accuracy of 95.33% with m = 2, r = 0.2*standard deviation of the time series and N = 2048. For ELM, high average classification accuracy in the range of 94.97%-95.67% are obtained for some combinations of m, r, and N (for example, m = 2, r = 0.1*Std, N = 1024; m = 2, r = 0.2*Std, N = 2048 and m = 3, r = 0.1*Std, N = 1024). The ELM obtains the best average accuracy of 95.67% with m = 3, r = 0.1*standard deviation of the time series and N =

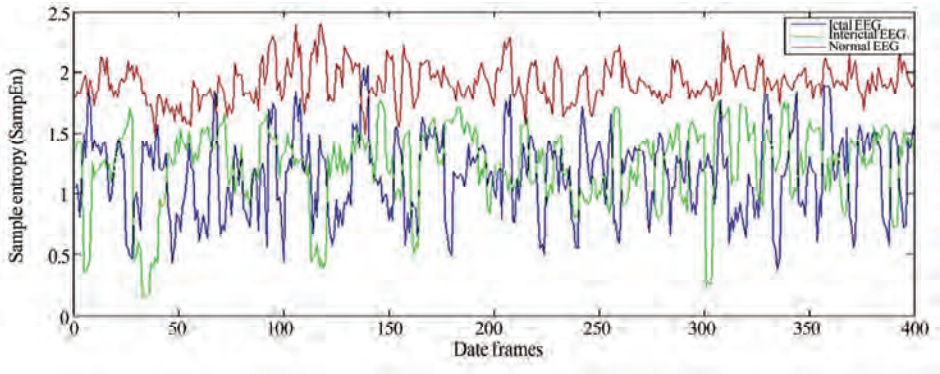


(a)

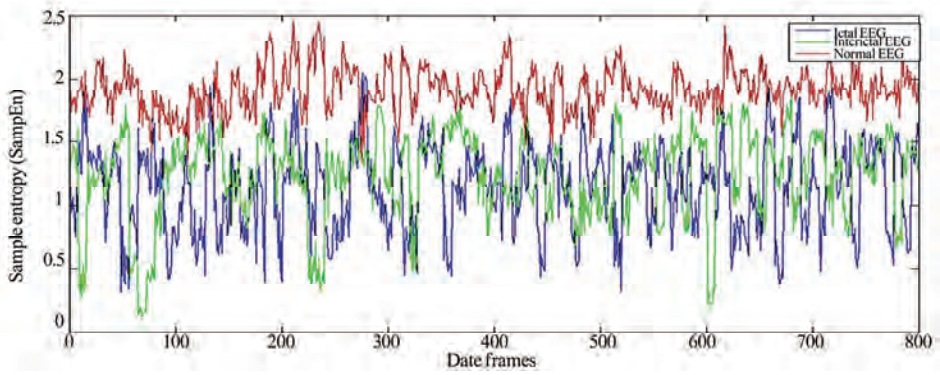


(b)

Figure 5. Sample figures of *SampEn* showing clear discrimination among normal, interictal and ictal EEG signals



(a)



(b)

Figure 6. Sample figures of *SampEn* showing partial overlap among normal, interictal and ictal EEG signals.

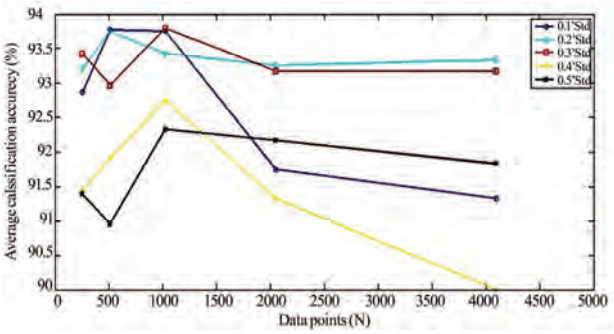


Figure 7. Average classification accuracy achieved by BPNN with m = 1.

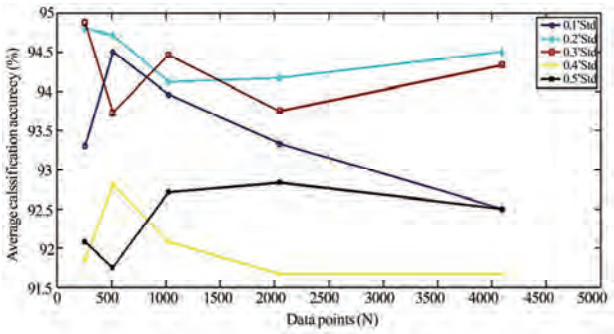


Figure 8. Average classification accuracy achieved by ELM with m = 1.

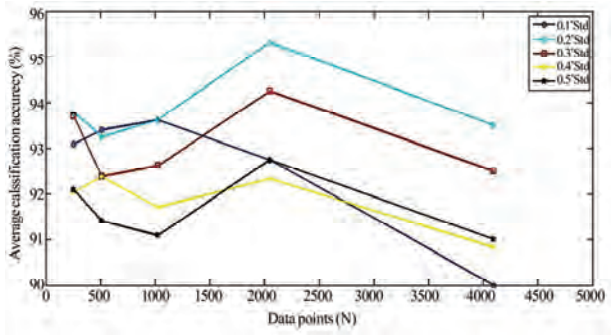


Figure 9. Average classification accuracy achieved by BPNN with m = 2.

1024. The average accuracies achieved by ELM for other parameter combinations range from 91.19%-94.83%, which are also acceptable for clinical diagnosis. From the results, it can be concluded that, generally, ELM outperforms BPNN for most of the parameter combinations.

Tables 2 and **3** show the classification results with the highest accuracies of the BPNN (95.33%) and the ELM (95.67%), respectively, by two confusion matrices. In terms of the confusion matrix for BPNN, all healthy segments were classified correctly by the BPNN, 2 seizure-free epileptogenic zone segments were classified incorrectly as healthy segments, 3 seizure-free epileptogenic zone segments were classified incorrectly as epileptic seizure segments and 2 epileptic seizure segments were classified incorrectly as seizure-free epileptogenic zone segments. In terms of the confusion matrix for ELM, all healthy segments were correctly classified, 2 seizure-free epileptogenic zone segments were classified incorrectly as healthy segments, 3 seizure-free epileptogenic zone segments were classified incorrectly as epileptic seizure segments and 2 epileptic seizure segments were classified incorrectly as seizue-free epileptogenic zone segments.

The values of statistical evaluation parameters introduced in Subsection 4.1.1 are given in **Table 4**. As can be

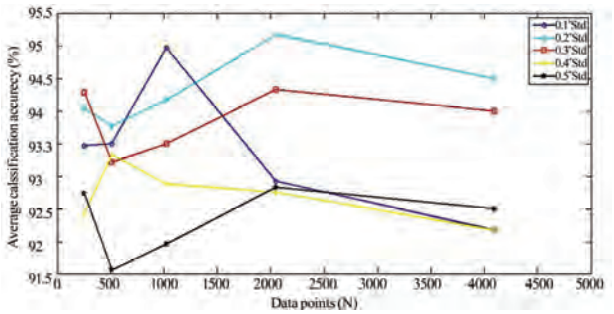


Figure 10. Average classification accuracy achieved by ELM with m=2.

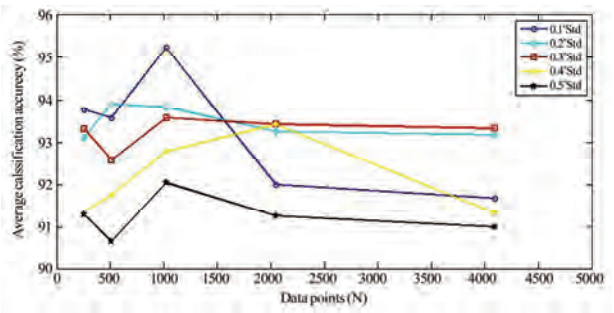


Figure 11. Average classification accuracy achieved by BPNN with m = 3.

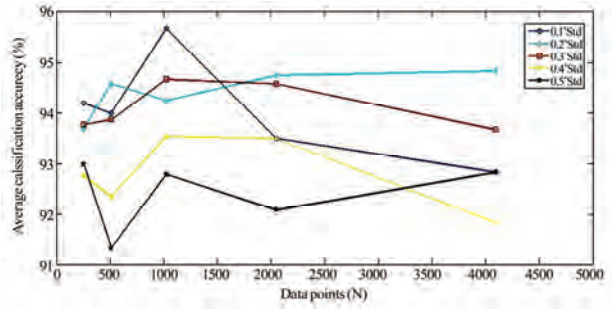


Figure 12. Average classification accuracy achieved by ELM with m = 3.

Table 2. Confusion matrix (BPNN).

Output/desired	Set A: healthy Segment	Set D: seizure-free Epileptogenic zone Segment	Set E: epileptic Seizure segment
Set A: healthy Segment	40	0	0
Set D: seizure-free Epileptogenic zone Segment	0	38	2
Set E: epileptic seizure Segment	0	3	37

Table 3. Confusion matrix (ELM).

Output/desired	Set A: healthy Segment	Set D: seizure-free Epileptogenic zone Segment	Set E: epileptic Seizure segment
Set A: healthy Segment	77	0	0
Set D: seizure-free Epileptogenic zone Segment	2	72	3
Set E: epileptic seizure Segment	0	2	84

Table 4. Performance comparison of ELM and BPNN.

Classification Model	Average Learning Time(Seconds)	Average Sensitivity (seizure free epileptogenic zone segments)	Average Sensitivity (epileptic seizure segments)	Average Specificity	Average Classification Accuracy
ELM	0.0250	91.06	97.26%	98.77%	95.67%
BPNN	86.4807	92.91%	95.81%	7.54%	95.33%

seen, the BPNN discriminated healthy segments, seizure-free epileptogenic zone segments and epileptic seizure segments with the average accuracies of 97.54%, 92.91% and 95.81%, respectively. The healthy segments, seizure-free epileptogenic zone segments and epileptic seizure segments were classified with the average accuracy of 95.33%. The average accuracies of the ELM were 98.77% for healthy segments, 91.06% for seizure-free epileptogenic zone segments, and 97.26% for epileptic seizure segments. The healthy segments, seizure-free epileptogenic zone segments and epileptic seizure segments were classified with an average accuracy of 95.67%. Hence, the average accuracy of the ELM classifier is slightly higher than that of the BPNN classifier. In addition, in **Table 4**, we find that the learning time of the ELM classifier is 0.0250 seconds while the learning time of the BPNN classifier is 86.4807 seconds. The ELM classifier can run 3459 times faster than the BPNN classifier. Thus, in the case of real-time implementation of epilepsy diagnosis support system, ELM classifiers are more appropriate than BPNN classifiers.

In **Table 5**, we present a comparison in classification performance achieved by different methods. We have quoted results from our present proposed method and also from recently reported in [30] and [31]. The datasets

Table 5. Comparison of classification accuracy obtained by our approach for the detection of epileptic seizures compared to the classification accuracy obtained by other researchers.

Authors	Method	Accuracy
Sadati *et al.* [30]	Discrete wavelet transform-Adaptive neural fuzzy network	85.90%
Übeyli [31]	Discrete wavelet transform-Combined neural network	94.83%
This work.	Sample entropy-Extreme learning machine	95.67%

used in these experiments are the same. It is shown in the table that the result obtained from our approach is the best presented for this dataset, indicating an improvement ment from 0.84% to 9.77% from other approaches proposed in the literature.

5. CONCLUSIONS

This study presents an attempt to develop a general-purpose EEG epilepsy detection scheme that can be used for classifying different kinds of EEG time series signals. Diagnosing epilepsy is not an easy task, which needs acquisition of patients' EEG recording and collecting additional clinical information. The proposed system employed a recently-proposed statistical parameter referred to as Sample entropy (*SampEn*), together with extreme learning machine (ELM) which is a recently-developed classification model, to classify subjects as normal subject, patients not having an epileptic seizure or patients having an epileptic seizure. This supplies a valuable diagnostic decision support tool for physicians treating potential epilepsy. Experimental results show that the proposed scheme achieves an excellent performance with not only the accuracy as high as 95.67% but also with very fast learning speed (0.0250 seconds), which demonstrates its potential for real-time implementation in an epilepsy diagnosis support system.

6. ACKNOWLEDGEMENTS

This work is financed by the EU 6 Framework Programme Project: Measuring and Modeling Relativistic-Like Effects in Brain and NCSs'.

REFERENCES

[1] Iasemidis, L.D., Shiau, D.S., Chaovalitwongse, W., Sackellares, J.C., Pardalos, P.M., Principe, J.C., Carney, P.R., Prasad, A., Veeramani, B. and Tsakalis, K. (2003) Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering*, **50(5)**, 616-627.

[2] Acir, N., Oztura, I., Kuntalp, M., Baklan, B. and Guzelis, C. (2005) Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks. *IEEE Transactions on Biomedical Engineering*, **52(1)**, 30-40.

[3] Ghosh-Dastidar, S., Adeli, H. and Dadmehr, N. (2007) Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. *IEEE Transactions on Biomedical Engineering*, **54(9)**, 1545-1551.

[4] Jahankhani, P., Kodogiannis, V. and Revett, D. (2006) EEG signal classification using wavelet feature extraction and neural networks. *IEEE John Vincent Atanasoff* 2006 *International Symposium on Modern Computing*, Sofia, 3-6 October 2006, 120-124.

[5] Kalayci, T. and Ozdamar, O. (1995) Wavelet preprocessing for automated neural network detection of EEG spikes. *IEEE Engineering in Medicine and Biology*, **14(2)**, 160-166.

[6] Übeyli, E.D. (2009) Statistics over features: EEG signals analysis. *Computers in Biology and Medicine*, **39(8)**, 733-741.

[7] Wang, B.J., Jun, L., Bai, J., Peng, L., Li, Y. and Li, G. (2006) EEG recognition based on multiple types of information by using wavelet packet transform and neural networks. *Proceedings of* 2005 *IEEE Engineering in Medicine and Biology Conference*, Shanghai, 17-18 June 2006, 5377-5380.

[8] Chandaka, S., Chatterjee, A. and Munshi, S. (2009) Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Systems with Applications*, **36(2)**, 1329-1336.

[9] Guler, I. and Übeyli, E.D. (2007) Multiclass support vector machines for EEG-signals classification. *IEEE Transactions on Information Technology in Biomedicine*, **11(2)**, 117-126.

[10] Lima, C.A.M., Coelho, A.L.V. and Chagas, S. (2009) Automatic EEG signal classification for epilepsy diagnosis with Relevance Vector Machines. *Expert Systems with Applications*, **36(6)**, 10054-10059.

[11] Kannathal, N., Rajendra, A.U., Lim, C.M. and Sadasivan, P.K. (2005) Characterization of EEG—A comparative study. *Computer Methods and Programs in Biomedicine*, **80(1)**, 17-23.

[12] Srinivasan, V., Eswaran, C. and Sriraam, N. (2007) Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Transactions on Information Technology in Biomedicine*, **11(3)**, 288-295.

[13] Richman, J.S. and Moorman, J. (2000) Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology*, **278(6)**, 2039-2049.

[14] Li, X.L., Ouyang, G.X. and Richards, D.A. (2007) Predictability analysis of absence seizures with permutation entropy. *Epilepsy Research*, **77(1)**, 70-74.

[15] Ocak, H. (2009) Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications*, **36(2)**, 2027-2036.

[16] Pravin, K.S., Sriraam, N. and Prabhu, B.G. (2008) Automated detection of epileptic seizures using wavelet

entropy feature with recurrent neural network classifier. *Proceedings of the* 2008 *IEEE Region* 10 *Conference of TeNCON*, Hyderabad, 19-21 November 2008, 1-5.

[17] Huang, G.B., Zhu, Q.Y. and Siew, C.K. (2004) Extreme learning machine: A new learning scheme of feedforward neural networks. *Proceedings of the* 2004 *International Joint Conference on Neural Networks*, **2**, 25-29 July 2004, 985-990.

[18] Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P. and Elger, C.E. (2001) Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, **64(6)**, 61907(1-8).

[19] Sleigh, J.W., Steyn-Ross, D.A., Grant, C. and Ludbrook, G. (2004) Cortical entropy changes with general anaesthesia: theory and experiment. *Physiological Measurement*, **25(4)**, 921-934.

[20] Abásolo, D., Hornero, R., Espino, P., Álvarez, D. and Poza, J. (2006) Entropy analysis of the EEG background activity in Alzheimer's disease patients. *Physiological Measurement*, **27(3)**, 241-253.

[21] Aboy, M., Cuesta-Frau, D., Austin, D. and Mico-Tormos, P. (2007) Characterization of sample entropy in the context of biomedical signal analysis. *Proceedings of the* 29*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, 22-26 August 2007, 5942-5945.

[22] Al-Angari, H.M. and Sahakian, A.V. (2007) Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *IEEE Transactions on Biomedical Engineering*, **54(10)**, 1900-1904.

[23] Wang D. and Huang, G.B. (2005) Protein sequence classification using extreme learning machine. *Proceedings of the* 2005 *IEEE International Joint Conference on Neural Networks*, Montreal, **3**, 31 July-4 August 2005, 53-59.

[24] R. Zhang, R., Huang, G.B., Sundararajan, N. and Saratchandran, P. (2007) Multicategory classifcation using an extreme learning machine for microarray gene expression cancer diagnosis. *IEEE Transactions on Computational Biology and Bioinformatics*, **4(3)**, 485-495.

[25] Liang, N.Y., Saratchandran, P., Huang, G.B. and Sundararajan, N. (2006) Classification of mental tasks from EEG sginals using extreme learning machine. *International Journal of Neural Systems*, **16(1)**, 29-38.

[26] Kim, J., Shin, H., Lee, Y. and Lee, M. (2007) Algorithms for classifying arrhythmia using extreme learning machine and principle component analysis, *Proceedings of the* 29*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, 22-26 August 2007, 3257-3260.

[27] Hagan, M.T. and Menhaj, M.B. (1994) Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, **5(6)**, 989-993.

[28] Lake, D.E., Richman, J.S., Griffin, M.P. and Moorman, J.R. (2002) Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology—Regulatory, Integrative and Comparative Physiology*, **283(3)**, 789-797.

[29] Molinaro, A., Simon, R. and Pfeiffer, R. (2005) Prediction error estimation: A comparison of resampling methods. *Bioinformatics*, **21(15)**, 3301-3307.

[30] Sadati, N., Mohseni, H.R. and Mamgshoudi, A. (2006) Epileptic seizure detection using neural fuzzy networks. *Proceedings of the* 2006 *IEEE International Conference on Fuzzy Systems*, Vancouver, 16-21 July 2006, 596-600.

[31] Übeyli, E.D. (2009) Combined neural network model employing wavelet coeffients for EEG signals classification. *Digital Signal Processing*, **19(2)**, 297-308.

J. Biomedical Science and Engineering, 2010, 3, 568-575 **JBiSE**
doi:10.4236/jbise.2010.36079 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Ridge penalized logistical and ordinal partial least squares regression for predicting stroke deficit from infarct topography

Jian Chen[1,2*], Thanh G. Phan[1], David C. Reutens[1,3]

[1]Stroke and Aging Research, Department of Medicine, Monash University, Melbourne, Australia;
[2]Developmental and Functional Brain Imaging, Murdoch Childrens Research Institute, Victoria, Australia;
[3]Experimental Neurology, Centre for Advanced Imaging, University of Queensland, Queensland, Australia.
Email: jian.chen@med.monash.edu.au

Received 17 October 2009; revised 15 December 2009; accepted 20 December 2009.

ABSTRACT

Improving the ability to assess potential stroke deficit may aid the selection of patients most likely to benefit from acute stroke therapies. Methods based only on 'at risk' volumes or initial neurological condition do predict eventual outcome but not perfectly. Given the close relationship between anatomy and function in the brain, we propose the use of a modified version of partial least squares (PLS) regression to examine how well stroke outcome covary with infarct location. The modified version of PLS incorporates penalized regression and can handle either binary or ordinal data. This version is known as partial least squares with penalized logistic regression (PLS-PLR) and has been adapted from its original use for high-dimensional microarray data. We have adapted this algorithm for use in imaging data and demonstrate the use of this algorithm in a set of patients with aphasia (high level language disorder) following stroke.

Keywords: Ridge Penalized; Logistical PLS; Stroke

1. INTRODUCTION

Correlations between brain lesions and clinical symptoms have yielded valuable insights into brain function in the past. In individual patient care, these clinico-lesion correlations may play a role in predicting neurological deficits following stroke. More recently, attempts have been made to utilize the information obtained from brain imaging studies to aid prediction of neurological outcome. Initial approaches depended upon measurement of infarct volume but volumetric approaches proved to be inaccurate predictors of neurological outcome. The correlation between infarct volume and the National Institutes of Health Stroke Scale (NIHSS) is moderate at best [1,2]. One factor ignored in volumetric approaches is the information on stroke location available in the images. We have recently demonstrated that the relationship between tissue damage assessed at the voxel level and neurological disability can be predicted using a new method: Ridge Penalized Logistic Partial Least Squares (RPL-PLS). This method allows both stroke extent and location to be incorporated into the predictive model for neurological deficit.

Previously, voxel-based statistical techniques have concentrated on the relationship between involvement of individual voxels or clusters of voxels and neurological deficit [3-5]. However, strokes often involve large ensembles of voxels and the task involved in prediction is to establish the relationship between involvement of ensemble as a whole and neurological deficit. The functional inter-correlation between different groups of voxels is likely to mean that their contribution to outcome is not independent. One method of dealing with this kind of issue is principal components regression (PCR), which uses orthogonal linear combinations of the original predictor variables as predictors in a multiple linear regression [6]. In PCR orthogonal linear combinations of the original predictor variables are first constructed as principle components (PCs) to maximize the variance of data. These PCs are then used as predictors in a multiple linear regression. Thus dimension reduction in PCR is achieved without regard the response variable. Partial least squares regression (PLS), as an alternative method, has the advantage over PCR in that it takes into account the response variable when performing the dimension reduction step [7,8].

Bookstein 1994 [9], McIntosh, *et al.* [10] and Leibovitch, *et al.* [11] introduced a variant of the partial least squares (PLS) approach to the brain imaging community. Here singular-value decomposition (SVD) is applied to the cross-correlation matrix between dependent and independent variables to yield latent variables which are

linear combinations of the original variables, and which maximise the explained covariance. This characteristic of PLS makes it more suited to the purpose of prediction on the basis of involvement of functionally related ensembles of voxels. The reduction in dimensionality achieved may reflect the functional relationships between brain regions.

There are several challenges in using the PLS technique to build a prediction model. First, there is a high degree of correlation among neighbouring voxels due to shared function and shared vascular blood supply. This leads to collinearity thus preventing stable estimates of regression coefficients. Second, the outcome variables are binary or ordinal and are correctly dealt with using logistic regression with the dependent variable being transformed into a logit variable describing the odds of a specific outcome. Thirdly, estimates of model coefficients using generalized least squares may still fail to converge. The solution of the first issue is the introduction of a ridge estimator to PLS and such analysis has recently been shown to provide stable estimate in microarray data analysis [12-14]. The solution of the second issue is achieved by embedding the usual PLS steps within the iterative re-weighted least square (IRLS) [15]. In this setting, the binary variables were transformed to the continuous-valued pseudo-response variable by logit conversion. Variables from logistic regression are further constrained to be finite by penalizing with a ridge estimator for overcoming the convergence issue before feeding to the PLS. Finally, standard PLS method has been extended to weighted partial least squares (WPLS) to further reduce noise effects and to improve the convergence of the PLS. WPLS penalizes or regularizes PLS model by giving samples different weights (based on their relevance to the study). This additional weight determines how much each observation in the data set influences the final parameter estimates and the 'dispersion matrix', from logistical regression, can be severed as weights for the WPLS (detailed in methods section).

We have successful demonstrated RPL-PLS in stroke deficit prediction study [16]. In this study we describe this modification of PLS method to take into account binary as well as ordinal outcome variables. To illustrate the use of this technique we describe its use in predicting stroke outcome using only knowledge of the location and extent of infarction. In Section 2 we describe the theory of the algorithm and its implementation; in Section 3 we describe an application of the method to stroke data, in Section 4 is result and in Section 5 is discussion.

2. METHODS

2.1. Partial Least Squares Regression (PLS) and Weighted PLS

PLS [7] is a dimension reduction technique, which addresses the issue of multiple regression when the number of variables are greater than the number of observations. The n observations described by p dependent variables are stored in a $n \times p$ matrix denoted **Y**, and the values of m predictors collected on these observations are in a $n \times m$ matrix **X**. PLS regression searches k number, with $k <= m$, of principle component scores and loadings (latent variables) by performing an iterative simultaneous decomposition of independent data **X** and dependent data **Y.**

In matrix form, PLS decomposes **X** and **Y** into the form:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (1)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} \quad (2)$$

where the **T** and **U** are ($n \times k$) score matrices, the ($m \times k$) **P** and the ($p \times k$) **Q** are matrices of loadings. **E** and **F** are matrices of residuals. The regression model is then step up between the scores:

$$\mathbf{U} = \mathbf{B}\mathbf{T} \quad (3)$$

These matrixes are column centered and normalized (the symbol means "to normalize the result of operation"). The PLS regression method described here is based in the nonlinear iterative partial least squares (NIPLALS) algorithm [7], Iterative decomposition starts with random initialization the Y-score vector **u,** with initial **E** = **X** and initial **F** = **Y**, and iteratively go though the following steps until a stopping criterion is met or **E** becomes a null matrix.

Step 1. $\mathbf{w} \propto \mathbf{E}^T\mathbf{u}$ (estimate **E** weights)
Step 2. $\mathbf{t} \propto \mathbf{E}\mathbf{w}$ (estimate **E** scores)
Step 3. $\mathbf{q} \propto \mathbf{F}^T\mathbf{t}$ (estimate **F** weights)
Step 4. $\mathbf{u} \propto \mathbf{F}\mathbf{q}$ (estimate **F** scores)
Step 5. Check covariance, if t has not converged, goes to Step1, else go to Step 6.
Step 6. $b \propto \mathbf{t}^T\mathbf{u}$ (compute regression coefficient)
Step 7. $\mathbf{E} = \mathbf{E} - \mathbf{t}\mathbf{p}^T$ (residual matrix of **E**)
Step 8. $\mathbf{F} = \mathbf{F} - \mathbf{b}\mathbf{t}\mathbf{q}^T$ (residual matrix of **F**)

By regressing **E** on **t** and **F** on **u**, the loading vectors $\mathbf{p} = (\mathbf{t}^T\mathbf{t})^{-1}\mathbf{E}^T\mathbf{t}$ and $\mathbf{q} = (\mathbf{u}^T\mathbf{u})^{-1}\mathbf{F}^T\mathbf{u}$ can be computed. In this way it finds the weight vectors, **w**, **q** such that

$$\begin{aligned}[\mathrm{cov}(\mathbf{t},\mathbf{u})]^2 &= [\mathrm{cov}(\mathbf{E}\mathbf{w},\mathbf{F}\mathbf{q})]^2 \\ &= \max_{|r|=1,|s|=1}[\mathrm{cov}(\mathbf{E}r,\mathbf{F}s)]^2\end{aligned} \quad (4)$$

where the sample covariance between two variables are kept maximized through maximizing the sample covariance between the two scores (components) at each decomposition step. In such a way, it minimizes the norm of **Y** while keeping the correlation between **X** and **Y** by the *inner relation* **Eq.3**. Once the relationship has been built, the dependent variables are predictable using multivariate regression formula, the *outer relation*, as:

$$\mathbf{Y} = \mathbf{TBQ}^{\mathrm{T}} = \mathbf{XB}_{\mathrm{PLS}} \quad (5)$$

with $\mathbf{B}_{\mathrm{PLS}} = (\mathbf{P}^{\mathrm{T}+})\mathbf{BQ}^{\mathrm{T}}$ (where $\mathbf{P}^{\mathrm{T}+}$ is Moore-Penrose pseudo-inverse of $\mathbf{P}^{\mathrm{T}}$).

Least squares solution of linear regression is only appropriate when the variances of the predictor variable are uniform [17]. When there are unreliable data or errors in the data measurement, unequal diagonal elements in the variance of the error matrix will lead to instability of parameter estimate for the least squares formula. Weighted partial least squares (WPLS) generalize PLS with an empirical weighted squared error in the same way that weighted least square regression generalized least squares regression. The main idea is to penalize or regularize the coefficients of WPLS model and to facilitate model interpretation and further reduce noise effects of the samples: instead of weighting all samples equally, they are weighted such that samples with great weight contribute more to fit. WPLS defines k number of $\mathbf{V}$ weighted orthogonal scores $\mathbf{t}_k$, linear combination of $\mathbf{X}$ such that for all k, $\prod_n^{\mathrm{T}} v\mathbf{t}_k$ and performs a $\mathbf{V}$ weighted least squares regression of $\mathbf{Y}$ through $\mathbf{U}$ on $\mathbf{T}$. $\mathbf{V}$ is a symmetric positive definite matrix with v_{ii} is the weight assigned to each sample, is induced with the belief that observations with small variances provide more reliable information about the regression function than those with large variances. PLS is a special case of WPLS with $\mathbf{V}$ as identical matrix. In this study we will use WPLS to compensate the problem of possible unequal variance in the error matrix. The element v_{ii} of $\mathbf{V}$ is a probability of occurrence of sample i obtained from logistic regression step (detailed in following).

2.2. Ridge Penalized Logistic Regression

PLS was originally designed for normal random response variables. In the presence of binary response variable, linear regression can result in regression coefficients, which cannot guarantee that response values only have two possible predicted values, 0 and 1. Logistic regression is one of the approaches to this issue. Let variable y_i indicates the class of sample i for response variable y and π_i the probability that $y_i = 1$. Consequently, the probability of sample represents a class 0 is then $1 - \pi_i$. Let x_{ij} indicate the jth independent variable in the ith sample. The logistic regression model is:

$$\eta_i = \log\frac{\pi_i}{1-\pi_i} = \beta_0 + \sum_{j=1}^{m}\beta_j x_{ij} \quad (6)$$

where η_i is called the linear predictor in the jargon of generalized linear model. It is connected to π_i by so-called link function f with

$$f(\pi) = \log(\frac{\pi}{1-\pi}) \quad (7)$$

In vector format $\eta_{\mathrm{i}} = \boldsymbol{\beta}[1 \ \ \mathbf{x}_i^{\mathrm{T}}]$. $\boldsymbol{\beta}$ is unknown parameter and could be estimated by the maximum likelihood estimator (MLE), $\hat{\boldsymbol{\beta}}$. The log-likelihood of the observations for the value $\boldsymbol{\beta}$ of the parameters $L(\boldsymbol{\beta})$ is given by

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n} y_i \log\pi_i + \sum_{i=1}^{n}(1-y_i)\log(1-\pi_i) \quad (8)$$

If $\mathbf{z}=[1 \ \ \mathbf{x}^{\mathrm{T}}]$ is full column-rank and the configuration of n samples in the observation space is overlap, the solution exists and is unique. This solution could be computed by the literately reweighted least squares (IRLS) [18]. Let $\mathbf{V}^{\mathrm{T}}$ be the $n \times n$ diagonal matrix with $\mathbf{v}_{ii}^{\mathrm{T}} = \boldsymbol{\pi}_i^{\mathrm{T}}[1-\boldsymbol{\pi}_i^{\mathrm{T}}]$ at iteration t and $\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{T}}$. Each iteration divides into two steps,

$$\mathbf{g}^{\mathrm{T}} = \mathbf{Z}\boldsymbol{\beta}^{\mathrm{T}} + [\mathbf{V}^{\mathrm{T}}]^{-1}(\mathbf{y} - \boldsymbol{\pi}^{\mathrm{T}}) \quad (9)$$

$$\boldsymbol{\beta}^{\mathrm{T}+1} = (\mathbf{Z}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{g}^{\mathrm{T}} \quad (10)$$

where $\mathbf{g}$ is the calculated new response variables (detailed in Appendix).

Multicollinearity can still exist even after dimension reduction in the setting of our study: many voxels will show nearly identical patterns across the samples and they may supply no additional information to the model. This issue can be further addressed by introducing the ridge estimator, the regularization on sum of the squares of regression coefficients [19], into the logistic regression [20].

The ridge estimator, $\hat{\boldsymbol{\beta}}^R$, is defined as the (unique) maximum of the penalized log-likelihood

$$L(\boldsymbol{\beta})^* = L(\boldsymbol{\beta}) - \frac{\lambda}{2}\boldsymbol{\beta}^T\mathbf{R}\boldsymbol{\beta} \quad (11)$$

where $\lambda > 0$ is the shrinkage parameter, the stronger its influence and the smaller the β_j^2's are forced to be. $\hat{\boldsymbol{\beta}}^R$, always existing, is unique. Ridge-IRLS (RIRLS) replaces the weighted regression (**Eq.10**) in IRLS by a weighted ridge regression

$$\boldsymbol{\beta}^{t+1} = (\mathbf{Z}^T\mathbf{V}^t\mathbf{Z} + \lambda\mathbf{R})^{-1}\mathbf{Z}^T\mathbf{V}^t\mathbf{g}^t \quad (12)$$

where $\mathbf{R}$ is a diagonal matrix with entries $R_{1,1} = 0$ and

$$\mathbf{R} = \sum_{i=1}^{n}(Z_{i,j} - \boldsymbol{\Pi}_n^{\mathrm{T}}\frac{Z_{\cdot,j}}{n})^2, \quad j \quad \{2,\ldots, m+1\} \quad (13)$$

with $Z_{\cdot,\mathrm{j}} = [Z_{1\mathrm{j}}, Z_{2\mathrm{j}}, \ldots, Z_{\mathrm{nj}}]$.

$\boldsymbol{g}^{\mathrm{T}}$ in **Eq.12** is built as in **Eq.9**. λ can be chosen as the minimum, over a given range, of the Bayesian information criterion (BIC) which gives the best balance between model complexity and the best fit to the data [21],

$$-2L(\hat{\boldsymbol{\gamma}}^R) + \log(n)trace[\mathbf{Z}(\mathbf{Z}^T\mathbf{V}(\hat{\boldsymbol{\beta}}^R)\mathbf{Z} + \lambda\mathbf{R}^2)^{-1}$$

$$\times\mathbf{Z}^T\mathbf{V}(\hat{\boldsymbol{\beta}}^R)]\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

2.3. Ridge Penalized Logistic Partial Least Squares Regression

Embedding ridge penalized logistic regression into PLS procedure forms RPL-PLS. This method involves two steps. The first step, ridge penalty logistic regression (RIRLS), builds a continuous response variable $\mathbf{g}^{\infty}$ and 'dispersion matrix' $[\mathbf{V}^{\infty}]^{-1}$ for the input of the second step. Second step is weighted PLS (WPLS) [12].

1) $(\mathbf{g}^{\infty}, \mathbf{V}^{\infty}) \leftarrow RIRLS(\mathbf{Y}, \mathbf{X}, \lambda)$

2) $\hat{\boldsymbol{\beta}}^{PLS} \leftarrow WPLS(\mathbf{g}^{\infty}, \mathbf{X}, \mathbf{V}^{\infty}, k)$

There are two parameters, shrinkage parameter λ and number of component k, to be determined in RPL-PLS. λ, as stated early, is determined by BIC in the first step. The optimal number k is empirically chosen by selecting the minimal number of components that give the minimum leave-one-out cross-validation (LOOCV) error rate for the training data. RPL-PLS provides unique an estimate $\hat{\boldsymbol{\beta}}^{PLS}$ for given $\mathbf{Y}$, $\mathbf{X}$, λ and k.

Binary logistic regression can be easily extended to ordinal response variables by creating a sequence of binary response variables, one for each response category [18]:

$$y_i^1 = \begin{cases} 1 & \text{if } i\text{th sample response is category 1} \\ 0 & \text{Otherwise} \end{cases}$$

$$y_i^2 = \begin{cases} 1 & \text{if } i\text{th sample response is category 2} \\ 0 & \text{Otherwise} \end{cases}$$

$$\vdots$$

$$y_i^c = \begin{cases} 1 & \text{if } i\text{th sample response is category c} \\ 0 & \text{Otherwise} \end{cases}$$

The same technique can be applied to RPL-PLS to form more generalized multi- ordinal RPL-PLS.

2.4. Choosing the Model

The maximum number of components from RPL-PLS is equal to the number of samples in the dataset. Since these components are sorted in a descending order according to the proportion of variance they explained, only the first of few components were needed and the rest were considered as noise. The number of components could be made up to number of samples and optimal number of components was determined by Leave-one-out cross-validation step and when the error rate became stable. These models were illustrated up to 6 components which have already comprised most of variance of the data. The optimal number of components for each model was selected by choosing the value of k minimizing LOOCV error rate in cross-validation of the training dataset.

3. MATERIALS

Patients were recruited if they had an ischemic stroke in the anterior circulation. 38 patients were used for development of the model (training dataset) and 22 patients were used for the model validation (validation dataset). Neurological deficit from stroke was measured on an ordinal scale of the NIHSS and assessment was performed immediately prior to MR imaging. The domain of interests for this demonstration was aphasia (higher language disorder). The NIHSS language component is rated 0 (normal), 1 (mild to moderate), 2 (severe) and 3 (mute and global aphasia). In our ordinal model, a score of 1 correspond to NIHSS language score of 0, a score of 1 correspond to NIHSS language score of 1-2 and a score of 3 correspond to NIHSS score of 3.

MR scans were acquired within three months after stroke onset. Fast spin echo T2-weighted images were acquired on 1.5T scanner (GE, Milwaukee, WI) with thickness 6 mm/1.7 mm, matrix 256 × 256, and TR/TE/ETL 2000 ms/102 ms/12. Images from different subjects were aligned to a standard brain template registration [22] by manual registration using 9-parameter linear transformation [23]. Infarcts were manually segmented on standard space images using interactive mouse driven software. Due to memory limitations of the PC, binary images were resampled to 4 mm × 4 mm × 4 mm as the input of RPL-PLS. The computation scripts were implemented in MATLAB (Mathworks, Inc., MA).

4. RESULTS

RPL-PLS is a robust method and has convergence for all three models. In LOOCV, the optimal number of the components, k, was 2 for aphasia (binary), and 3 for aphasia (ordinal). The algorithm correctly identified 37 of 38 samples for aphasia (binary) using two components and 37 of 38 samples for aphasia (ordinal) using 3 components. In a model, the coefficients of each voxel in the components indicate the relative importance of that voxel (anatomical locations) to the associated neurological deficit. The cross-validation results of models consisted different number of components were illustrated in **Table 1**.

In external validation with new data set consist of 22 samples. Binary aphasia model produced 4 errors (81.8% correct) and ordinal aphasia model produced 5 errors (77.3% correct).

Figure 1 corresponds to the binary aphasia model. Since the optimal number of components for this model was 2, left column is the first component of the model wiliest the second column is the second component of the model. The brighter the voxel is, the higher the weighting of the voxel with respect to aphasia deficit.

In **Figure 2** we presented coefficient images of three

Table 1. Number of errors in LOOCV of 38 training data samples.

Neuropsychological assessment		Number of components					
		1	2	3	4	5	6
Number of errors	Aphasia (ordinal)	7	4	1	1	1	1
	Aphasia (Binary)	5	1	1	1	1	1

components model of aphasia ordinal model. Images in the first the column of **Figure 2** showed each voxel relate to the aphasia score = 1 when using a model compromise three components, the second column images related to aphasia score = 2, and the third column images relate to aphasia score = 3.

5. DISCUSSIONS

In this study, we developed novel approach of generalized regression method, RPL-PLS, for predicting neurological deficit from MRI image data. The method incorporates dimension reduction techniques and ridge penalized logistic regression for addressing the problem of large collinearity dataset with binary and ordinal response variables. The PLS algorithm described in this paper is known as the 'standard' PLS algorithm and has been presented in detail elsewhere [7,8,24-26].

The model built from the training dataset has produced encouraging results for predicting different neurological domain following stroke for the new dataset. It only uses information presented in the MR image and has

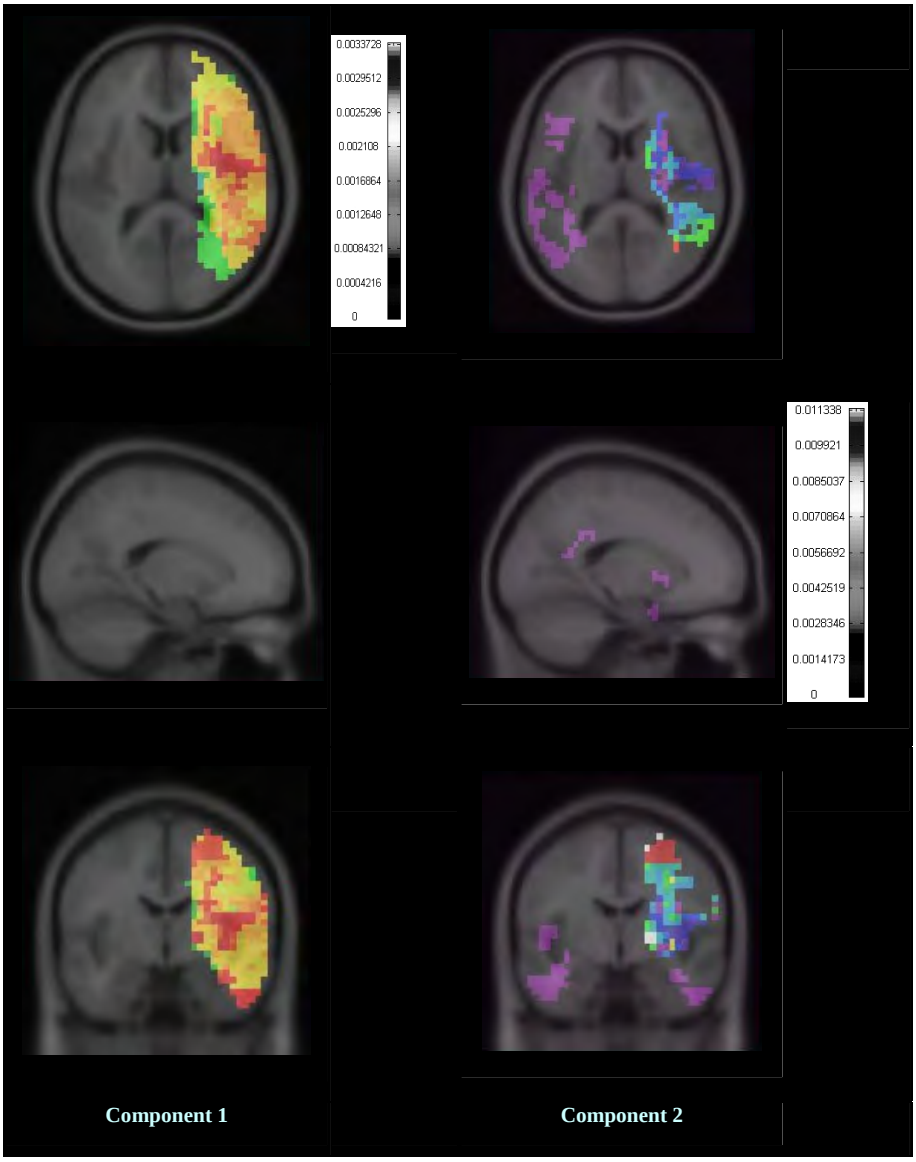


Figure 1. Image representation of first 2 components of aphasia binary model (left side image corresponds to right side of the patient, radiological conversion).

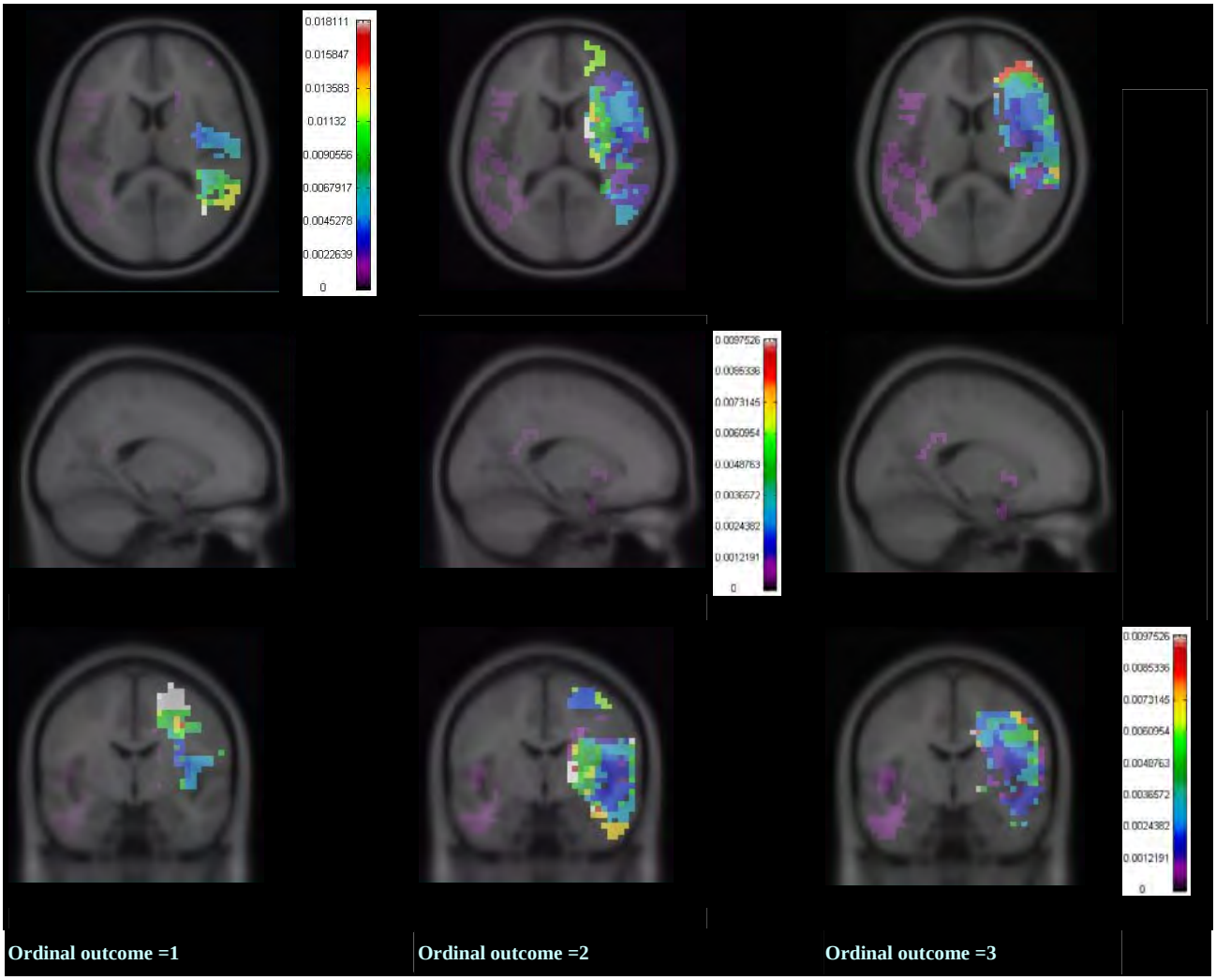


Figure 2. Image representation of 3 components aphasia ordinal model (left side image corresponds to right side of the patient, radiological conversion).

no requirement from human expert observer. This novel approach of using infarct topography to describe neurological deficit is an improvement of cruder volumetric methods. This study provides support of the concept that information presented in image can be used to predict the outcome of stroke. This concept paves way for the development of similar model for understanding the neuroanatomy of neurological deficits and determ ining the outcome of rehabilitation and acute stroke trial.

For this proof-of-concept study we examined patients with well-defined infarcts on MRI scans acquired 3 months after infarction to predict outcome at 3 month. In this aspects, the model described here does not conform to a typical definition of a prediction model which is to use early MRI scans (< 1 week) to predict long term outcome (at 3 months). Nevertheless, the concept developed here can be used to obtain the "holy grail" of prediction. We would anticipate that with the appropriate training set, the method would also perform well at other time points after infarction, for example in the acute stage (less than 1 week). To increase the homogeneity of the group for this proof of concept study, we restricted the analysis to patients with infarcts in the anterior circulation. Future studies involving other infarct territories will be required to assess whether this method of correlating infarct extent and location will perform as well for other brain regions.

REFERENCES

[1] Barber, P.A., Darby, D.G., Desmond, P.M., Yang, Q., Gerraty, R.P., Jolley, D., Donnan, G.A., Tress, B.M. and Davis, S.M. (1998) Prediction of stroke outcome with echoplanar perfusion- and diffusion-weighted MRI. *Neurology*, **51(2)**, 418-426.

[2] Wardlaw, J.M., Keir, S.L., Bastin, M.E., Armitage, P.A. and Rana, A.K. (2002) Is diffusion imaging appearance an

independent predictor of outcome after ischemic stroke? *Neurology*, **59(9)**, 1381-1387.

[3] Kertesz, A. (1979) Aphasia and associated disorder: Taxonomy, localization and recovery. Grune & Stratton, Inc., New York.

[4] Dronkers N.F. (1996) A new brain region for coordinating speech articulation. *Nature*, **384**, 159-161.

[5] Bates, E. Wilson, S.M. Saygin, A.P. Dick, F. Sereno, M.I. Knight, R.T. and Dronkers, N.F. (2003) Voxel-based lesion-symptom mapping. *Nature Neuroscience*, **6(5)**, 448-450.

[6] Frank, I. and Friedman, J. (1993) A statistic review of some chemometrics regression tools, with discussion, *Technometrics*, **35(2)**, 109-148.

[7] Wold, H. (1975) Soft modelling by latent variables: Non-linear iterative partial least squares (NIPALS) approach. In: Gani, M.S.B., Ed., *Perspectives in Probability and Statistics*, Academic Press, London, 117-142.

[8] Naes, T. and Martens, H. (1985) Comparison of prediction methods for multicollinearity data. *Communication Statist Assoc*, **60**, 234-246.

[9] Bookstein, F.L. (1994) Partial least squares: A dose-response model for measurement in the behavioral and brain sciences. *Psycoloquy*, **5(23)**, least squares (1).

[10] McIntoch, A.R., Bookstein, F.L., Haxby, J.C. and Grady, C.L. (1996) Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage*, **3(3)**, 143-157.

[11] Leibovitch, F.S., *et al*. (1999) Brain SPECT imaging and left hemispatial neglect covaried using partial least squares: the sunnybrook stroke study. *Human Brain Mapping*, **7(4)**, 244-253.

[12] Fort, G. and Lambert-Lacroix, S. (2005) Classification using partial least squares with penalized logistic regression. *Bioinformatics*, **21(7),** 1104-1111.

[13] Shen, L. and Tan, E.C. (2005) PLS and SVD based penalized logistic regression for cancer classification using microarray data. *Proceedings of the* 3*rd Asia-Pacific Bioinformatics Conference*, Singapore, 17-21 January 2005, 219-228.

[14] Huang, X.H., Pan, W., Han, X.Q., Chen, Y.J., Miller, L.W. and Hall, J. (2005) Borrowing information from relevant microarray studies for sample classification using weighted partial least squares. *Computational Biology and Chemistry*, **29(3)**, 204-211.

[15] Marx, B.D. (1996) Iterative reweighted least squares estimation for generalized linear regression. *Technometrics*, **38(4),** 374-381.

[16] Phan, T.G., Chen, J., Donnan, G., Srikanth, V., Wood, A. and Reutens, D.C. (2009) Development of a new tool to correlate stroke outcome with infarct topography: A proof-of-concept study. *NeuroImage*, **49(1)**, 127-133.

[17] Draper, N.R. and Smith, H. (1998) Applied Regression Analysis, 3rd Edition, Wiley, New York.

[18] Kutner, M.H., Neter, J., Nachtsheim, C.J. and Li, W. (2004) Applied linear statistical models, 5th Edition. McGraw- Hill Irwin, Boston.

[19] Hoerl, A.E. and Kennard, R.W. (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, **12(1)**, 55-67.

[20] Le Cessie, S. and van Houwelingen, J.C. (1992) Ridge estimators in logistic regression, *Applied Statistics*, **41(1)**, 191-201.

[21] Kass, R. and Raftery, A. (1995) Bayes factor. *Journal of the American Statistical Association*, **90(430)**, 773-795.

[22] Talairach, J. and Tournoux, P. (1988) Co-planar stereotactic atlas of the human brain. Thieme Medical Publishers, New York.

[23] Woods, R.P., Grafton, S.T., Watson, J.D., Sicotte, N.L. and Mazziotta, J.C. (1998) Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, **22(1)**, 153-165.

[24] Wold, S., Martens, H. and Wold, H. (1983) The multivariate calibration problem in chemistry solved by the PLS method. In: Ruhe, A. and Kagstrom, B. Eds., *Proceedings of the Conference on Matrix Pencils*, Pite Havsbad, 22-24 March 1983, 286-293.

[25] Geladi, P. and Kowalski, B.P. (1986) Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, **185(1)**, 1-17.

[26] Abdi, H. (2003) Partial least squares (PLS) regression. In Bryman, A. Futing, T. and Lewis-Beck, M. Eds., *Encyclopedia of Social Sciences Research Methods*, London.

APPENDIX

Maximum likelihood (*ML*) *estimate of logistic regression and Iterative reweighted least squares* (*IRLS*)

When response variable $y_1, y_2, \ldots, y_n$ are binary, taking on the values 0 and 1 with probabilities and $1-\pi$, respectively, with expect value $E\{y\} = \pi$, covariates $x_{ij}\{i = 1, 2, \ldots, n; j = 1, 2,\ldots, m\}$ are also available, the logistic regression model would construct by a canonical link function

$$\eta = \log\frac{\pi}{1-\pi} = \alpha + \sum_{j=1}^{m}\beta_j x_j = \alpha + \boldsymbol{\beta}^T\mathbf{x} \qquad (A.1)$$

$$E\{\mathbf{y}\} = \pi = \frac{e^{(\alpha+\boldsymbol{\beta}^T\mathbf{x})}}{1+e^{(\alpha+\boldsymbol{\beta}^T\mathbf{x})}} \qquad (A.2)$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_m]$, $\mathbf{x} = [x_1, x_2, \ldots x_m]$. Thus the probability distribution of y is

$$f(y) = \pi^y(1-\pi)^{1-y} \qquad (A.3)$$

Since the y_i observations are independent, their joint probability function is

$$h(y_1, y_2, \ldots y_n) = \prod_{i=1}^{n} f_i(y_i) = \prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i} \qquad (A.4)$$

It is often more convenient to work with the logarithm of the joint probability function to find the maximum likelihood estimate

$$\begin{aligned}\log h(y_1, y_2, \ldots y_n) &= \log\prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i}\\ &= \sum_{i=1}^{n}[y_i\log(\frac{\pi_i}{1-\pi_i})] + \sum_{i=1}^{n}\log(1-\pi_i)\end{aligned} \qquad (A.5)$$

If we substitute (A.2) to (A.5) and consider equation a1, we have

$$\log L(\alpha, \boldsymbol{\beta}) == \sum_{i=1}^{n} y_i(\alpha + \boldsymbol{\beta}\mathbf{x}_i^T) - \sum_{i=1}^{n}\log[1+e^{(\alpha+\boldsymbol{\beta}\mathbf{x}_i^T)}] \qquad (A.6)$$

where $\mathbf{x}_i$ is shorthand of $[x_{i1}, x_{i2}, \ldots x_{ip}]$ and $L(\alpha, \boldsymbol{\beta})$ replaces $h(y_1, y_2,\ldots y_n)$ to show explicitly that we now view this function as the likelihood function of the parameters to be estimated. Denote $\mathbf{Z}_i = [1 \quad \mathbf{x}_i^T]$ and $\boldsymbol{\gamma} = [\alpha \,|\, \boldsymbol{\beta}]$, we have

$$\ell(\gamma) = \log L(\boldsymbol{\gamma}) == \sum_{i=1}^{n}[y_i(\boldsymbol{\gamma}\mathbf{Z}_i) - \log(1+e^{\gamma\mathbf{Z}_i})] \qquad (A.7)$$

Taylor's series tells us that an analytic function like (A.7) can be approximated as

$$\ell(\boldsymbol{\gamma}) \approx \ell(\boldsymbol{\gamma}^0) + (\boldsymbol{\gamma}-\boldsymbol{\gamma}^0)\ell'(\boldsymbol{\gamma}^0) + \frac{1}{2}(\boldsymbol{\gamma}-\boldsymbol{\gamma}^0)^2\ell''(\boldsymbol{\gamma}^0) \qquad (A.8)$$

where γ^0 is an estimated initial value of γ^0. To maximize $\ell(\boldsymbol{\gamma})$ we can differentiate with respect to γ and solve for γ

$$\ell'(\boldsymbol{\gamma}) \approx \ell'(\boldsymbol{\gamma}^0) + (\boldsymbol{\gamma}-\boldsymbol{\gamma}^0)\ell''(\boldsymbol{\gamma}^0) = 0 \qquad (A.9)$$

$$\rightarrow \boldsymbol{\gamma} = \boldsymbol{\gamma}^0 - \frac{\ell'(\boldsymbol{\gamma}^0)}{\ell''(\boldsymbol{\gamma}^0)} \qquad (A.10)$$

This suggests that we can start with an initial $\boldsymbol{\gamma}^0$ and iteratively apply (A.10) until the algorithm reaches convergence, at which point $\ell''(\boldsymbol{\gamma}^0) = 0$ and (A.10) does not change. This is what called Newton optimization and in the linear modeling setting is a vector. Newton's method has a generalization (Newton-Raphson) using the multivariate Taylor's series.

$$\boldsymbol{\gamma} = \boldsymbol{\gamma}^0 - [\frac{\partial^2}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^T}\ell(\boldsymbol{\gamma}^0)]^{-1}\frac{\partial\ell(\boldsymbol{\gamma}^0)}{\partial\boldsymbol{\gamma}} \qquad (A.11)$$

where $\frac{\partial^2}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^T}\ell(\boldsymbol{\gamma})$ is the matrix of second derivates and $\frac{\partial\ell(\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}}$ is the vector of first derivates.

The logistic log-likelihood for linear model becomes

$$\ell(\boldsymbol{\gamma}) = \sum_{i=1}^{n} y_i\mathbf{x}_i^T\boldsymbol{\gamma} - \log(1+e^{\mathbf{x}_i^T\gamma}) \qquad (A.12)$$

$$\begin{aligned}\frac{\partial}{\partial(\boldsymbol{\gamma})}\ell(\boldsymbol{\gamma}) &= \sum_{i=1}^{n} y_i\mathbf{x}_i^T - (1+e^{\mathbf{x}_i^T\gamma})^{-1}e^{\mathbf{x}_i^T\gamma}\mathbf{x}_i^T\\ &= \sum_{i=1}^{n}(y_i - \frac{1}{1-e^{-\mathbf{x}_i^T\gamma}})\mathbf{x}_i^T\\ &= \mathbf{X}^T(\mathbf{y}-\boldsymbol{\pi})\end{aligned} \qquad (A.13)$$

$$\begin{aligned}\frac{\partial^2}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^T}\ell(\boldsymbol{\gamma}) &= -\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^T\pi_i(1-\pi_i)\\ &= -\mathbf{X}^T\mathbf{V}\mathbf{X}\end{aligned} \qquad (A.14)$$

where $\mathbf{V}$ is a diagonal matrix with element $W(i, i)$ equal to $\pi_i(1-\pi_i)$. We can plug these results into (A.11)

$$\begin{aligned}\boldsymbol{\gamma} &= \boldsymbol{\gamma}^0 + (\mathbf{X}^T\mathbf{V}\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{y}-\boldsymbol{\pi})\\ &= (\mathbf{X}^T\mathbf{V}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\mathbf{X}\boldsymbol{\gamma}^0 + \mathbf{V}^{-1}(\mathbf{y}-\boldsymbol{\pi}))\\ &= (\mathbf{X}^T\mathbf{V}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}\mathbf{g}\end{aligned} \qquad (A.15)$$

where $\mathbf{g} = \mathbf{X}\boldsymbol{\gamma}^0 + \mathbf{V}^{-1}(\mathbf{y}-\boldsymbol{\pi})$. This process is called iteratively reweighted least squares (IRLS).

J. Biomedical Science and Engineering, 2010, 3, 576-583 **JBiSE**
doi:10.4236/jbise.2010.36080 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

A robust system for melanoma diagnosis using heterogeneous image databases

Khaled Taouil[1], Zied Chtourou[1], Nadra Ben Romdhane[2]

[1]UR CMERP National Engineering School of Sfax, Sfax, Tunisia;
[2]Faculty of Sciences of Sfax, Sfax, Tunisia.
Email: Khaled.taouil@cmerp.net

Received 5 April 2010; revised 8 May 2010; accepted 15 May 2010.

ABSTRACT

Early diagnosis of melanoma is essential for the fight against this skin cancer. Many melanoma detection systems have been developed in recent years. The growth of interest in telemedicine pushes for the development of offsite CADs. These tools might be used by general physicians and dermatologists as a second advice on submission of skin lesion slides via internet. They also can be used for indexation in medical content image base retrieval. A key issue inherent to these CADs is non-heterogeneity of databases obtained with different apparatuses and acquisition techniques and conditions. We hereafter address the problem of training database heterogeneity by developing a robust methodology for analysis and decision that deals with this problem by accurate choice of features according to the relevance of their discriminative attributes for neural network classification. The digitized lesion image is first of all segmented using a hybrid approach based on morphological treatments and active contours. Then, clinical descriptions of malignancy signs are quantified in a set of features that summarize the geometric and photometric features of the lesion. Sequential forward selection (SFS) method is applied to this set to select the most relevant features. A general regression network (GRNN) is then used for the classification of lesions. We tested this approach with color skin lesion images from digitized slides data base selected by expert dermatologists from the hospital "CHU de Rouen-France" and from the hospital "CHU Hédi Chaker de Sfax-Tunisia". The performance of the system is assessed using the index area (Az) of the ROC curve (Receiver Operating Characteristic curve). The classification permitted to have an Az score of 89,10%.

Keywords: Melanoma; Computer Aided Diagnosis System; Segmentation; Feature Selection; Classification; Generalized Regression Neural Network

1. INTRODUCTION

Melanoma is the most deadly form of skin cancer. The World Health Organization estimates that more than 65000 people a year worldwide die from too much sun, mostly from malignant skin cancer [1].

The five-year survival rate for people whose melanoma is detected and treated before it spreads to the lymph nodes is 99 percent. Five-year survival rates for regional and distant stage melanomas are 65 percent and 15 percent, respectively [2]. Thus the curability of this type of skin cancer depends essentially on its early diagnosis and excision.

The ABCD (asymmetry, border, colour and dimension) clinical rule is commonly used by dermatologists in visual examination and detection of early melanoma [3]. The visual recognition by clinical inspection of the lesions by dermatologists is 75% [4]. Experienced ones with specific training can reach a recognition rate of 80% [5].

Several works has been done on translating knowledge of expert physicians into a computer program. Computer-aided diagnosis (CAD) systems were introduced since 1987 [6]. It has been proved that such CAD systems can improve the recognition rate of the nature of a suspect lesion particularly in medical centres with no experience in the field of pigmented skin lesions [7,8]. For these systems to be efficient, the shots of the suspected lesion have to be taken using the same type of apparatuses than the one used for the learning database [9] and with identical lighting and exposure conditions. This could become very challenging in the majority of cases.

In order to overcome the lack of standardization in stand alone CADs and to provide an open access to dermatologists, web-based melanoma screening systems were proposed [10,11]. These systems have to consider the heterogeneity in databases collected in different cen-

tres.

This work describes an enhanced CAD system that addresses the problem of robustness of such tools under the use of different databases.

2. PROPOSED CAD SYSTEM

The proposed software combines automated image segmentation and classification procedures and is designed to be used by dermatologists as a complete integrated dermatological analysis tool. CAD systems in melanoma detection are usually based on image processing and data classification techniques. Five steps are generally needed: data acquisition, pre-processing, segmentation, feature extraction and classification (**Figure 1**).

2.1. Data Acquisition

The main techniques used for this purpose are the epiluminence microscopy (ELM, or dermoscopy), transmission electron microscopy (TEM), and the image acquisition using still or video cameras. The use of commercially available photographic cameras is also quite common in skin lesion inspection systems, particularly for telemedicine purposes [12].

2.2. Segmentation of Lesion Images

Image segmentation is the most critical step in the entire process. It consists of the extraction of the region of interest (ROI) which is the lesion. The result of segmentation is a mask image. This mask is the base for the computation of several shape and colour features.

The computer has a great difficulty in finding lesion edge accurately. This task alone has formed the basis of much research [13]. The difficulty of segmentation is due to low contrast between the lesion and the surrounding skin and irregular and fuzzy lesion borders. Artefacts (light reflections, shadows, overlapping hair, etc) can also give a false segmentation result. Some works rely on the physician to outline the suspicious area [14].

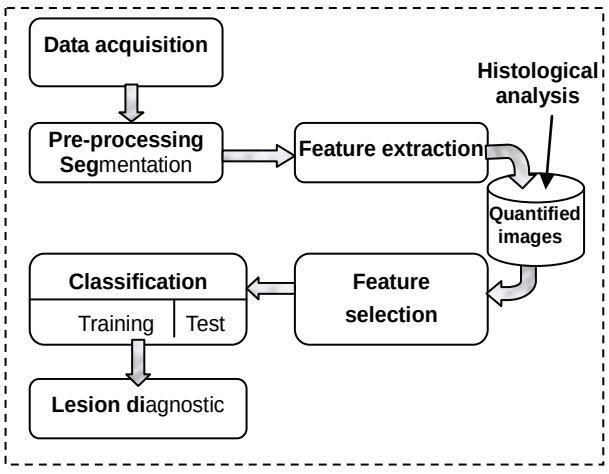


Figure 1. CAD system in melanoma detection.

We use a hybrid segmentation approach based on two steps. The first consists in applying morphological pre-processing filters to facilitate the extraction of the approximate region of the lesion from the safe skin. The second consists in applying active contour method on the approximate mask to have the final contour of the lesion [15].

Active contours or snakes are curves defined within an image domain that can move under the influence of internal forces coming from within the curve itself and external forces computed from the image data. Snakes were introduced by Kass *et al.* [16]. Snakes are parameterized curves:

$$v(s) = [x(s), y(s)], s \in [0,1] \quad (1)$$

This curves move through the spatial domain of an image to minimize the functional energy [17]:

$$E_{snake} = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s))\, ds \quad (2)$$

$$E_{int}(v(s)) = \alpha \left| v'(s) \right|^2 + \beta \left| v''(s) \right|^2 \quad (3)$$

$$E_{ext}(x, y) = -\left| \nabla I(x, y) \right|^2 \quad (4)$$

where:

- v(s) is a set of coordinates to form a snake contour.
- v'(s) and v"(s) denote the first and second derivatives of v(s) with respect to s .
- α and β are weighting parameters that control respectively the snake's tension and rigidity.
- $\nabla I(x, y)$ is the gradient of grey-level image I.

A snake that minimizes E_{snake} must satisfy the Euler equation.

$$\alpha v''(s) - \beta v''''(s) - \nabla E_{ext} = 0 \quad (5)$$

The internal force F_{int} discourages stretching and bending while the external potential force F_{ext} pulls the snake toward the desired image edges.

To find a solution to (4), the snake v(s) is made dynamic by adding the parameter of time t to the equation of the curve that becomes:

$$v_t(s,t) = \alpha v''(s,t) - \beta v''''(s,t) - \nabla E_{ext} \quad (6)$$

indicating how the snake must be modified at the instant t+1 according to its position at the instant t.

When v(s, t) stabilizes, we achieve a solution of (6).

2.3. Features Extraction from Lesion Images

To characterize the different types of lesions we consider a parametric approach. In such approach, the skin lesion is resumed in a vector of features which dimension depends on the number of extracted primitives. We use

quantitative parameters from the descriptions of dermatologists based on the ABCD rule to model clinical signs of malignancy.

The preliminary developed set of parameters underwent a series of tests to evaluate their robustness when quantifying multiple shots of the same lesion acquired under different lighting conditions with different apparatuses as is the case when using different slides from heterogeneous databases [18]. Our set of parameters was besides used by [19] to develop an automatic recognition of melanoma which reached a correct classification rate of 79.1%.

The most important clinic signs that were kept to characterize melanoma and the different lesions are the irregularity of the contour, the asymmetry of colour and shape, as well as the heterogeneity of the colour. We classify parameters in two categories: geometric and photometric parameters.

2.3.1. Geometric Parameters

Geometric parameters are extracted from the binary shapes obtained after image segmentation. These parameters permit to characterize the shape of the lesion, its elongation and the regularity of its contour. All these parameters are standardized and independent of translation, zoom and rotation effects and therefore compensate for rigid transformations introduced by the optics conditions and the scene selection and framing.

2.3.2. Photometric Parameters

Photometric parameters are calculated from true colour and binary images. These parameters permit to describe homogeneity and symmetry of the colour as well as the deviation between the mean colour of the lesion and the mean colour of the surrounding safe skin. We tested different colour spaces representations calculated from the red green blue components. We reflect the correlation between the level of a colour component of a pixel and its position on the digital image witch is supposed to be independent of lighting conditions and spectral ensitiveity of the camera sensor.

2.4. Feature Selection

Feature selection allows choosing the most relevant parameter subset to perform the classification step. This subset must contain the more robust and the most discriminative primitives [20]. Three criteria's must be fixed: the assessment method of the variables set relevance, the research procedure to follow and the stopping criteria of the selection. In this work we report the use of a sequential forward selection with a stopping criteria based on the minimum error generated by the classifier.

2.4.1. Sequential Forward Selection

The SFS [21] is an ascending research method (bottom-up) of the set of most discriminative parameters from an initial set of parameters (Ei) with:

$$Ei = \{pj,\ j = 1, 2,\ \ldots,\ N\}$$

$$E_{SFSn} \subset Ei\ ,\ n \le N$$

For this method one parameter pj is added at a time to the E_{SFS} subset. If the assessment criterion is an artificial neural network, for each step, we insert one by one remaining parameters pj of E_i in E_{SFS} and we calculate the corresponding classification error (Err) with:

$$Err = \frac{1}{q}\sum_{i=1}^{q}(d_i - a_i)^2 \qquad (7)$$

with q: total image number of the training database.

di: desired output.

ai: real output.

Initially, the subset $E_{SFS0} = \varphi$;

For every step, parameter pj that will be selected is the one for which the new E_{SFS} subset permits to minimize the classification error:

$$Err(E_{SFS}U_{pj}) \le Err(E_{SFS}U_{pj}') \qquad (8)$$

Thus, the first selected parameter is the most discriminative one of the initial set of parameters. The selection of parameters stops when while adding a new parameter to E_{SFS}, the classification error increases.

2.5. Classification of Lesion Images

After having summarized information contained in the different images of our databases in vectors of parameters we use a classifier based on a general regression network (GRNN) [22]. GRNN network is much faster to train than a multilayer perceptron network (MPN). GRNN gave better recognition rates than MPN for melanoma classification [23]. The architecture of this network is illustrated on **Figure 2**. The GRNN is composed of four layers: the input layer, the first intermediate

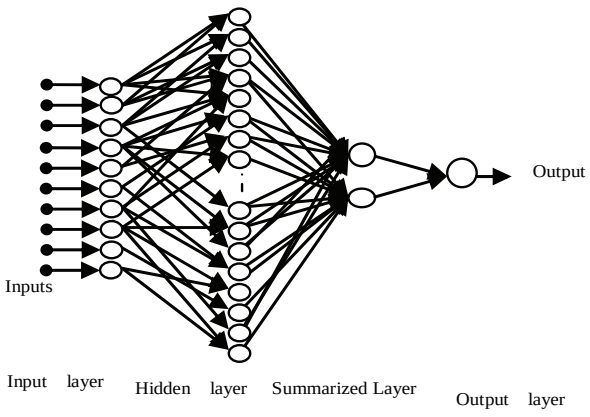


Figure 2. Architecture of the used GRNN.

layer constituted of radial units, the second intermediate layer constituted of summarized units and the output layer.

3. ASSESSMENT OF THE CLASSIFICATION

The performance of our CAD system is evaluated in term of sensitivity and specificity. These measures are defined as follow:

$$\text{Sensitivity} = \frac{\#\text{True Positives}}{\#\text{True Positives} + \#\text{False Negatives}} \quad (9)$$

$$\text{Specificity} = \frac{\#\text{True Negatives}}{\#\text{True Negatives} + \#\text{False Positives}} \quad (10)$$

With #true-Positive and #False-Negative corresponds respectively to the number of malign lesions well classified and badly classified. #True-Negative and #False-Positive corresponds respectively to the number of benign lesions well classified and badly classified.

A ROC curve consists in representing the value of the sensitivity according to (1-specificity) [24]. The area under the ROC curve or area index (Az) represents the probability to correctly identify the image with anomaly when an image with anomaly and an image without are presented simultaneously to the observer.

We use two databases of image lesions whose malign or benign nature is perfectly known after histological analysis. The first one has been collected in CHU Rouen France with the collaboration of the research laboratory PSI-INSA Rouen and has been supported by the French National League against Cancer. This database was digitized in true colours by a 35 mms slides Nikon LS-1000S scanner. It was used in previous works [19,25] and [26,27]. We divide this database in a training (B0) and test (B1) sets used in the first assessment of classification.

The second image database has been collected in Tunisia from the dermatology service of CHU Hédi CHAKER in Sfax. Images were digitized in true colours with a HP Scanjet 3570c scanner (cf. **Table 1**).

Our approach for efficiency assessment of the developed tool has been achieved in three steps:

3.1. First Assessment of the System

For the first step, we evaluate the diagnosis results of our system while using two sets of images (B0 for training and B1 for test) from the same CHU Rouen image database.

Table 1. Distribution of images databases.

Database	Set	Class	Nbr of BL	Nbr of ML	Total
CHU Rouen	B0	Training	59	27	86
	B1	Test	34	15	49
CHU Sfax	B2	Test	31	20	51
	Total		124	62	186

3.2. Comparison of Our System's Diagnosis with Dermatologist's Visual Diagnosis

This step consists in comparing diagnostic results of our system with the opinion of four expert dermatologists for the same test database (B1). Dermatologists are part of the dermatology service of the CHU Sfax. We asked every dermatologist to give his diagnosis for each lesion.

3.3. Third Assessment of the System

For this step, we evaluate our system while using the second image database (B2) collected at CHU Sfax. This test has been done while using the artificial neural network having the best recognition rate according to the first assessment.

4. RESULTS AND DISCUSSION

4.1. Results of the Segmentation Step

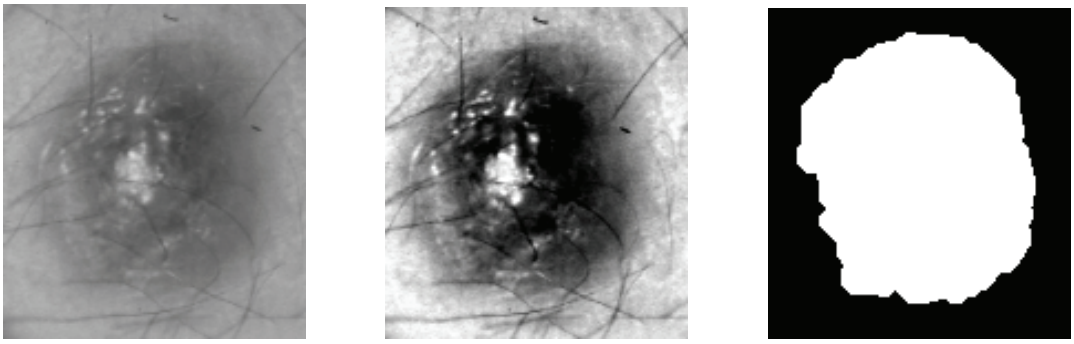
For image lesion segmentation, we propose a hybrid method that combines the advantages of morphological treatments, histogram thresholding and active contours techniques.

First the contrast of the gray level original image is enhanced using top-hat and bottom-up filtering (**Figure 3**).

The extraction of the image mask is based on the detection of regional minima of the complementary image of the contrast enhanced image. The detection of these regions requires the application of a threshold. This threshold is obtained with histogram thresholding using Otsu method [28]. We apply then a morphological opening on the obtained image. Lakes of the resulting image are eliminated by filling holes. The approximate zone or approximate mask of the lesion is finally obtained following a labelling, conservation of the biggest element.

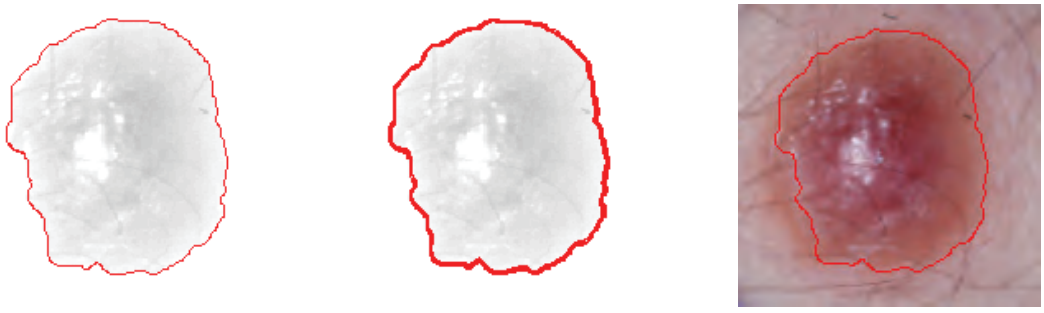
We initialize a snake at the approximate boundary of the safe skin (cf. **Figure 4**). The snake begins with the calculation of a field of external forces over the image domain. The forces drive it toward the boundary of the lesion. The process is iterated until it matches the contour of the lesion. We superpose the obtained contour on the original color image.

Figures 5 and **6** show some examples of segmentation results of lesions collected from both CHU Rouen and CHU Sfax databases.



(a) original image (b) contrast enhancement (c) Approximate mask of the safe skin

Figure 3. Extracting the approximate mask of the safe skin.



(a) Snake initialization (b) Snake progression (c) segmented image

Figure 4. Application of the active contour on the approximate mask.

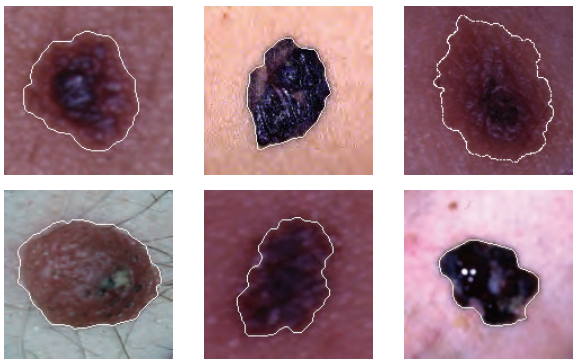


Figure 5. Segmentation results of lesions collected from CHU Rouen database.

4.2. Results of Features Selection

For features extraction, a set of 68 parameters are extracted for every lesion. Through correlation and robustness study, a set of 42 parameters have been kept. To find the most discriminative ones for the classification step, we apply the sequential forward selection (SFS) method. Training and test databases of images have been randomly selected.

For the SFS method, the assessment of parameter selection is based on the comparison of the error generated by the general regression neural network (GRNN) for the different subsets of selected variables.

The chosen set of variables is the one that generates the minimal error. We pursued research until the selection of all parameters. Then we chose the smallest subset gotten with the minimal error.

The result of this selection method is illustrated on **Figure 7**. It illustrates the variation of the classification's mean square error (MSE) according to the number of included parameters, during trainings and tests of the GRNN.

According to the test curve, we note that the set of the

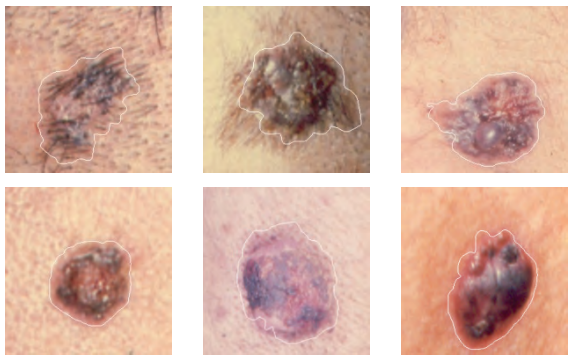


Figure 6. Segmentation results of lesions collected from CHU Sfax database.

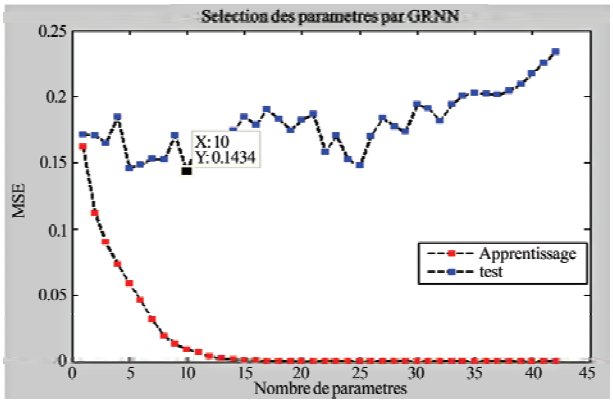


Figure 7. Parameter selection using SFS method.

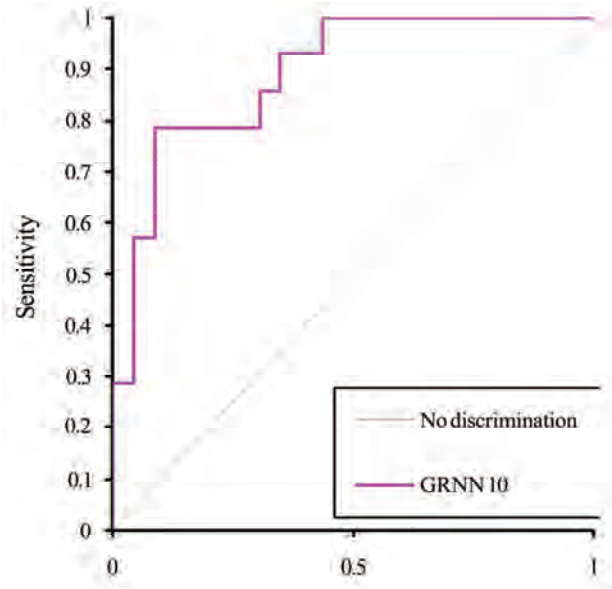


Figure 8. The ROC curve obtained using GRNN classifier and the selected parameters.

Table 2. Order of the selected parameters using SFS method.

Rg	N°	Description	
01	24	scG	symmetry of the green component
02	09	rmoyr	mean of the normalized red in the lesion
03	03	DeltaD	expanse of the distance to the center
04	35	Beta	spherical coordinate beta
05	23	scR	symmetry of the red component
06	01	comp	compactness of the shape
07	39	RB	Ratio between lesion and safe skin for the red
08	29	gamma1B	Fisher coefficient for the blue
09	16	sigmamoyLb	Standard deviation of normalized bleu in the safe skin
10	06	scB	symmetry of the blue component

most discriminative parameters permitting to minimize the MSE of test database are gotten after the insertion of the first ten parameters (MSE = 0.1434).

The selection by the SFS method permitted a reduction of 76.19% of the total number of parameters. The list of parameters selected is presented in **Table 2**.

4.3. Results of the Classification Step

The classification is based on parameters kept in features selection. **Figure 8** illustrates results of the classification while using the 10 most discriminative parameters selected by the SFS method.

The performances of classification using B1 set of images extracted from CHU Rouen database are based on the comparison of the value of the area index (Az) of the ROC curve. We obtain a value of Az equal to 89.1%. The recognition rate of the system is 92.05%.

To validate the efficiency of our system, we compare the obtained classification results with the diagnosis of four dermatologists from CHU Sfax. We asked every dermatologist to give his diagnosis for every lesion of the same test set B1. Results of their diagnoses are given in **Table 3**.

According to these results, we notice that the mean value of sensitivity provided by dermatologists is equal to the mean percentage of visual recognition of the true positives by dermatologists that is 75% [4].

The recognition rate of our CAD system (92.05%) exceeds that obtained by dermatologists. It even exceeds that of experimented trained ones which is about 80% [5].

Results of the second assessment of the system are given in **Table 4**. We note that even when using a test set of image lesions selected randomly from a different database the recognition rate our system is 90.15%. It remains better than the one of the visual diagnosis of experimented dermatologists.

5. CONCLUSIONS

In this paper we have described the different steps used in the CAD system that we propose for melanoma detection. To make this tool useful by the dermatologist community outside specialized centres, each stage of processing had to be automatic and robust to different conditions of acquisition and apparatus. The system segments and extracts parameters of description of the lesion. These parameters are normalized and used as inputs for the neural network classifier which decides if the lesion is suspicious. We have also described the different steps used for the evaluation of our system. This evaluation had proved the robustness of our system when using different databases in training and test. This property makes it a suitable and an efficient candidate for use in a context of a telemedicine dermatological application.

Table 3. Recognition rate dermatologists (D).

RR / D	D A	D B	D C	D D	mean
FP	12	0	12	8	
VP	9	8	14	14	
FN	6	7	1	1	
VN	22	34	22	26	
Sensitivity	60	53,3	93,3	93,3	74,9
Specificity	64,7	100	6 4,7	76,5	76,4
RR	62,3	76,6	79,0	84,9	**75,7**

- FP (False Positive)
- TP (True Positive)
- FN (False Negative)
- TN (True Negative)
- RR: Recognition Rate

Table 4. Assessment of the system with database B2.

TR	Base2
FP	03
VP	18
FN	02
VN	28
Sensitivity	90
Specificity	90,3
Recognition rate	90,15

6. ACKNOWLEDGEMENTS

We sincerely thank:

- Dermatology Services of CHU Rouen, France and CHU Hédi Chaker, sfax, Tunisia.

- Laboratory PSI (Perseption Systems Information), INSA Rouen, France.

- Research Unit: Sciences and Technologies of Image and Telecommunications, High Institute of Biotechnology, Sfax.

REFERENCES

[1] (2006) Solar ultraviolet radiation: Global burden of disease from solar ultraviolet radiation. Environmental Burden of Disease Series, World Health Organization, **13**.

[2] (2009) Cancer Facts and Figures. American Cancer Society, Atlanta.

[3] Friedman, R.J. (1985) Early detection of malignant melanoma: the role of the physician examination and self examination of the skin. *Cancer Journal for Clinicians*, **35**, 130-151.

[4] Schmid-Saugeon, P., Guillod, J. and Thiran, J.P. (2003) Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*, **27**, 65-78.

[5] Blanzieri, E. (2000) Exploiting classifier combination for early melanoma diagnosis support. *Proceedings of the* 11*th European Conference on Machine Learning*, Barcelona, 31 May-2 June 2000, 55-62.

[6] Cascinelli, N., Ferrario, M., Tonelli, T., *et al.* (1987) A possible new tool for clinical diagnosis of melanoma. *Journal of the American Academy of Dermatology*, **16**, 361-367.

[7] Ercal, F., Chawla, A., Stoecker, W.V., Lee, H.C. and Moss, R.H. (1994) Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on Biomedical Engineering*, **41(9)**, 837-845.

[8] Ganster, H., Pinz, A., Röhrer, R., *et al.* (2001) Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, **20(3)**, 233-238.

[9] Burroni, M. (2004) Melanoma computer-aided diagnosis: Reliability and feasibility study. *Clinical Cancer Research*, **10**, 1881-1886.

[10] Guillod, J., Schmid-Saugeon, P., Déccaillet, F., *et al.* (2003) An open internet platform to distributed image processing applied to dermoscopy. *Studies in Health Technology and Informatics*, **95**, 107-112.

[11] Oka, H., Hashimoto, M., Iyatomi, H. and Tanaka, M. (2004) Internet-based program for automatic discrimination of dermoscopic images between melanoma and Clark nevi. *British Journal of Dermatology*, **150(5)**, 1041.

[12] Maglogiannis, I. and Charalampos, N.D. (2003) Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transaction on Information Technology in Biomedicine*, **13(5)**, 721-733.

[13] Hall, P.N., Claridge, E. and Morris Smith, J.D. (1995) Computer screening for early detection of melanoma—Is there a future? *British Journal of Dermatology*, **132**, 325-338.

[14] White, R., Rigel, D.S. and Friedman, R.J. (1991) Computer applications in the diagnosis and prognosis of malignant melanoma. *Dermatologic Clinics*, **9**, 695-702.

[15] Taouil, K. and Romdhane, N.B. (2006) A new automatic approach for edge detection of skin lesion images. *International Conference on Information & Communication Technologies*: *From Theory to Applications*, Damascus.

[16] Kass, M., Witkin, A. and Terzopoulos, D. (1987) Snakes: Active contour models. *International Journal of Computer Vision*, **1**, 321-331.

[17] Xu, C.Y. and Prince, J.L. (1998) Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, **7(3)**, 359.

[18] Taouil, K., Bouhlel, M.S., Elloumi, M. and Kamoun, L. (2002) Quantification des caractéristiques de mélanomes en vue d'une classification. JTEA'2002, Sousse, 179-185.

[19] Zagrouba, E. and Barhoumi, W. (2004) A prelimary approach for the automated recognition of malignant melanoma. *Image Analysis and Stereology Journal*, **23(2)**, 121-135.

[20] Taouil, K. and Romdhane, N.B. (2006) Automatic segmentation and classification of skin lesion images. *The* 2*nd International Conferences on Distributed Frameworks for Multimedica Applications*, Pulau Pinang.

[21] Pudil, P., Novoviovca, J. and Kittler, J. (1994) Floating search methods in feature selection. *Pattern Recognition Letters*, **15**, 1119-1125.

[22] Specht, D.F. (1991) A general regression neural network. *IEEE Transactions on Neural Networks*, **2**, 568-576.

[23] Romdhane, N.B., Taouil, K., Boudaya, S. Turki, H. and Bouhlel, M.S. (2007) Sélection des Variables et Classification par Réseaux de Neurones des Lésions Derma-

tologiques. 4*th International Conference*: *Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia.

[24] Tokan, F., Türker, N. and Yıldırım, T. (2006) ROC analysis as a useful tool for performance evaluation of artificial neural networks. *Artificial Neural Networks — ICANN* 2006, **4132**, 923-931.

[25] Zagrouba, E. and Barhoumi, W. (2005) An accelerated system for melanoma diagnosis based on subset feature selection. *Journal of Computing and Information Technology*, **13(1)**, 69-82.

[26] Vannoorenberghe, P., Colot, O. and de Brucq, D. (1999) Dempster-shafer's theory as an aid to color information processing application to melanoma detection in dermatology. *Proceedings of the* 10*th International Conference on Image Analysis and Processing*, Venice, 774-779.

[27] Lefevre, E., Colot, O., Vannoorenberghe, P. and de Brucq, D. (2000) Knowledge modeling methods in the framework of evidence theory: An experimental comparison for melanoma detection Systems. 2000 *IEEE International Conference on Man and Cybernetics*, Nashville, **4**, 2806- 2811.

[28] Sezgin, M. and Sankur, B. (2004) Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, **13(1)**, 146-165.

J. Biomedical Science and Engineering, 2010, 3, 584-601 **JBiSE**
doi:10.4236/jbise.2010.36081 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Novel host markers in the 2009 pandemic H1N1 influenza a virus

Wei Hu

Department of Computer Science Houghton College, Houghton, USA.
Email: wei.hu@houghton.edu

Received 21 March 2010; revised 28 April 2010; accepted 8 May 2010.

ABSTRACT

The winter of 2009 witnessed the concurrent spread of 2009 pandemic H1N1 with 2009 seasonal H1N1. It is clinically important to develop knowledge of the key features of these two different viruses that make them unique. A robust pattern recognition technique, Random Forests, was employed to uncover essential amino acid markers to differentiate the two viruses. Some of these markers were also part of the previously discovered genomic signature that separate avian or swine from human viruses. Much research to date in search of host markers in 2009 pandemic H1N1 has been primarily limited in the context of traditional markers of avian-human or swine-human host shifts. However, many of the molecular markers for adaptation to human hosts or to the emergence of a pandemic virus do not exist in 2009 pandemic H1N1, implying that other previously unrecognized molecular determinants are accountable for its capability to infect humans. The current study aimed to explore novel host markers in the proteins of 2009 pandemic H1N1 that were not present in those classical markers, thus providing fresh and unique insight into the adaptive genetic modifications that could lead to the generation of this new virus. Random Forests were used to find 18 such markers in HA, 15 in NA, 9 in PB2, 11 in PB1, 13 in PA, 10 in NS1, 1 in NS2, 11 in NP, 3 in M1, and 1 in M2. The amino acids at many of these novel sites in 2009 pandemic H1N1 were distinct from those in avian, human, and swine viruses that were identical at these positions, reflecting the uniqueness of these novel sites.

Keywords: 2009 Pandemic H1N1; Host Switch; Influenza; Mutation; Random Forests

1. INTRODUCTION

In addition to the common seasonal H1N1 influenza virus, an antigenically novel swine-origin pandemic H1N1 influenza virus marked the flu season in 2009. It is likely that both 2009 pandemic H1N1 and seasonal influenza will coexist for some time. Elucidation of the characteristics of this new virus has become an important part of the current flu research. The identification of molecular markers for drug resistance, virulence, viral transmission and replication, human adaptation, and evolution can shed new light into the nature of this virus.

There are eight single-stranded RNA segments of the influenza A virus genome. They code 11 proteins: hemagglutinin (HA), neuraminidase (NA), matrix 1 (M1), matrix 2 (M2), nucleoprotein (NP), non-structural protein 1 (NS1), non-structural protein 2 (NS2; also termed nuclear export protein, NEP), polymerase acidic protein (PA), polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2), and polymerase basic protein 1 – F2 (PB1-F2). Segments 1, 3, 4, 5, and 6 each encode a single protein, *i.e.*, PB2, PA, HA, NP, and NA, respectively, whereas segments 2, 7, and 8 each encode two proteins, *i.e.*, PB1 and PB1-F2, M1 and M2, NS1 and NS2, respectively. The life cycle of influenza virus has the following steps with several proteins involved in each: entry into the host cell (HA, M1 and M2), entry of viral ribonucleoproteins (vRNP) into the nucleus (NP, PA, PB1 and PB2), transcription and replication of the viral genome (PA, PB1, PB2, NS1, and NP), export of the vRNPs from the nucleus (NP, NS2 and M1), and assembly and budding at the host cell plasma membrane (HA, NA, M1 and M2) [1].

Besides mutations, viruses with segmented genomes can generate genetic diversity by exchanging gene segments between different viruses to produce a new virus. Comprehensive phylogenetic analysis suggested that the genes of 2009 pandemic H1N1 were derived from avian (PB2 and PA), human H3N2 (PB1), classical swine (HA, NP and NS), and Eurasian avian-like swine H1N1 (NA and M) lineages [2].

The symptoms of the 2009 pandemic H1N1 flu are

similar to the 2009 seasonal flu with a possibility of additional symptoms such as vomiting and diarrhea [3]. The Center for Disease Control and Prevention (CDC) and Mayo Clinic have developed several molecular tests to detect and discriminate the novel 2009 pandemic H1N1 virus and the 2009 seasonal virus [4]. The matrix (M) gene is highly conserved compared to other gene segments, which makes it an ideal target for RT-PCR assays used to detect the presence of influenza. A mutation in the M gene of the 2009 pandemic virus could invalidate these tests [5].

Sequence survey suggested that there were two distinct evolutionary trends in antigenic drift of H1N1 HAs at two residues 190 and 225. The epidemic H1N1 HAs favor position 190 while the 1918 pandemic and swine HAs favor position 225 [6]. In contrast to these two trends, the 2009 pandemic H1N1 strains are highly conserved at both HA 190 and 225 and possess the signature markers Asp190 and Asp225 that are known to confer specificity to the human α2-6 sialylated glycan receptors [7]. Further analysis indicated the 2009 pandemic H1N1 HAs possess residues that can be positioned to bind to avian α2-3 sialylated glycan receptors as well [7,8]. By homology modeling of the HA structure, the antigenic similarity between the 1918 H1N1 and the pandemic 2009 H1N1 viruses was confirmed, and the future amino acid substitutions on the antigenic sites of 2009 pandemic H1N1 HA were also predicted [9], raising the concerns that these two pandemic H1N1 viruses may share a similar evolutionary path. With informational spectrum method [10], a bioinformatics technique, highly conserved domains and mutations in the 2009 pandemic H1N1 HAs were identified and the contributions of these mutations to the changes of binding specificity of the 2009 pandemic H1N1 HAs were quantified [11-13].

Many 2009 seasonal H1N1 strains carry a NA mutation H275Y that confers high-level resistance to oseltamivir. Although most 2009 pandemic H1N1 strains are susceptible to oseltamivir, the co-circulation of pandemic and seasonal H1N1 viruses might provide opportunities for 2009 pandemic H1N1 to develop oseltamivir resistance through mutations and reassortments between pandemic and seasonal H1N1 viruses. Positive natural selection was detected in the NA proteins of 2009 pandemic H1N1 at codons 275 and 248 and seasonal H1N1 at codon 275, with statistically significant bias of nonsynonymous mutations relative to synonymous mutations [14]. Besides position 275, mutations at other positions in NA such as 116, 117, 119, 136, 150, 151, 199, 223, 275, and 295 could also alter NA inhibitor susceptibility [15,16].

Two recent reports [17,18] revealed three NA variant groups in 2009 pandemic H1N1. The first group had V106 and N248, the second included I106 and N248, and the third contained I106 and D248, highlighting the rapid genetic variation of this surface antigen under host immune pressure and the need for close monitoring. The NA protein of the avian viruses has, in addition to the catalytic site, a separate sialic acid binding site that is not present in human viruses, which could enhance the catalytic efficiency of NA [19]. Although the second binding site was not conserved in swine NA strains, a recent report found the 2009 pandemic H1N1 strain of swine origin appeared to have retained some of the key features of the second binding site. Their data showed possible lowered HA activity for this second site, which might be an important event in the emergence of the 2009 pandemic strain [20].

The interaction of NP and the influenza polymerase, containing the PA, PB1 and PB2 proteins, catalyses viral RNA replication (vRNA→cRNA→vRNA) and transcription (vRNA→mRNA) in the nucleus of infected cells. The PB2 protein of human viruses tend to possess a lysine at position 627 (K627), whereas avian viruses generally have glutamic acid at this position (E627). The mutation E627K allows avian virus to efficiently grow in humans and was identified experimentally as a crucial host range and pathogenicity determinant [21,22]. The 2009 pandemic H1N1 strains could transmit in humans efficiently, but exclusively possess the avian signature E627. Therefore, there might be alternative strategies employed by the novel 2009 H1N1 polymerase to maintain the efficient replication rate. A recent study discovered that serine at position 590 (S590) and arginine at position 591 (R591) might serve as a regulator of polymerase activity that contributes to the increased replication efficiency of 2009 pandemic H1N1. The paired mutations S590 and R591, termed the SR polymorphism, were present in only three of the 2849 PB2 sequences of human viruses before 2009 [23]. Other sites might affect the polymerase activity as well. The mutations at position 504 in PB2 (I504V) and position 550 in PA (I550L) could result in enhanced virulence [24]. A special region (residues 360-374) in the NP protein was found to play a vital role in overcoming species barrier for 2009 pandemic H1N1 [25].

Compared to other proteins in the influenza viruses, PB1-F2 is a newly discovered protein, which is unique in that this protein is coded by a subset of the nucleotides that code for PB1 due to the use of a different reading frame. The PB1-F1 protein has been implicated in pathogenicity and the induction of cell death [26-28]. The 2009 pandemic H1N1 virus has a truncated PB1-F1 protein, because its genome contains three stop codons preventing PB1-F2 expression. Recently, studies found that its function is not universal, but cell type and virus

strain dependent [29], and it plays a critical role in the pathogenicity and tranmissibility of 2009 pandemic H1N1 [30,31].

NS1 is a multifunctional protein that contributes to viral pathogenesis by neutralizing the interferon (IFN) -based defense system of the host cell [32], and serves as a strong inducer of apoptosis in infected human respiratory epithelial cells [33]. The NS1 protein of 2009 pandemic H1N1 is truncated and therefore missing a domain responsible for increased pathogenicity of avian virus [34]. There are two molecular markers that account for the virulence of the highly pathogenic avian H5N1 viruses are not present in 2009 pandemic H1N1. They are a lysine (K) at position 627 of PB2, and glutamic (E) acid in position 92 of NS1 that might increase the replication efficiency and block host inhibition of viral replication, respectively [35,36].

A recent study on M gene identified sites of high selective pressure between human and avian influenza, which were 115, 121, 137 in M1, and 11, 16, 20, 54, 57, 78, 86, and 93 in M2 [37]. The 2009 pandemic H1N1 virus contains the adamantine-resistant mutation S31N in its M2 protein, thus making the NA Inhibitors oseltamivir and zanamivir the only options available to treat the infections caused by the pandemic virus [38].

Several studies focused on determining which amino acid changes best distinguish an avian or swine influenza virus from a human virus. An entropy analysis revealed the human-avian host shift genomic signature of 52 markers in ten proteins of the influenza virus [39]. This signature extended in [40] provided the basis for finding the amino acids of 2009 pandemic H1N1 at the host species-specific positions to illustrate the adaptive mutations of this virus. By comparison of the protein sequences of 2009 pandemic H1N1 with those in the previous pandemics and human, swine, and avian influenza viruses, the mutation trend of the residues at the signature positions was discovered, and the potential roles of the mutated residues in human adaptation and virulence was probed in [41]. With mutual information analysis, the characteristic sites for human-to-human transmission in PB2 of influenza viruses were uncovered [42], and subsequently a catalogue of 68 such sites in eight internal proteins were found to derive adaptation signatures of viral proteomes [43], which included many of the 32 and 34 markers identified in [44,45], respectively.

Many of the molecular determinants associated with adaptation to human hosts or to the emergence of a pandemic virus are not present in 2009 pandemic H1N1, suggesting that other previously unrecognized molecular markers are responsible for its ability to infect humans [17]. Therefore, uncovering new molecular features of 2009 pandemic H1N1 is of prime significance. In this study, we collected all the protein sequences of the 2009 pandemic H1N1, 2009 seasonal H1N1, avian, human, and swine influenza viruses available from the National Center for Biotechnology Information (NCBI). Our objective was to explore the novel host markers in 2009 pandemic H1N1 that were not present in the classical avian-human or swine-human host shift markers, and the top markers that could differentiate 2009 pandemic H1N1 from 2009 seasonal H1N1.

2. MATERIALS AND METHODS

2.1. Sequence Data

All influenza virus protein sequences were retrieved from the Influenza Virus Resource (http://www.ncbi/nlm.nih.giv/genomes/FLU/FLU.html) of the National Center for Biotechnology Information (NCBI). Detailed information about these sequences is in **Table 1**. All the sequences used in the study were aligned with MAFFT [46].

2.2. Random Forests

Random Forest, proposed by Leo Breiman in 1999 [47], is an ensemble classifier based on many decision trees. Each tree is built on a bootstrap sample from the original training set and is unpruned to obtain low-bias trees. The variables used for splitting the tree nodes are a random subset of the whole variable set. The classification decision of a new instance is made by majority voting over all trees. About one-third of the instances are left of the bootstrap sample and not used in the construction of the tree. These instances in the training set are called “out-of-bag” instances and are used to evaluate the performance of the classifier, which can achieve both low bias and low variance with bagging and randomization.

2.3. Feature Selection Using Random Forests

Random Forest calculates several measures of variable importance. The mean decrease in accuracy measure was employed in [48] to rank the importance of the features in prediction. This measure is based on the decrease of classification accuracy when values of a variable in a node of a tree are permuted randomly. In this study, two packages of R, randomForest and varSelRF [48], were utilized to compute the importance of the amino acids in a given protein sequence dataset. The effectiveness and robustness of this technique as a feature selection method has been demonstrated in various studies [49-54].

2.4. Procedure to Find Novel Host Sites in 2009 Pandemic H1N1

Four steps were created to locate the novel sites associated

Table 1. Counts of the influenza protein sequences used in the current study.

Host	Subtype	Protein	Number of Sequences	Years	Host	Subtype	Protein	Number of Sequences	Years
Human	Pandemic H1N1	HA	710	2009	Human	All types	NS1	509	All years
Human	Pandemic H1N1	NA	643	2009	Human	All types	NS2	383	All years
Human	Pandemic H1N1	NP	394	2009	Human	All types	PA	279	All years
Human	Pandemic H1N1	M1	490	2009	Human	All types	PB1	289	All years
Human	Pandemic H1N1	M2	482	2009	Human	All types	PB2	269	All years
Human	Pandemic H1N1	NS1	366	2009	Avian	H1	HA	120	All years
Human	Pandemic H1N1	NS2	358	2009	Avian	N1	NA	1821	All years
Human	Pandemic H1N1	PA	295	2009	Avian	All types	NP	2888	All years
Human	Pandemic H1N1	PB1	311	2009	Avian	All types	M1	4232	All years
Human	Pandemic H1N1	PB2	311	2009	Avian	All types	M2	3182	All years
Human	Seasonal H1N1	HA	128	2009	Avian	All types	NS1	4610	All years
Human	Seasonal H1N1	NA	125	2009	Avian	All types	NS2	3422	All years
Human	Seasonal H1N1	NP	25	2009	Avian	All types	PA	3106	All years
Human	Seasonal H1N1	M1	129	2009	Avian	All types	PB1	2979	All years
Human	Seasonal H1N1	M2	129	2009	Avian	All types	PB2	2643	All years
Human	Seasonal H1N1	NS1	25	2009	Swine	H1	HA	379	All years
Human	Seasonal H1N1	NS2	25	2009	Swine	N1	NA	278	All years
Human	Seasonal H1N1	PA	23	2009	Swine	All types	NP	420	All years
Human	Seasonal H1N1	PB1	25	2009	Swine	All types	M1	516	All years
Human	Seasonal H1N1	PB2	25	2009	Swine	All types	M2	406	All years
Human	H1	HA	640	All years	Swine	All types	NS1	506	All years
Human	N1	NA	1127	All years	Swine	All types	NS2	351	All years
Human	All types	NP	393	All years	Swine	All types	PA	343	All years
Human	All types	M1	1512	All years	Swine	All types	PB1	368	All years
Human	All types	M2	1415	All years	Swine	All types	PB2	327	All years

with host adaptation in 2009 pandemic H1N1.

Step 1: For each protein, the consensus sequence of avian, 2009 pandemic H1N1, human, and swine viruses were calculated separately, and the positions with different amino acids of the four consensus sequences were identified, since the different amino acids at these positions have the potential to contribute to host switches.

Step 2: For each protein, Random Forests were used to identify the top 20 positions that have highest importance in separating avian from human viruses, and swine from human viruses, respectively.

Step 3: Finding the intersection of the top positions with importance larger than 0.005 for separating 2009 pandemic H1N1 from human viruses and the positions with different consensus amino acids found in step one.

Step 4: The positions discovered in step three minus the positions found in step two will be the novel positions important for separating 2009 pandemic H1N1

Table 2. This table contains the Hamming distances of consensus protein sequences of avian, human, 2009 pandemic H1N1, and swine viruses.

Proteins	HA	NA	NP	M1	M2	NS1	NS2	PA	PB1	PB2
Dist(Avian,2009_pandemic)	90	35	28	8	6	35	11	17	23	17
Dist(Human,2009_pandemic)	109	61	42	19	15	42	13	30	21	31
Dist(Swine,2009_pandemic)	43	52	10	11	6	17	7	15	20	17
Dist(Avian,Human)	90	50	31	11	13	23	5	20	4	19
Dist(Avian,Swine)	61	37	21	3	2	22	7	4	6	2
Dist(Human,Swine)	97	52	35	8	11	36	7	19	6	21

from human viruses.

The purpose of steps 2 and 3 was to calculate the adaptation signatures for various virus groups, which were then used in step 4. The amino acids at most of these novel sites in 2009 pandemic H1N1 turned out to be different from those in avian, human, and swine viruses that were the same at these positions.

Random Forests produce non-deterministic outcomes. To compensate this bias, the Random Forests algorithm was run multiple times and then the average of the results was taken. The importance of each residue in the protein sequences was based on the averaged calculations by using the function randomVarImpsRF in varSelRF repeated 5 times.

3. RESULTS

3.1. Comparison of Consensus Protein Sequences of Influenza Viruses

In considering the relationship among the proteins of influenza viruses, the Hamming distance, defined as the number of positions at which the corresponding amino acids of two sequences are different, of any two consensus protein sequences of avian, human, 2009 pandemic H1N1, and swine viruses was calculated. The distance information in **Table 2** provided insight into the sequence similarity between the proteins of 2009 pandemic H1N1 and those of other virus groups. In particular, the distances between 2009 pandemic H1N1 and avian, human, and swine viruses reflected the origin of 2009 pandemic H1N1 [2].

3.2. Novel Host Sites in the Proteins of 2009 Pandemic H1N1

Our analysis discovered a catalogue of novel host markers in the proteins of 2009 pandemic H1N1 that included 18 markers in HA, 15 in NA, 9 in PB2, 11 in PB1, 13 in PA, 10 in NS1, 1 in NS2, 11 in NP, 3 in M1, and 1 in M2. In the following sections, each of the ten proteins of 2009 pandemic H1N1 was compared to that of avian, human, and swine viruses. Random Forests were employed to identify the top important positions in the proteins of influenza that could separate 2009 pandemic H1N1 from avian, human, and swine viruses, and the top positions that could discriminate 2009 pandemic H1N1 and 2009 seasonal H1N1.

The novel host markers in 2009 pandemic H1N1 were uncovered with the procedure outlined in Section 2.3. Some of the markers that could classify 2009 pandemic H1N1 and 2009 seasonal H1N1 were also part of the previously discovered genomic signature that separate avian or swine from human viruses. Because the sequences of 2009 seasonal H1N1 were a subset of those of human viruses, there were common important sites in each protein between the sites in 2009 pandemic versus 2009 seasonal and the sites in 2009 pandemic versus human viruses.

To render a complete picture of host shift markers of different types, the novel sites in each of the ten proteins of 2009 pandemic H1N1 were exhibited along with the avian-human and swine-human sites in a single table. The conservation of residues comprising these sites in each protein as represented by their frequency at these positions was also displayed in the table. The top important sites in each protein for differentiating 2009 pandemic H1N1 from avian, human, and swine viruses were displayed in a single figure, which were used in the procedure to find the novel sites.

Due to high genetic variation of the HA and NA proteins, only the HA protein sequences of H1 subtype and the NA protein sequences of N1 subtype of avian, human, and swine viruses were used to compare those of 2009 pandemic H1N1 in the current analysis. Therefore, the novel markers in HA and NA of 2009 pandemic H1N1 found in this study were subtype-specific. Because all the PB1-F2 proteins of 2009 pandemic H1N1 were truncated and nonfunctional, they were excluded in this study.

3.2.1. HA Protein

As the primary target of host immune responses, the surface protein HA is under high selection pressure, as evidenced by the large number of amino acid substitutions in this protein. There was a clear distinction of amino acids at position 127, where the human HA had a

Table 3. This table contains the consensus amino acids and their frequency at positions in HA that have high importance in separating 2009 pandemic H1N1 from human H1 viruses. The single letter 'a' (for avian) or 's' (for swine) in parenthesis after a position number indicates whether the same position is also important for separating 2009 pandemic H1N1 from avian or swine viruses or both. The novel host sites in this protein are the positions without an 'a' or a 's' or both.

Position	71	84	120(a)	127(a,s)	128(s)	129(s)	130(a,s)	142	168	216	239	250
Avian	L(93.3%)	N(99.2%)	A(96.7%)	E(97.5%)	T(100%)	T(93.3%)	K(94.2%)	S(92.5%)	N(99.2%)	A(94.2%)	T(97.5%)	A(94.2%)
Human	I(92.5%)	N(98.4%)	E(95.0%)	-(100%)	T(95.8%)	V(95.0%)	T(97.8%)	S(88.4%)	N(98.4%)	K(96.1%)	T(99.2%)	A(100%)
2009 H1N1	S(100%)	S(100%)	T(100%)	D(99.7%)	S(97.0%)	N(100%)	K(100%)	K(100%)	D(100%)	I(99.9%)	K(100%)	V(99.9%)
Swine	F(58.6%)	N(91.8%)	A(44.9%)	E(57.3%)	T(81.0%)	N(64.12%)	R(62.0%)	N(66.2%)	N(85.5%)	A(46.7%)	T(80.0%)	V(62.5%)
Position	257	258	260	261	298	302	314	365	374	493	527	
Avian	L(85.8%)	N(95.0%)	G(94.2%)	S(98.3%)	I(93.3%)	E(98.33%)	M(99.2%)	Q(94.2%)	G(100%)	S(96.7%)	L(99.2%)	
Human	L(96.7%)	S(96.7%)	G(98.0%)	F(97.3%)	V(99.5%)	E(100%)	M(99.5%)	Q(98.9%)	G(99.5%)	S(99.2%)	L(99.1%)	
2009 H1N1	M(100%)	E(99.9%)	N(99.0%)	A(99.6%)	I(100%)	K(100%)	L(99.9%)	L(100%)	E(99.4%)	A(100%)	V(99.9%)	
Swine	M(54.1%)	N(41.7%)	G(78.1%)	S(72.0%)	V(73.1%)	E(94.2%)	M(95.3%)	Q(61.7%)	G(90.8%)	S(67.0%)	L(78.1%)	

Table 4. This table contains the consensus amino acids and their frequency at positions in NA that have high importance in separating 2009 pandemic H1N1 from human N1 viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	84	126	149	163(s)	166	189	257	269	285(s)	321
Avian	T(76.4%)	H(98.9%)	V(95.2%)	V(94.1%)	A(99.6%)	S(93.5%)	K(96.8%)	L(99.7%)	A(97.5%)	V(94.6%)
Human	T(92.8%)	H(100%)	V(95.6%)	L(83.8%)	A(99.8%)	G(86.4%)	K(99.7%)	L(100%)	T(85.5%)	V(99.9%)
2009 H1N1	K(100%)	P(100%)	I(100%)	I(100%)	V(99.7%)	N(99.8%)	R(99.8%)	M(99.7%)	S(100%)	I(100%)
Swine	I(55.8%)	H(89.2%)	V(73.0%)	I(84.5%)	A(64.0%)	G(58.3%)	K(90.3%)	L(90.7%)	T(55.04%)	V(71.6%)
Position	331	365(a,s)	369(a,s)	385	389	395	397	398	436	
Avian	G(99.7%)	T(91.4%)	S(99.4%)	S(88.3%)	V(90.4%)	A(99.2%)	T(99.2%)	D(99.5%)	T(99.6%)	
Human	G(98.8%)	N(84.7%)	K(84.5%)	S(99.7%)	V(94.5%)	A(99.7%)	T(99.7%)	D(99.7%)	T(99.5%)	
2009 H1N1	K(100%)	I(99.8%)	N(100%)	N(100%)	I(100%)	G(100%)	N(99.7%)	E(100%)	-(100%)	
Swine	G(68.7%)	I(63.0%)	S(82.0%)	S(69.4%)	V(38.1%)	A(63.3%)	T(87.8%)	D(92.1%)	T(99.6%)	

deletion whereas the other three virus groups had not (**Table 3**). However, as will be seen in the NA protein section below (**Table 4**), the NA protein of 2009 pandemic H1N1 had a deletion at position 436 though the other three virus groups had not. The positions in Table 3 including 71, 84, 130, 257, 258, and 314 had significant effects on the receptor binding specificity of HA of 2009 pandemic H1N1[13]. HA has two functional domains HA1 (residues 1-327) and HA2 (residues 328-549). Evidently, most of the sites in **Table 3** were in HA1, illustrating a much higher selection pressure of HA1 relative to HA2. The HA active site located in a cleft is composed of the residues 91, 150, 152, 180, 187, 191, and 192. The active site cleft of HA is formed by its right edge (131_GVTAA) and left edge (221_RGQAGR) [55]. Four sites 127, 128, 129, and 130 in **Table 3** were near the right edge of the active site (**Table 3**).

3.2.2. NA Protein

In addition to the surface protein HA, the influenza A virus also has NA as another surface protein, and the balanced interplay between them is essential for the life cycle of this virus. Because of its critical role in viral replication and its highly conserved active sites, NA is the main target for drug design against influenza virus. The NA Inhibitors oseltamivir and zanamivir were the only drugs available to treat the infections caused by 2009 pandemic H1N1, because the novel virus had an adamantine-resistant mutation S31N in its M2 protein [38]. As a result, the surveillance of any potential drug-resistant mutations in the NA protein of 2009 pandemic H1N1 received high priority. The mutation H275Y (N1 numbering, H274Y in N2 numbering) in NA is well known for its resistance to NA Inhibitors. There were 123 Ys and 2 Hs in 125 NA sequences of 2009 seasonal H1N1 and 12 Ys and 631 Hs in 643 NA sequences of 2009 pandemic H1N1 used in the current study. Both NAs in 2009 pandemic and 2009 seasonal H1N1 did not have the novel NA mutation Q136K [41] that confers zanamivir resistance.

NA is also is constantly evolving under host immune pressure, and the mutations in Table 4 illustrated its genetic variation. As mentioned in the HA protein section above, the NA of 2009 pandemic H1N1 had a deletion at

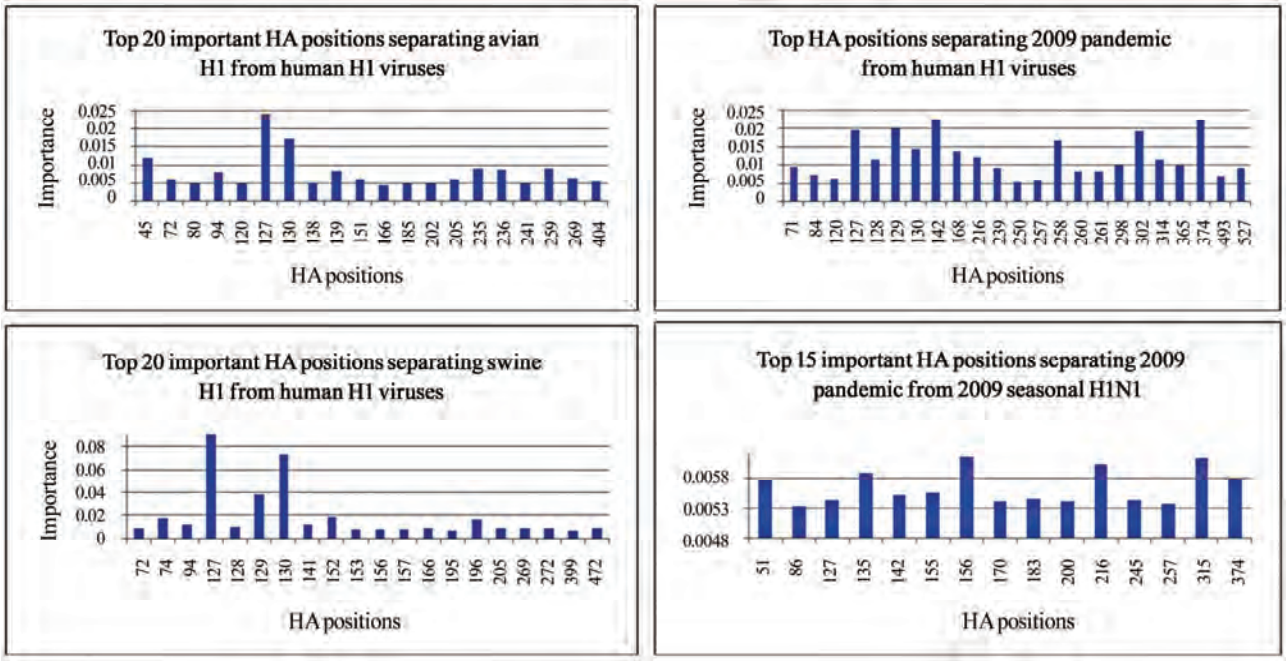


Figure 1. Top important HA positions in distinguishing avian H1, human H1, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine H1 viruses.

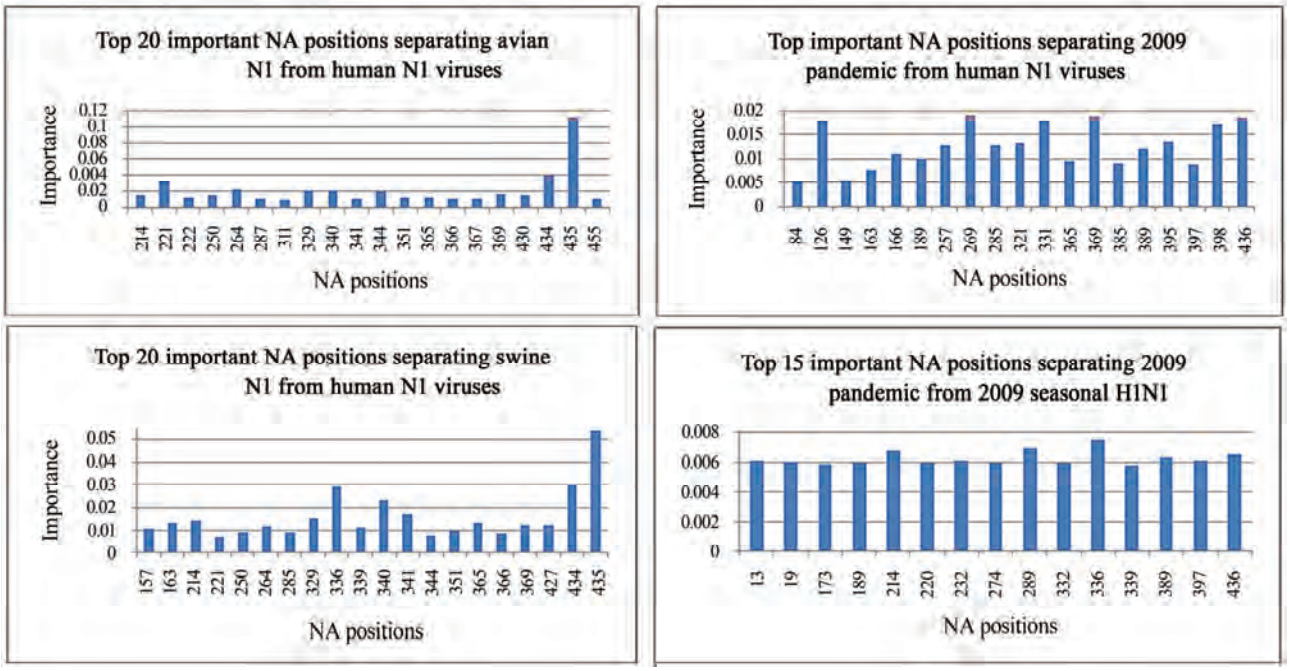


Figure 2. Top important NA positions in distinguishing avian N1, human N1, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine N1 viruses.

position 436 while the other three virus groups had not. However, the HA of human virus had a deletion at position 127 but the other three virus groups had not.

The NA active site is a shallow pocket constructed from conserved residues, some of which contact the substrate directly and participate in catalysis, while others provide a structural framework [56]. According to the numbering in [57], these residues of N1 are 118, 119,

151, 152, 156, 179, 180, 223, 225, 228, 247, 277, 278, 293, 295, 368, and 402. The antigenic sites of N1 are residues 83-143, 156-190, 252-303, 330, 332, 340-345, 368, 370,387-395, 431-435, 448-468. Novel host sites 84, 126, 166, 189, 257, 269, 389, and 395 were at the antigenic sites of N1 (**Table 4**).

Because of a common deletion in the stalk region of the NA proteins of avian viruses, only the residues after position 82 were included in the Random Forest analysis on avian, human, 2009 pandemic H1N1, and swine viruses. However, the whole NA sequences were used in the analysis of 2009 pandemic and 2009 seasonal H1N1.

3.2.3. M1 Protein

M1 protein forms a shell inside the viral envelope to offer strength and rigidity to the viral structure. M1 interacts with HA, NA, M2, and lipid membranes during budding of new virions from the cell surface, and functions in the formation of vRNP complexe and the dissociation of vRNP from the nuclear matrix, and in assembly by recruiting the viral components to the site of assembly. The dissociation of M1 from vRNP is triggered by transport of hydrogen ions across the viral membrane by M2, an early step preceding entry of vRNPs into the cytoplasm of the host cells. M1 also binds to NS2 to facilitate nuclear export of the vRNP [37]. There was a mutation R101K in the M1 protein of 2009 pandemic H1N1 (**Table 5**). It would be of interest to exam the impact of this mutation on viral replication. The basic amino acids 101RKLKR105 of M1 were involved in vRNP binding and nuclear localization. In [58], the functions of 101RKLKR105 were studied by introducing mutations into the M gene of influenza virus A/WSN/33. Individual substitution, R101S or R105S, had a minimal effect on viral replication, but the double mutation R101S-R105S reduced viral replication at a restrictive temperature.

The M1 is a highly conserved protein. Therefore, the changes of M1 may reflect host-specific adaptation. Positions 115, 121, and 137 were identified as avian-human host shift markers in [43]. Our investigation indicated position 218 was as important as position 121. Position 137 was a swine-human marker in [40], but our study also revealed positions 115 and 218 were as important as positions 137 as swine-human markers (**Figure 3**). The novel site 30 was in the membrane binding domain [43], and sites 207 and 209 were in the C-terminal part of M1 (residues 165-252) that binds to vRNP [59].

3.2.4. M2 Protein

This 97 amino acid-long integral membrane protein has three domains, one N-terminal extracellular domain (24 residues) recognized by host immune system, one 19-residue transmembrane domain responsible for ion channel activity, and one 54-residue cytoplasmic tail interacting with M1 and required for genome packing and formation of virus particles [37]. Two M2 inhibitors (adamantine and rimantadine) affect two steps in the replication cycle, viral uncoating and viral maturation. There are five known adamantine-resistanant mutations in M2 (L26F, V27A, A30V, A30T, S31N, and G34E). The 2009 pandemic H1N1 virus contains a mutation S31N. They also contain a mutation L43T in M2 (**Table 6** and **Figure 4**), which is not present in seasonal, triple-reassortant swine or H5N1 influenza viruses [15]. The replacement of the non-polar residue L43 by the polar residue T43 in M2 may influence a nearby functional residue W44, the channel lock and the binding site of rimantadine [60]. Positions 11, 14, 20, 28, 54, 55, 57, 78, and 86 were avian-human host shift sites found in [43]. However, the positions 18, 50, 86, and 93 were as important as these sites in our examination. Positions 57, 86, and 93 were swine-human shift markers in [40], but our analysis also included positions 28, 54, 77, 78, 79, and 89 as swine-human markers with high importance (**Figure 4**). The only novel site in this protein was 13 which was in the extracellular domain (**Table 6**).

3.2.5. NP Protein

The NP protein of the influenza virus binds the RNA genome and functions as an adaptor between the virus and the host cell. The interaction of the NP protein with the viral polymerase is required for viral RNA replication, but not for the synthesis of viral messenger RNAs(transcription). Previous experiments implicated

Table 5. This table contains the consensus amino acids and their frequency at positions in M1 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	15(a)	30	101(a,s)	115(a,s)	116(s)	121(a,s)	137(a,s)	142(a,s)	166(a,s)	207	209	214(s)	218(a,s)
Avian	I(52.4%)	D(99.9%)	R(52.7%)	V(99.5%)	A(97.5%)	T(96.2%)	T(99.4%)	V(91.6%)	V(53.7%)	S(68.4%)	A(98.9%)	Q(99.2%)	T(99.8%)
Human	V(67.6%)	D(99.7%)	R(93.1%)	I(92.3%)	A(98.7%)	A(92.9%)	A(93.1%)	V(68.2%)	V(93.2%)	S(92.7%)	A(99.9%)	Q(99.8%)	A(84.5%)
2009 H1N1	I(100%)	S (99.8%)	K(100%)	V(99.4%)	S(100%)	T(100%)	T(100%)	A(99.6%)	A(100%)	N(99.8%)	T(100%)	H(100%)	T(100%)
Swine	V(65.3%)	D(77.3%)	R(64.7%)	V(90.9%)	A(68.4%)	A(59.7%)	T(93.2%)	V(77.3%)	V(66.1%)	S(91.9%)	A(78.5%)	Q(67.1%)	T(92.4%)

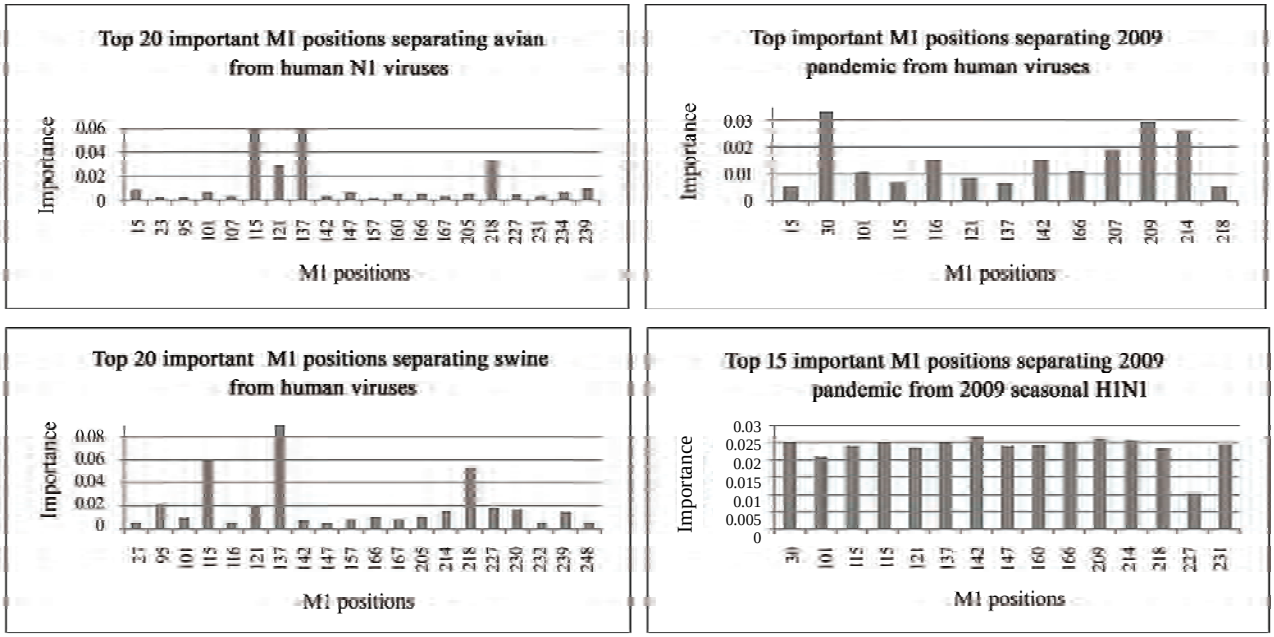


Figure 3. Top important M1 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

Table 6. This table contains the consensus amino acids and their frequency at positions in M2 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	11(a,s)	13	16(a,s)	20(a,s)	28(a,s)	31(s)	43(a,s)	77(s)	78(a,s)	86(a,s)
Avian	T(92.0%)	N(91.70%)	E(92.4%)	S(96.2%)	I(54.5%)	S(88.6%)	L(97.0%)	R(98.1%)	Q(99.4%)	V(99.6%)
Human	I(91.6%)	N(99.29%)	G(91.2%)	N(92.2%)	V(96.8%)	S(67.6%)	L(67.7%)	R(99.6%)	K(60.9%)	A(91.7%)
2009 H1N1	T(100%)	S(99.79%)	E(100%)	S(100%)	I(100%)	N(100%)	T(100%)	Q(100%)	Q(100%)	V(100%)
Swine	T(52.7%)	N(72.66%)	E(57.9%)	N(52.7%)	I(38.9%)	S(60.8%)	L(93.8%)	R(63.5%)	Q(95.6%)	V(94.6%)

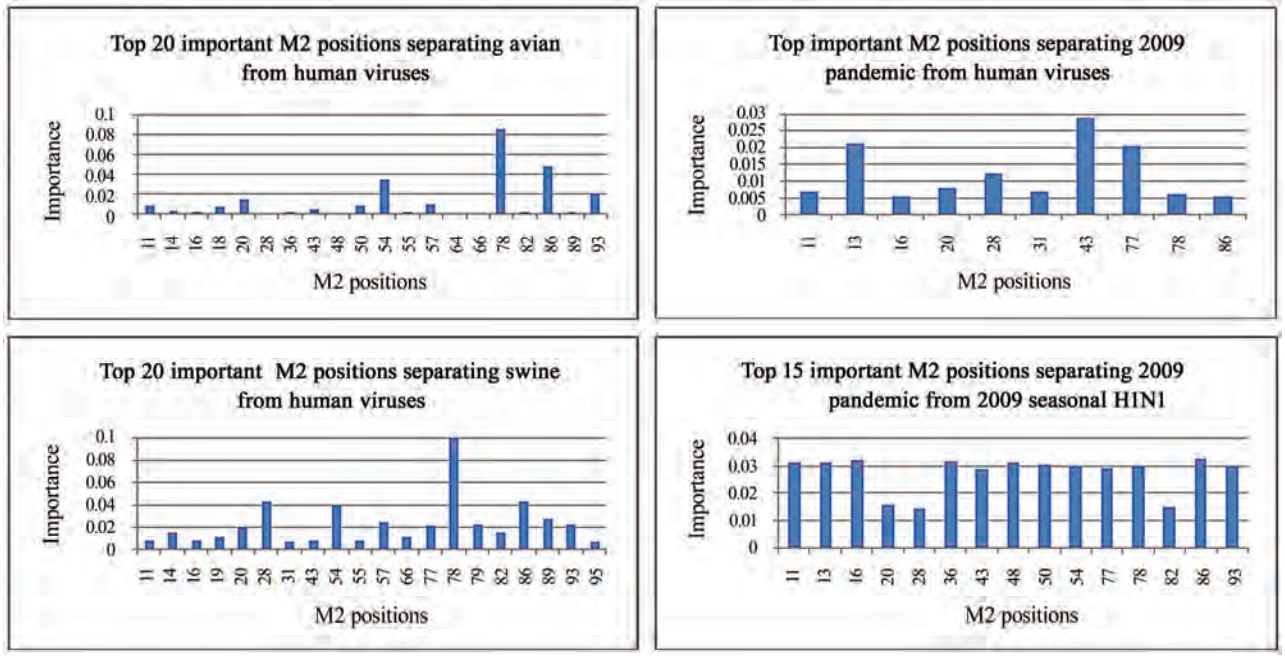


Figure 4. Top important M2 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

three NP regions (residues 1-160, 256-340 and 340-498) in binding to PB1 and PB2 [61]. Novel sites 21, 53, 119, 316, 353, 371, 377, 433, 444, and 498 were scattered in these three regions (**Table 7**). One region, residues 360-374, in NP of 2009 pandemic H1N1 was deemed extremely important for host range restriction, and is a common feature of pandemic viruses [25]. Two positions 371 and 373 in Table 7 were in this region. Residue 100 was involved in the NP-PB2 interaction [62], and ranked second in separating 2009 pandemic H1N1 from avian viruses. The consensus amino acids at 100 of PB2 proteins of avian, human, 2009 pandemic H1N1, and swine viruses were R, V, I and V respectively. The mutation V100I might contribute to the increased transmissibility or infection of 2009 pandemic H1N1 [41].

Positions 16, 33, 61, 100, 136, 214, 283, 305, 313, 357, 375, and 423 were avian-human host shift markers in [43]. Furthermore, we found positions 31, 217, 373, and 455 significant for discriminating avian and human viruses (**Figure 5**).

3.2.6. NS1 Protein

All of the proteins in influenza virus are structural except for NS1 and PB1-F2. This protein is designated as non-structural because it is synthesized in infected cells, but is not incorporated into virions. NS1 is a multifunctional protein involved in both protein-protein and protein-RNA interactions. Its N-terminal region has an RNA-binding domain (residues 1-73) and its C-terminal region (residues 74-237) contains the effector domain that inhibits the maturation and exportation of the host cellular antiviral mRNAs [63].

Because of a truncation in the NS1 proteins of 2009 pandemic H1N1, only the first 219 residues of the NS1 proteins were included in our analysis. Positions 22, 60, 81, 84, 215, and 227 were avian-human host shift sites in [43], whereas our Random Forests analysis implied positions 79, 81, 114, 171, and 215 were as significant as

Table 7. This table contains the consensus amino acids and their frequency at positions in NP that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	21	31(a,s)	53	119	189(s)	190	217(a)	289(s)	313(a,s)	316
Avian	N(99.20%)	R(99.8%)	E(99.90%)	I(97.65%)	M(99.1%)	V(98.61%)	I(94.8%)	Y(99.2%)	F(99.0%)	I(99.58%)
Human	N(97.5%)	K(65.4%)	E(100%)	I(97.46%)	M(97.5%)	V(97.20%)	S(48.1%)	Y(97.5%)	Y(78.1%)	I(99.75%)
2009 H1N1	D(100%)	R(100%)	D(100%)	V(100%)	I(99.75%)	A(100%)	V(98.2%)	H(99.7%)	V(100%)	M(100%)
Swine	D(61.19%)	R(82.6%)	E(98.10%)	V(57.14%)	I(60.00%)	A(56.42%)	I(76.90%)	H(64.29%)	F(86.66%)	I(95.24%)
Position	350(s)	353	371	373(a)	377	430(s)	433	444	456(s)	498
Avian	T(94.39%)	V(90.30%)	M(94.77%)	T(69.18%)	S(69.67%)	T(94.8%)	T(95.36%)	I(99.00%)	V(98.4%)	N(96.09%)
Human	T(97.20%)	S(52.42%)	M(91.35%)	A(34.35%)	S(80.66%)	T(83.72%)	T(88.04%)	I(98.22%)	V(82.95%)	N(96.18%)
2009 H1N1	K(100%)	I(99.75%)	V(100%)	T(76.40%)	N(100%)	S(100%)	N(100%)	V(100%)	L(100%)	S(99.24%)
Swine	K(64.52%)	V(53.81%)	V(59.29%)	A(57.38%)	S(51.90%)	S(38.57%)	N(60.95%)	I(64.76%)	L(61.19%)	N(69.29%)

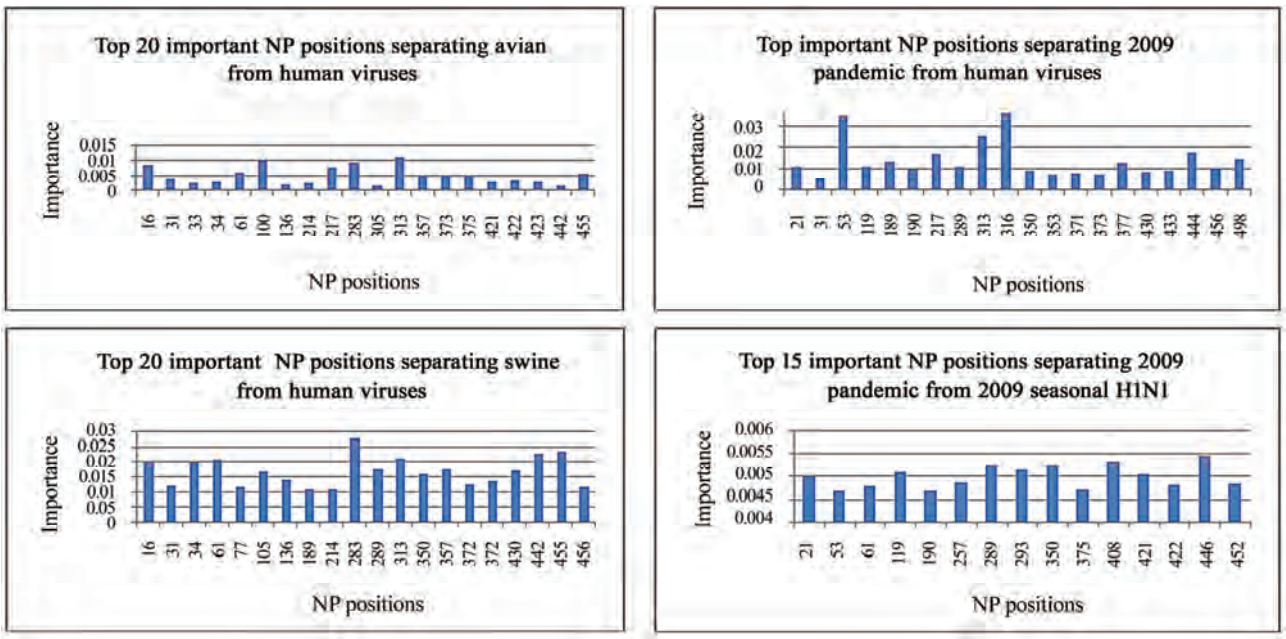


Figure 5. Top important NP positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

Table 8. This table contains the consensus amino acids and their frequency at positions in NS1 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	6	25(s)	59	67(s)	74	76	91	111	112(a)
Avian	V(77.85%)	Q(79.35%)	R(73.41%)	R(78.85%)	D(97.87%)	A(77.79%)	T(97.33%)	V(79.76%)	A(59.11%)
Human	V(95.48%)	Q(95.09%)	H(45.19%)	R(54.62%)	D(97.64%)	A(97.84%)	T(96.66%)	V(96.07%)	E(55.00%)
2009 H1N1	M(99.73%)	N(99.73%)	L(100%)	W(100%)	S(99.73%)	T(100%)	S(99.73%)	I(100%)	I(99.45%)
Swine	V(95.26%)	N(58.70%)	L(59.49%)	W(59.68%)	D(37.49%)	T(58.70%)	A(60.28%)	V(50.59%)	A(36.96%)
Position	119	129(s)	171(a,s)	198	205	206(s)	207	213(s)	217(s)
Avian	M(99.39%)	I(75.55%)	D(48.87%)	L(52.52%)	S(69.50%)	S(67.66%)	D(59.20%)	P(92.56%)	K(68.00%)
Human	M(85.85%)	M(51.28%)	I(55.20%)	L(84.28%)	S(92.93%)	S(91.55%)	N(78.00%)	P(97.05%)	K(69.16%)
2009 H1N1	L(100%)	V(100%)	Y(100%)	I(100%)	N(100%)	C(100%)	D(100%)	S(100%)	E(100%)
Swine	M(90.12%)	I(64.23%)	D(59.29%)	L(97.04%)	S(64.62%)	R(56.13%)	N(92.09%)	P(51.38%)	E(57.51%)

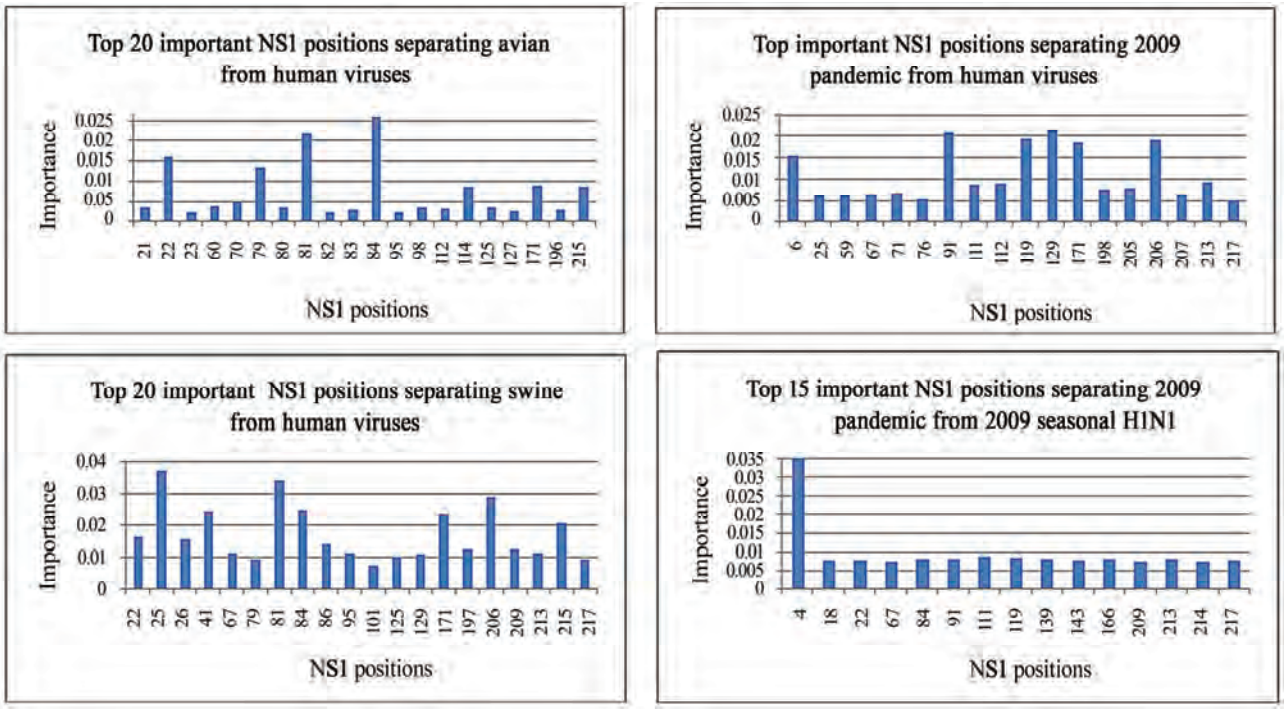


Figure 6. Top important NS1 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

these sites (**Figure 6**). There were two novel sites 6 and 59 in the RNA-binding domain and the other novel sites in the effector domain (**Table 8**).

3.2.7. NS2 Protein

Influenza virus replicates its RNA genome in the nucleus of infected cells. The NS2 protein mediates the nuclear export of virion RNAs, with help from M1 and NP. A recent report indicated that it also has a role in the regulation of viral transcription and replication [64]. NS2 contains a highly conserved nuclear export signal motif in its amino-terminal region (residues 12-21) [65], and site 14 in **Table 9** was in this region.

Positions 60, 70, and 107 were avian-human host shift markers in [43]. We found position 14 important as a host marker as well. Position 107 was a swine-human host switch marker in [40], but our analysis also pointed to positions 14, 32, 49, and 57 as such sites of high significance (**Figure 7**). The NS2 protein of 2009 pandemic H1N1 contained so many important avian-human or swine-human sites, resulting in only one site as a novel site (**Table 9**).

3.2.8. PA Protein

Compared to the well-defined functions of PB1 and PB1, PA is involved in a diverse range of functions of the polymerase complex, including protein stability, endonuclease activity, and cap binding and promoter binding

[66]. Positions 28, 55, 57, 65, 66, 100, 225, 268, 321, 337, 356, 382, 400, 404, 409, 421, and 552 were avian-human host shift markers in [43]. Additionally, we found positions 241 and 383 equally important as these positions as avian-human markers. Positions 268 and 552 were swine-human markers uncovered in [40]. Our analysis suggested the positions 28, 225, 337, and 400 were equally crucial as these two sites as swine-human markers (**Figure 8**).

The N-terminal domain of PA (residues 1-256) harbors several functional domains, including an endonuclease active site with a putative active site motif, two putative nuclear transport motifs (residues 124-139 (NLS1) and residues 186-247 (NLS2)), and a proteolytic domain that can induce generalized proteolysis of both viral and host proteins. The C-terminal domain of PA (residues 257-716) binds to PB1 for complex formation and nuclear transport [66]. There were three novel sites 186, 204, and 213 within the second putative nuclear localization signals (NLS2), and one novel site 626 within the PB1 binding domain (**Table 10**).

3.2.9. PB1 Protein

The influenza virus polymerase is responsible for replication and transcription of the eight gene segments of the viral RNA genome in the infected host cell. PB1 can interact with PB2, PA, and NP and binds to viral promoter, and is accountable for viral RNA elongation and cap RNA cleavage activities [66,67].

Position 336 was the only avian-human host shift markers in [43]. We found positions 212, 327, 361, 375, 384, 401, 473, and 584 equally significant as position 336 (**Figure 9**). There were one novel site 12 within the PB1-PA binding domain (residues 1-25) and two novel sites 618 and 728 in the PB1-PB2 binding domain (residues 600-757) (**Table 11**) [68].

PB1-PA binding domain (residues 1-25) and two novel

Table 9. This table contains the consensus amino acids and their frequency at positions in NS2 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	6	14(a,s)	32(s)	34(s)	40(a,s)	48(a)	57(a,s)
Avian	V(79.1%)	M(56.2%)	I(99.0%)	Q(95.6%)	L(71.54%)	A(73.2%)	S(97.6%)
Human	V(98.2%)	L(55.6%)	I(99.0%)	Q(98.7%)	L(61.9%)	A(96.6%)	S(60.3%)
2009 H1N1	M(99.7%)	M(100%)	V(99.7%)	R(100%)	I(100%)	T(100%)	Y(100%)
Swine	V(95.2%)	M(83.2%)	V(68.1%)	Q(53.3%)	I(67.8%)	A(70.1%)	Y(65.8%)
Position	60(a,s)	63(a,s)	83(a)	89(a,s)	107(a,s)	115(a)	
Avian	S(55.7%)	G(75.8%)	V(71.7%)	I(69.8%)	L(99.9%)	T(84.4%)	
Human	N(68.7%)	G(96.3%)	V(98.2%)	T(63.2%)	F(74.9%)	T(89.8%)	
2009 H1N1	S(100%)	E(95.8%)	M(99.7%)	A(97.8%)	L(100%)	A(99.4%)	
Swine	N(61.0%)	E(62.1%)	V(97.4%)	M(32.5%)	L(90.9%)	T(98.3%)	

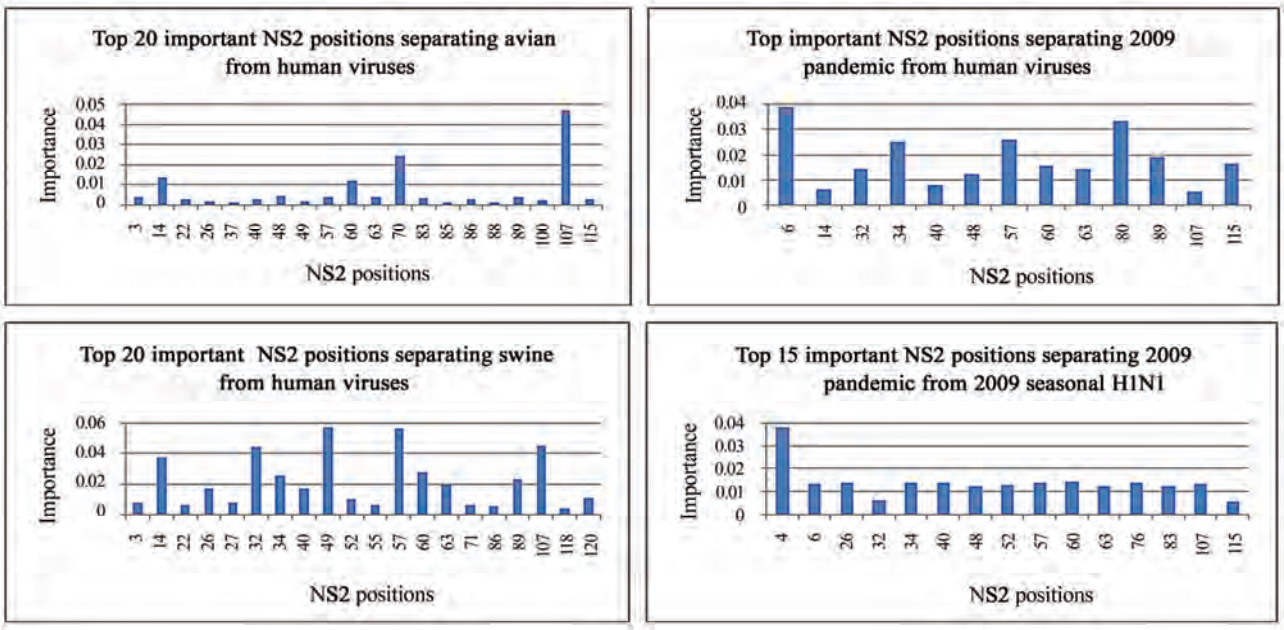


Figure 7. Top important NS2 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

Table 10. This table contains the consensus amino acids and their frequency at positions in PA that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	28(a,s)	55(a)	85	100(a,s)	186	204	213	256	262	275	277
Avian	P(99.5%)	D(98.9%)	T(96.5%)	V(94.5%)	G(98.0%)	R(80.6%)	R(97.8%)	R(98.6%)	K(96.9%)	P(97.5%)	S(97.1%)
Human	L(70.3%)	N(71.3%)	T(91.0%)	A(70.6%)	G(100%)	R(53.0%)	R(97.5%)	R(54.5%)	K(95.0%)	P(98.9%)	S(35.8%)
2009 H1N1	P(100%)	D(100%)	I(100%)	V(100%)	S(100%)	K(100%)	K(100%)	K(100%)	R(100%)	L(98.6%)	H(100%)
Swine	P(79.9%)	D(53.4%)	T(60.9%)	V(86.6%)	G(96.8%)	R(90.7%)	R(92.1%)	R(63.8%)	K(77.8%)	P(91.5%)	S(53.4%)
Position	336	337(a,s)	356(a)	362	388	400(a,s)	404(a,s)	407	552(a,s)	626	
Avian	L(99.5%)	A(88.7%)	K(98.9%)	K(99.5%)	S(80.1%)	S(40.9%)	A(93.2%)	I(95.9%)	T(99.7%)	K(83.7%)	
Human	L(97.1%)	S(35.8%)	R(69.9%)	K(98.6%)	S(84.2%)	L(79.2%)	S(72.0%)	I(98.6%)	S(71.7%)	K(98.2%)	
2009 H1N1	M(100%)	A(99.0%)	R(99.7%)	R(100%)	G(99.0%)	P(100%)	A(100%)	V(99.3%)	T(99.7%)	R(99.3%)	
Swine	L(95.6%)	A(86.9%)	K(53.9%)	K(71.1%)	S(52.8%)	P(30.6%)	A(86.9%)	I(71.1%)	T(91.5%)	K(96.2%)	

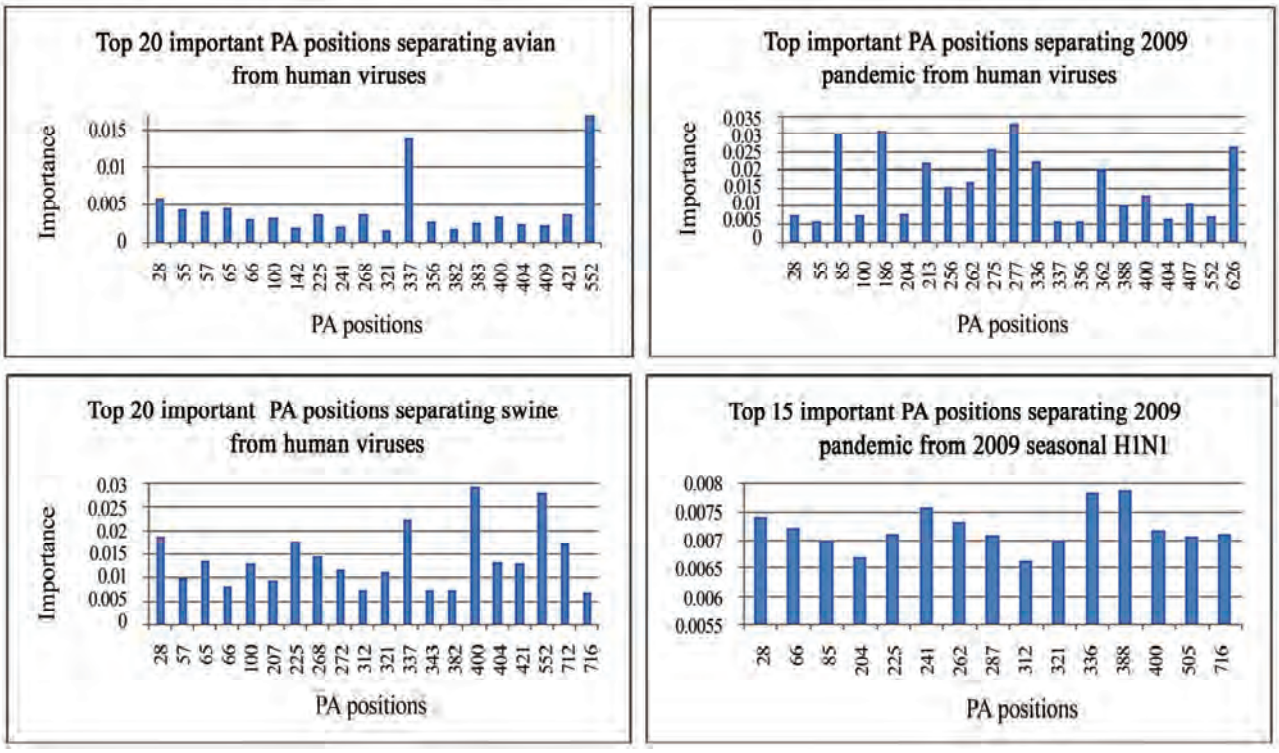


Figure 8. Top important PA positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

sites 618 and 728 in the PB1-PB2 binding domain (residues 600-757) (**Table 11**) [68].

3.2.10. PB2 Protein

PB2 interacts with PB1 and NP, but not PA. Its primary function is binding to cap structures on host cell pre-mRNAs before they are cleaved to provide primers for viral mRNA synthesis [66]. Positions 9, 44, 64, 81, 105, 199, 271, 292, 368, 475, 567, 588, 613, 627, 661, 674, and 702 were avian-human host shift markers in [43]. Positions 108, 197, and 684 were as significant as these sites in our finding. Position 44 was a swine- human marker in [40], but our analysis implied positions 64, 65, 81, 105, 199, 292, 567, 627, 649, 661, and 674 were equally important as position 44 (**Figure 10**). Position 702, an avian-human marker selected in [40,43,69], ranked 21th in our Random Forests analysis, and therefore it was not included in our plot in **Figure 10**. In addition to the SR polymorphism, S590 and R591, found in [23], novel sites in PB2 discovered here provided additional polymorphism that might convey enhanced polymerase activity in human cells. Position 627 in PB2 was considered critical for host shifts in our analysis, a well-known host marker discussed in [21,22], and was located in the PB2-PB1 and PB2-NP binding domains [43].

The PB2-NP binding domain contains residues 1-269 and 580-683, and the PB2-PB1 binding domain contains residues 51-259 and 580-759. There were novel sites 54, 590, 645, and 667 in the PB2-PB1 and PB2-NP binding

Table 11. This table contains the consensus amino acids and their frequency at positions in PB1 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	12	175	179	216	298	327(a,s)	339(s)	361(a,s)	364	386(s)
Avian	V(99.0%)	D(95.9%)	M(95.9%)	S(95.7%)	L(98.6%)	R(98.7%)	I(98.7%)	S(99.0%)	L(99.5%)	R(57.0%)
Human	V(99.7%)	D(97.2%)	M(69.6%)	S(65.7%)	L(79.2%)	K(53.6%)	I(96.9%)	S(61.9%)	L(99.3%)	R(65.4%)
2009 H1N1	I(99.7%)	N(100%)	I(100%)	G(100%)	I(100%)	R(100%)	M(100%)	R(100%)	I(100%)	K(100%)
Swine	V(95.1%)	D(96.7%)	M(66.0%)	S(60.1%)	L(96.7%)	R(89.7%)	I(44.6%)	N(32.3%)	L(96.5%)	R(95.4%)
Position	435	486	517(s)	584(a,s)	587	618	638(s)	728	741(a,s)	
Avian	T(99.0%)	R(98.6%)	I(99.1%)	R(97.1%)	A(98.4%)	E(97.7%)	E(98.8%)	I(99.1%)	A(96.2%)	
Human	T(99.3%)	R(64.7%)	I(81.3%)	R(63.7%)	A(98.6%)	E(99.7%)	E(98.6%)	I(100%)	A(59.2%)	
2009 H1N1	I(99.4%)	K(100%)	V(100%)	Q(100%)	V(97.4%)	D(100%)	D(100%)	V(100%)	S(100%)	
Swine	T(66.8%)	R(64.1%)	I(75.5%)	R(40.5%)	A(86.7%)	E(61.4%)	E(69.0%)	I(98.4%)	A(59.0%)	

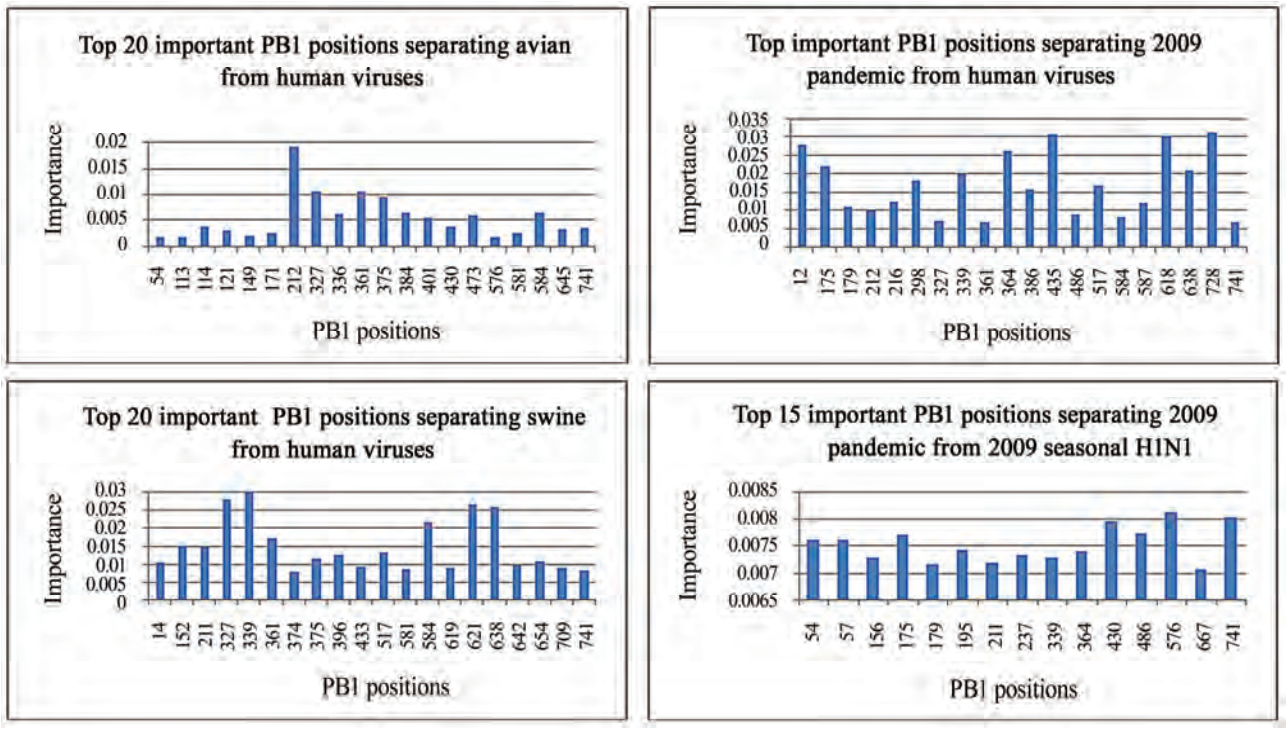


Figure 9. Top important PB1 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

domains and sites 147 and 225 in the PB2-NP binding domain [43] (**Table 12**).

4. DISCUSSIONS

Extensive research to date provided highly informative knowledge about the origin and genetic lineages of 2009 pandemic H1N1, but the host markers of this new virus remained elusive. Recent studies indicated that human host adaptation is complex and multigenic, and the well-known host shift markers are lacking in this new virus. The hypothesis in the current study was that these markers of 2009 pandemic H1N1 might exist outside of the space of traditional host switch markers. To test this hypothesis in this study, Random Forests were applied to uncover novel important markers in each of the ten proteins of influenza that could differentiate 2009 pandemic H1N1 from human viruses, but were not present in the previous avian-human or swine-human host switch markers.

Our approach naturally led to a systematic discovery of new host markers like the SQ polymorphism found in [23] that could enrich our current knowledge of 2009 pandemic H1N1 and complement the repertoire of existing host shift signatures. Among others, this study revealed the novel host sites 54, 147, 225, 315, 453, 559, 590, 645, and 667 in PB2 of 2009 pandemic H1N1. They provided ample potential sites to investigate experimentally whether they also compensate the lack of

amino acid lysine at residue 627, as the SR polymorphism. Their prospective broader roles in enhancing this new virus's replication and transmission in humans are worthy of further research. In this regard, the three positions 54, 315, and 559 in PB2 were particularly of interest because they had much higher importance than the two positions 590 and 591 associated with the SR polymorphism.

Four proteins are involved and required in the synthesis of influenza virus RNA, which are PB2 and PA of avian lineage, PB1 of human origin, and NP derived from classical swine viruses in 2009 pandemic H1N1. To gain insight into the adaptive strategies employed by these four proteins of different origins to evade restriction in human cells will be a challenge. The novel sites identified in this study provided a starting point for future integrative examination of the interactions of these proteins.

It was expected that 2009 pandemic H1N1 would co-circulate with seasonal H1N1 for some time. Our catalogue of amino acid markers that could effectively separate 2009 pandemic H1N1 from 2009 seasonal H1N1 presented a valuable view of these two viruses that share similar clinical courses but are unique genetically.

Table 12. This table contains the consensus amino acids and their frequency at positions in PB2 that have high importance in separating 2009 pandemic H1N1 from human viruses. The novel host sites in this protein are the positions without an 'a' (for avian) or a 's' (for swine) or both.

Position	9(a)	54	64(a,s)	65(s)	81(a,s)	105(a,s)	147	184(s)	199(a,s)
Avian	D(97.5%)	K(99.7%)	M(74.9%)	E(97.8%)	T(97.3%)	T(90.9%)	I(82.1%)	T(96.3%)	A(99.2%)
Human	N(71.0%)	K(100%)	T(68.0%)	E(98.5%)	M(52.0%)	V(52.4%)	I(87.7%)	T(99.6%)	S(72.9%)
2009 H1N1	D(100%)	R(100%)	M(100%)	D(99.7%)	T(100%)	T(100%)	T(100%)	A(99.0%)	A(100%)
Swine	D(63.9%)	K(98.5%)	M(53.2%)	E(69.4%)	T(84.7%)	T(87.8%)	I(68.2%)	T(57.2%)	A(55.0%)
Position	225	292(a,s)	315	340(s)	453	475(a)	559	567(a,s)	588(a,s)
Avian	S(99.4%)	I(88.6%)	M(95.2%)	R(52.2%)	P(94.7%)	L(99.2%)	T(91.0%)	D(98.0%)	A(95.8%)
Human	S(98.9%)	T(73.6%)	M(99.6%)	R(60.2%)	H(52.0%)	M(70.3%)	T(71.7%)	N(70.3%)	I(68.4%)
2009 H1N1	G(100%)	V(99.4%)	I(100%)	K(97.4%)	S(99.7%)	L(100%)	I(100%)	D(100%)	T(98.4%)
Swine	S(72.2%)	I(56.0%)	M(97.2%)	R(59.9%)	P(57.8%)	L(54.1%)	T(70.9%)	D(90.2%)	A(55.7%)
Position	590	591(s)	613(a,s)	627(a,s)	645	661(a,s)	667	674(a,s)	684(a)
Avian	G(87.6%)	Q(97.9%)	V(98.4%)	E(91.7%)	M(99.4%)	A(96.2%)	V(92.4%)	A(97.1%)	A(96.9%)
Human	G(69.9%)	Q(98.9%)	T(64.3%)	K(80.3%)	M(98.9%)	T(78.4%)	I(62.1%)	T(69.1%)	A(51.3%)
2009 H1N1	S(99.7%)	R(100%)	V(100%)	E(100%)	L(100%)	A(100%)	V(100%)	A(99.7%)	S(100%)
Swine	G(71.3%)	Q(67.3%)	V(74.9%)	E(53.8%)	M(74.9%)	A(48.0%)	V(69.4%)	A(86.5%)	A(74.3%)

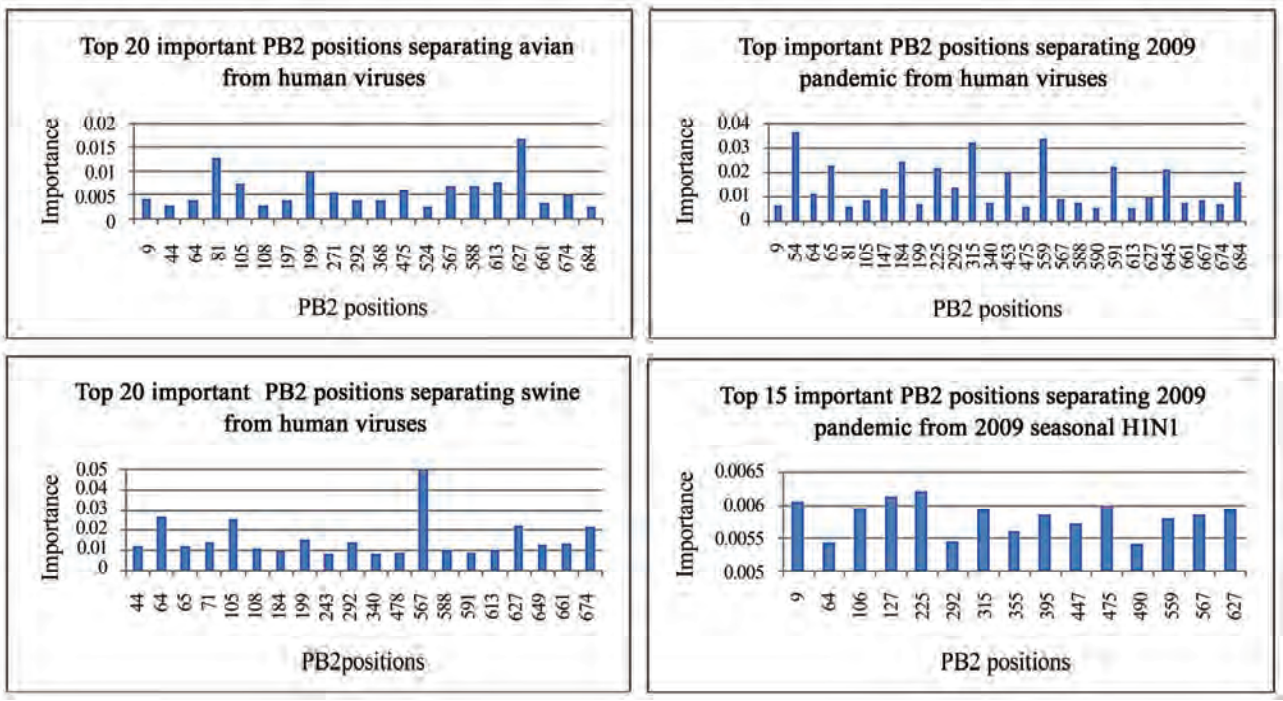


Figure 10. Top important PB2 positions in distinguishing avian, human, 2009 pandemic H1N1, 2009 seasonal H1N1, and swine viruses.

Various computational techniques including entropy [39,40], mutual information [42,43], statistical tests [44], and support vector machines [45] were utilized to discover molecular markers in influenza viruses. To demonstrate the validity of using Random Forests as a feature selection technique in identifying novel host markers in 2009 pandemic H1N1, the top markers found by Random Forests to distinguish the human virus from avian or swine viruses were also included in this report, which contained many known host adoption markers from previous studies. There were fewer novel sites in M1, M2, and NS2 than in the other proteins under this study resulting from many avian-human or swine-human sites among these proteins.

5. CONCLUSIONS

Our findings confirmed that there are novel host sites in the proteins of 2009 pandemic H1N1 that could separate this new virus from human viruses with high confidence. These markers could not be found in the search space of traditional avian-human or swine-human host shift markers, thus offering new potential sites for further experimental verification to elucidate their biological functions.

6. ACKNOWLEDGEMENTS

We thank Houghton College for its financial support.

REFERENCES

[1] Samji, T. (2009) Influenza A: Understanding the viral life cycle. *Yale Journal of Biology Medicine*, **82(4)**, 153-159.

[2] Gavin, J.D., Smith, D.V., Bahl, J., Lycett, S.J., *et al.* (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*, **459**, 1122-1125.

[3] Chang, Y.S., van Hal, S.J., Spencer, P.M, Gosbell, I.B. and Collett, P.W. (2010) Comparison of adult patients hospitalised with pandemic (H1N1) 2009 influenza and seasonal influenza during the PROTECT phase of the pandemic response. *The Medical Journal of Australia*, **192(2)**, 90-93.

[4] Dhiman, N., Mark, J.E., Irish, C., Wright, P., Smith, T.F. and Pritt, B.S. (2010) Mutability in the matrix gene of novel influenza A H1N1 virus detected using a fret probe-based real-time reverse transcriptase PCR assay. *Journal of Clinical Microbiology*, **48(2)**, 677-679.

[5] Zheng, X., Todd, K.M., Yen-Lieberman, B., Kaul, K., Mangold, K. and Shulman, S.T. (2009) Unique finding of a 2009 H1N1 influenza virus—positive clinical sample suggests matrix gene sequence variation. *Journal of Clinical Microbiology*, **48(2)**, 665-666.

[6] Shen, J., Ma, J. and Wang, Q. (2009) Evolutionary trends of A (H1N1) influenza virus hemagglutinin since 1918. *PLoS One*, **4(11)**, e7789.

[7] Soundararajan, V., Tharakaraman, K., Raman, R., Raguram, S., Shriver, Z., Sasisekharan, V. and Sasisekharan, R. (2009) Extrapolating from sequence—the 2009 H1N1 'swine' influenza virus. *Nature Biotechnology*, **27**, 510-513.

[8] Childs, R.A., Palma, A.S., Wharton, S., Matrosovich, T., Liu, Y., Chai, W.G., Campanero-Rhodes, M.A., *et al.* (2009) Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. *Nature Biotechnology*, **27**, 797-799.

[9] Igarashi, M., Ito, K., Yoshida, R., Tomabechi, D., Kida, H. and Takada, A. (2009) Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. *PLoS One*, **5(1)**, e8553.

[10] Cosic, I. (1997) The resonant recognition model of macromolecular bioreactivity, theory and application. Birkhauser Verlag, Berlin.

[11] Veljkovic, V., Niman, H.L., Glisic, S., Veljkovic, N., Perovic, V. and Muller, C.P. (2009) Identification of hemagglutinin structural domain and polymorphisms which may modulate swine H1N1 interactions with human receptor. *BMC Structural Biology*, **9**, 62.

[12] Hu, W. (2010) Identification of highly conserved domains in hemagglutinin associated with the receptor binding specificity of influenza viruses: 2009 H1N1, avian H5N1 and swine. *Journal of Biomedical Science and Engineering*, **3**, 114-123.

[13] Hu, W. (2010) Quantifying the effects of mutations on receptor binding specificity of influenza viruses. *Journal of Biomedical Science and Engineering*, **3**, 227-240.

[14] Janies, D.A., Voronkin, I.O., Studer, J., Hardman, J., Alexandrov, B.B., Treseder, T.W. and Valson, C. (2010) Selection for resistance to oseltamivir in seasonal and pandemic H1N1 influenza and widespread co-circulation of the lineages. *International Journal of Health Geographics*, **9(1)**, 13.

[15] Deyde, V.M., Sheu, T.G., Trujillo, A.A., Okomo-Adhiambo, M., Garten, R., Klimov, A.I. and Gubareva, L.V. (2010) Detection of molecular markers of drug resistance in 2009 pandemic influenza A (H1N1) viruses by pyrosequencing. *Antimicrob Agents Chemother*, **54(3)**, 1102-1110.

[16] Hurt, A.C., Holien, J.K., Parker, M., Kelso, A. and Barr, I.G. (2009) Zanamivir-resistant influenza viruses with a novel neuraminidase mutation. *The Journal of Virology*, **83(20)**, 10366-10373.

[17] Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., *et al.* (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science*, **325(5937)**, 197-201.

[18] Itoh, Y., Shinya, K., Kiso, M., Watanabe, T., Sakoda, Y., Hatta, M., Muramoto, Y., *et al.* (2009) In vitro and in vivo characterization of new swine-origin H1N1 influenza viruses. *Nature*, **460**, 1021-1025.

[19] Uhlendorff, J., Matrosovich, T., Klenk, H.D. and Matrosovich, M. (2009) Functional significance of the hemadsorption activity of influenza virus neuraminidase and its alteration in pandemic viruses. *Archives of Virology*, **154(6)**, 945-957.

[20] Sung, J.C., van Wynsberghe A.W., Amaro, R.E., Li, W.W. and McCammon, J.A. (2010) Role of secondary

sialic acid binding sites in influenza N1 neuraminidase. *Journal of the American Chemistry Society*, **132(9)**, 2883-2885.

[21] Steel, J., Lowen, A., Mubareka, S., Palese, P. and Baric, R. (2009) Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog*, **5**, e1000252.

[22] Subbarao, E.K., London, W. and Murphy, B.R. (1993) A single amino-acid in the Pb2-gene of influenza-A virus is a determinant of host range. *Journal of Virology*, **67**, 1761-1764.

[23] Mehle, A. and Doudna, J.A. (2009) Adaptive strategies of the influenza virus polymerase for replication in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **106(50)**, 21312-21316.

[24] Rolling, T., Koerner, I., Zimmermann, P., Holz, K., Haller, O., Staeheli, P. and Kochs, G. (2009) Adaptive mutations resulting in enhanced polymerase activity contribute to high virulence of influenza A virus in mice. *Journal of Virology*, **83 (13)**, 6673-6680.

[25] Liu, X. and Zhao, Y.P. (2010) Switch region for pathogenic structural change in conformational disease and its prediction. *PLoS One*, **5(1)**, e8441.

[26] Chen, W., Calvo, P.A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J.R. and Yewdell, J.W. (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nature Medicine*, **7**, 1306-1312.

[27] Lamb, R.A. and Takeda, M. (2001) Death by influenza virus protein. *Nature Medicine*, **7**, 1286-1288.

[28] Zell, R., Krumbholz, A., Eitner, A., Krieg, R., Halbhuber, K.J. and Wutzler, P. (2007) Prevalence of PB1-F2 of influenza A viruses. *Journal of General Virology*, **88**, 536-546.

[29] McAuley, J.L., Zhang, K. and McCullers, J.A. (2010) The effects of influenza A virus PB1-F2 protein on polymerase activity are strain specific and do not impact pathogenesis. *Journal of Virology*, **84(1)**, 558-564.

[30] Ramakrishnan, M.A., Gramer, M.R., Goyal, S.M. and Sreevatsan, S. (2009) A Serine12Stop mutation in PB1-F2 of the 2009 pandemic (H1N1) influenza A: A possible reason for its enhanced transmission and pathogenicity to humans. *Journal of Veterinary Science*, **10(4)**, 349-351.

[31] Trifonov, V. and Rabadan, R. (2009) The contribution of the pb1-f2 protein to the fitness of influenza a viruses and its recent evolution in the 2009 influenza A (H1N1) pandemic virus. *PLoS Current*: *Influenza*, **21**, RRN1006.

[32] Hale, B.G., Randall, R.E., Ortín, J. and Jackson, D. (2008) The multifunctional NS1 protein of influenza A viruses, *Journal of General Virology*, **89**, 2359-2376.

[33] Zhang, C.F., Yang, Y.T., Zhou, X.W., Liu, X.L., Song, H.B., He, Y.X. and Huang, P.T. (2010). Highly pathogenic avian influenza A virus H5N1 NS1 protein induces caspase-dependent apoptosis in human alveolar basal epithelial cells. *Virology Journal*, **7**, 51.

[34] Jackson, D., Hossain, M.J., Hickman, D., Perez, D.R. and Lamb, R.A. (2008) A new infl uenza virus virulence determinant: The NS1 protein four C-terminal residues modulate pathogenicity. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4381-4386.

[35] Seo, S.H., Hoffmann, E. and Webster, R.G. (2002) Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nature Medicine*, **8**, 950-954.

[36] Salomon, R., Franks, J., Govorkova, E.A., Ilyushina, N.A., Yen, H.L., Hulse-Post, D.J., Humberd, J., Trichet, M., Rehg, J.E., Webby, R.J., Webster, R.G. and Hoffmann, E. (2006) The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04, *Journal of Experimental Medicine*, **203(3)**, 689-697.

[37] Furuse, Y., Suzuki, A., Kamigaki, T. and Oshitani, H. (2009) Evolution of the M gene of the influenza A virus in different host species: Large-scale sequence analysis. *Virology Journal*, **6**, 67.

[38] Furuse, Y., Suzuki A. and Oshitani, H. (2009) Large-scale sequence analysis of M gene of influenza A viruses from different species: Mechanisms for emergence and spread of amantadine resistance. *Antimicrobial Agents and Chemotherapy*, **53(10)**, 4457-4463.

[39] Chen, G.W., Chang, S.C., Mok, C.K., Lo, Y.L., Kung, Y.N., *et al.* (2006) Genomic signatures of human versus avian influenza A viruses. *Emerging Infectious Diseases*, **12**, 1353-1360.

[40] Chen, G.W. and Shih, S.R. (2009) Genomic signatures of influenza A pandemic (H1N1) 2009, Virus. *Emerging Infectious Diseases*, **15**, 1897-1903.

[41] Pan, C., Cheung, B., Tan, S., Li, C., Li, L., *et al.* (2010) Genomic signature and mutation trend analysis of pandemic (H1N1) 2009, Influenza A virus. *PLoS One*, **5(3)**, e9549.

[42] Miotto, O., Heiny, A., Tan, T.W., August, J.T., Brusic, V. (2008) Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large-scale mutual information analysis. *BMC Bioinformatics*, **9**, S18.

[43] Miotto, O., Heiny, A.T., Albrecht, R., García-Sastre, A., Tan, T.W., August, J.T. and Brusic, V. (2010) Complete-proteome mapping of human influenza A adaptive mutations: implications for human transmissibility of zoonotic strains. *PLoS One*, **5(2)**, e9025.

[44] Finkelstein, D.B., Mukatira, S., Mehta, P.K., Obenauer, J.C., Su, X., Webster, R.G. and Naeve, C.W. (2007) Persistent host markers in pandemic and H5N1 influenza viruses. *Journal of Virology*, **81(19)**, 10292-10299.

[45] Allen, J.E., Gardner, S.N., Vitalis, E.A., Slezak, T.R. (2009) Conserved amino acid markers from past influenza pandemic strains. *BMC Microbioloy*, **9**, 77.

[46] Katoh, K., Kuma, K., Toh, H. and Miyata, T. (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33**, 511-518.

[47] Breiman, L. (2001) Random Forests. *Machine Learning*, **45(1)**, 5-32.

[48] Díaz-Uriarte, R. and Alvarez de Andrés, S. (2006) Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, **7**, 3.

[49] Archer, K.J. and Kimes, R.V. (2008) Empirical characterization of random forest variable importance measures. *Computational Statistics and Data Analysis*, **52**, 2249-2260.

[50] Reif, D.M. Motsinger, A.A., McKinney, B.A., Crowe, J.E. and Moore, J.H. (2006) Feature selection using a random forests classifier for the integrated analysis of

multiple data types. *Proceedings of* 2006 *IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology*, Toronto.

[51] Granittoa, P.M., Furlanellob, C., Biasiolia, F. and Gasperia, F. (2006) Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemometrics and Intelligent Laboratory Systems*, **83**, 83-90.

[52] Menze1, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W. and Hamprecht, F.A. (2009) A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, **10**, 213.

[53] Gao, D., Zhang, Y.X. and Zhao, Y.H. (2009) Random forest algorithm for classification of multi-wavelength data. *Research in Astronomy and Astrophysics*, **9(2)**, 220-226.

[54] Hu, W. (2009) Identifying predictive markers of chemosensitivity of breast cancer with random forests. *Journal of Biomedical Science and Engineering*, **3(1)**, 59-64.

[55] KováccaronOVá, A., Ruttkay-Nedecký, G., Karol HaverlíK1, I. and Janecccaronek, S. (2002) Sequence similarities and evolutionary relationships of influenza virus A hemagglutinins. *Virus Genes*, **24**, 57-63.

[56] Colman, P.M., Hoyne, P.A. and Lawrence, M.C. (1993) Sequence and structure alignment of paramyxovirus hemagglutinin-neuraminidase with influenza virus neuraminidase. *Journal of Virology*, **67**, 2972-2980.

[57] Maurer-Stroh, S. Ma, J.M., Lee, R.T.C., Sirota, F.L. and Eisenhaber, F. (2009) Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites. *Biology Direct*, **4**, 18.

[58] Liu, T. and Ye, Z.P. (2005) Attenuating mutations of the matrix gene of influenza A/WSN/33 Virus. *Journal of Virology*, **79(3)**, 1918-1923.

[59] Baudin, F., Petit, I., Weissenhorn, W. and Ruigrok, R.W.H. (2001) In vitro dissection of the membrane binding and RNP binding activities of influenza virus M1 protein. *Virology*, **281**, 102-108.

[60] Dua, Q.S., Wang, S.Q., Huang, R.B. and Chou, K.C. (2010) Computational 3D structures of drug-targeting proteins in the 2009-H1N1 influenza A virus. *Chemical Physics Letters*, **485**, 191-195.

[61] Ye Q., Krug R.M. and Tao Y.J. (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*, **444**, 1078-1082.

[62] Biswas, S.K., Boutz, P.L. and Nayak, D.P. (1998) Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *Journal of Virology*, **72**, 5493-5501.

[63] Lin, D., Lan, J. and Zhang, Z. (2007) Structure and function of the NS1 protein of influenza A virus. *Acta Biochim Biophys Sin* (*Shanghai*), **39(3)**, 155-162.

[64] Robb, N.C., Smith, M., Vreede, F.T. and Fodor, E. (2009) NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *Journal of General Virology*, **90**, 1398-1407.

[65] Iwatsuki-Horimoto, K., Horimoto, T., Fujii, Y. and Kawaoka, Y. (2004) Generation of influenza A virus NS2 (NEP) mutants with an altered nuclear export signal sequence. *Journal of Virology*, **78(18)**, 10149-10155.

[66] Yuan, P.W., Bartlam, M., Lou, Z.Y., Chen, S.D., Zhou, J., He, X.J., Lv, Z.Y., Ge, R.W., Li, X.M., Deng, T., Fodor, E., Rao, Z.H. and Liu, Y.F. (2009) Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature*, **458**, 909-913.

[67] Biswas, S.K. and Nayak, D.P. (1994) Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *Journal of Virology*, **68**, 1819-1826.

[68] Ohtsu, Y., Honda, Y., Sakata, Y., Kato, H. and Toyoda, T. (2002) Fine mapping of the subunit binding sites of influenza virus RNA polymerase. *Microbiology and Immunology*, **46**, 167-175.

[69] Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G. and Fanning, T.G. (2005) Characterization of the 1918 influenza virus polymerase genes. *Nature*, **437(7060)**, 889-893.

J. Biomedical Science and Engineering, 2010, 3, 602-611 **JBiSE**
doi:10.4236/jbise.2010.36082 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Medication compliance among mentally Ill patients in public clinics in Kingston and St. Andrew, Jamaica

Andrea E. Pusey-Murray[1], Paul A. Bourne[1], Stan Warren[2], Janet LaGrenade[1], Christopher A. D. Charles[3]

[1]Department of Community Health and Psychiatry, Faculty of Medical Sciences University of the West Indies, Kingston, Jamaica;
[2]Department of Sociology, Psychology and Social Work, The University of the West Indies, Mona, Jamaica;
[3]King Graduate School, Monroe College, Bronx, New York and Center for Victim Support, Harlem Hospital Center, New York, USA.
Email: paulbourne1@yahoo.com

Received 26 February 2010; received 5 April 2010; accepted 15 April 2010.

ABSTRACT

The Bellevue and the Hagley Park mental health outpatient clinics in Jamaica serve the majority of psychiatric patients in the country, but there is a dearth of research on medication compliance, which is a very important mental health issue. Medication compliance affects intervention outcomes. Therefore, this study seeks to examine medication compliance among psychiatric patients in Jamaica. A 33-item questionnaire which included items on demographics, health conditions, medication compliance and insightfulness was administered to a sample of 370 participants with a response rate of 93%. The majority of the participants have schizophrenia, followed by depression, bipolar disorder and drug-induced psychosis. The majority of the participants (65.3%) did not comply with their prescribed medication regimen. Medication compliance was significantly related to: gender ($P < 0.05$) where males were more likely to take the prescribed medication, family support ($P < 0.05$) where the participants who received family support (the majority being males) were more likely to take the prescribed medication, and insightfulness ($P < 0.05$) where the majority of participants with insightfulness were females. Locus on control was not statistically tested but a majority of the non- compliant participants reported that factors external to themselves had greater control over their disorder. Conclusion: There are three significant factors that explain the medication compliance of psychiatric patients in Jamaica. An important non-tested factor is locus of control so there needs to be more research to understand the range of factors that can inform and improve patient education about medication compliance.

Keywords: Medication Compliance; Mentally Ill; Public Clinics; Kingston; St. Andrew

1. INTRODUCTION

Mental illness such as schizophrenia, bipolar disorders and uni-polar depression presents a serious health care problem in Caribbean countries such as Jamaica and worldwide. The economic, clinical, and personal burdens associated with schizophrenia make it a leading public health problem. The earliest epidemiological report on schizophrenia in the Caribbean was from Jamaica [1], and indicated that the annual population incidence rate was 150 per 100,000. Burke [2] confirmed this finding, and reported an annual incidence rate of 136 per 100,000. Both studies were based on mental hospital admission rates, and it is likely that these studies failed to access the total number of admissions for the island. A study of psychiatric admissions to private and public hospitals across Jamaica [3] reported that the incidence rate of schizophrenia had fallen from 69 per 100,000 in 1960 to 35 per 100,000 in 1990.

Antipsychotic drugs, the most effective treatment for acute episodes or exacerbations of schizophrenic illness [4], allow many patients to leave institutions and live in the community [5]. Rates of relapse among patients with schizophrenia who receive medication are two to three times lower than those among patients receiving placebo [6], and non-compliance increases the frequency of acute psychotic episodes and psychiatric hospitalization [7]. Although antipsychotic drugs can have serious adverse effects [8], many clinicians prescribe them at moderate doses for as long as possible to prevent relapse. In addition to antipsychotic agents, patients with schizophrenia may receive lithium, antidepressants, or benzodiazepines for concomitant psychiatric disorders [4].

Mental health research and pharmaceutical innovation have developed a class of drugs referred to as "atypical

antipsychotics", which are used for the treatment of severe schizophrenics who are considered treatment- resistant to traditional or conventional antipsychotic medications, or who experience side-effects severe enough to require that the patient discontinue use of conventional antipsychotics [9]. These atypical antipsychotics are tremendously effective in combating the symptoms of schizophrenia while avoiding the severe side-effects often experienced through treatment with conventional antipsychotics [10]. In Jamaica medications for treating mental health problems are limited and their availability varies district by district. Older antipsychotic drugs are usually available, but there is only limited availability of the newer, more expensive anti-psychotics. In the North East, some patients are treated with clozapine by special arrangement by the district psychiatrist [11].

Adherence to drug regimen is a very important factor for improvement. Adherence may be defined as the extent to which a person's behaviour confirms to medical or health advice [12]. Patient who do not follow the treatment schedule and drug regimens prescribed to them by physician can be described as noncompliant or not adherent [13]. Medication compliance among bipolar disorder patients is related to the constructs of the health belief model (HBM) such as benefits and barriers, susceptibility and perceived seriousness [14]. However, the compliance/non-compliance of patients is only moderately predicted by the HBM. The HBM does not cover some determinants of medication compliance such as social influence and treatment alliance [15]. Patients with bipolar disorder who are medication compliant perceive the quality of their life to be higher, have greater resources to cope with stress and have a stronger belief that their behaviour controls their health status, unlike non-compliant patients [16]. The patients' perception of the medication is important because cognitive dissonance suggests that the perceptual properties of the medication have particular meanings for the patients. These meanings can support or distract the patients from medication compliance [17]. Patients with a major depressive disorder who have a superior medication compliance index are more likely to show improved scores on the Hamilton Depression Scale [18]. The medication events monitoring system (MEMS) used in a study with schizophrenic disorder patients reveals a 63% mean compliance rate for the first month and a decline from 56% to 45% over the next five months. The medication compliance of these patients can be monitored with electronic monitoring devices, but data recovery and compliance require methodological improvements [19].

Among patients with psychotic disorders there was a significant relationship between medication adherence and involuntary admission, substance abuse, not graduating from school, and a history of abusive behaviour. Patients with paranoid or negative symptoms were less compliant in taking their medication. However, patients who were changed from a typical to an atypical antipsychotic medication were more compliant than patients who remained on the typical antipsychotic medication. The patients who had higher medication compliance experienced much greater improvement of their psychiatric symptoms [20]. Medication self-management among patients with schizophrenia or affective disorder is important for medication compliance. However, patients who are self-managing their medication, guided by the principles of motivational interviewing, have better attitudes towards medications and insights about their illness compared to controls, but the group difference is not significant [21].

Patients with schizophrenia, bipolar disorder and unipolar depression, who were educated about the nature of their disorder and its pharmacological management, were more compliant in outpatient follow-ups, and displayed less fear of being addicted to the medication and dealing with the side effects of the drug [22]. Severely mentally ill patients who were a part of the Medication Usage Skills for Effectiveness (MUSE) Program that taught cueing, life skill techniques, visual feedback and about the data displayed on the medication cap showed an overall mean compliance rate of 76%, compared to a 57% compliance rate for the control group [23].

The problem of poor adherence to therapeutic regimen has been a matter of concern to the professionals for years. The paucity of data on cost-related medication adherence problems has important implications not only for estimating their clinical significance, but also for understanding the extent to which adherence problems vary across socioeconomic groups. To our knowledge there has not been any study that has examined drug compliance among mentally ill patients in Jamaica who are treated with oral antipsychotics. The aim of this study was to examine the medication compliance of psychiatric patients in Jamaica.

2. MATERIALS AND METHODS

2.1. Sample

In Jamaica the single mental hospital is Bellevue Mental Hospital in Kingston. Its three acute care wards of 50 beds each serve a designated geographic catchment area of the parishes of Kingston, St. Andrew, and St. Catherine, a population that is 70 percent urban [24]. Of the hospital's patients, 40% are older than 65 years, 60% of them are regarded as chronically ill, and 300 of them have lived in the hospital for most of their adult lives. The University Hospital West Indies (UHWI), also located in Kingston, has a 20-bed psychiatric unit. It is an acute unit with an average length of stay of 15 days. Ja-

maica's rural and western regions are served partly by Cornwall Regional Hospital. This 60-bed unit offers a full range of services, but only 30 beds are used because of staff shortages [11].

The Bellevue Hospital in Kingston and the Hagley Park Health Centre in Saint Andrew investigated in this study are two of the main psychiatric clinics in Jamaica. They attend to in excess of 90% of the psychiatric out-patient cases in the country. The current study used cross-sectional survey to gather the data from participants. Using a sampling error of ± 3.0% and a 95% confidence interval, the calculated sample size was 370 participants. Based on the proportionality of cases seen over the last 2 months, this was used to proportion the sample size from each institution: 200 participants from the Bellevue Hospital Clinic and 170 from the Hagley Park Health Centre. The response rate was 93.0% with 186 participants from the Bellevue Outpatient Clinic and 158 from the Hagley Park Health Centre.

Data was collected over a three-month period, from March to May, 2008. A team of trained data collectors were retrained in keeping with the peculiarities of the task. The researcher was a part of the process, and regular checks were done to ensure consistency among interviewers. The inclusion/exclusion criterion was based on mental illness and medication use. Mentally ill patients who were visiting the clinics for the first time were excluded from the study because they would not have developed a pattern of medication compliance (or non-compliance).

2.2. The Instrument

A 33-item questionnaire was used to collect the data. The questions included: 1) demographic information, 2) medication compliance, 3) reasons for taking (or not taking) medication, 4) accessibility and affordability of medication, 5) use and prevalence of home remedies, 6) family support and 7) perception of the management of mental illness (Appendix 1).

2.2.1. Test-and-Pretesting the Instrument

The instrument was developed in consultation with academics at the University of the West Indies, Mona, Jamaica, caregivers/guardians at the Harbour View Health Centre in St. Andrew, health practitioners, and non-prospective mentally ill patients for the sample. The instrument was modified in keeping with the comments of the various stakeholders, with specific emphasis on mentally ill patients' perspectives, suggestions and recommendations. A pilot of 50 questionnaires was tested at the Harbour View Health Centre on ill patients, and they provided useful information regarding ambiguities, which was then used to modify the final instrument.

2.3. Measure

Non-compliance in medication denotes the failure or refusal of an individual to take the prescribed medication(s) as recommended by the medical practitioner. Compliance is adhering to the prescription of oral or other forms of medication as stipulated by the medical practitioner. It was measured using the item "I take my medication as instructed by the doctor." The options ranged from always, most times, occasionally and rarely, to never. Compliance for this study is taken to be only a response of 'always'.

Health condition is an illness or ailment diagnosed by the medical practitioner, and this was taken from the medical record of the participant. Insightfulness was measured by the questions: 1) "Do you have a sense of the likelihood of being ill or of when your illness is likely to occur?" 2) Should medication only be taken when one is experiencing symptoms of illness?" and 3) "Will counselling on how to take medications help you to take them better?" The options were yes, no, or unsure. If the answers to all three questions were yes, it was coded as insight, if no was selected in any question it was coded as lack of insight, and unsure was coded as undecided. Non-responses were treated as missing, and not included in any categorization.

2.4. Ethics and Informed Consent

The survey was submitted and approved by the University of the West Indies Medical Faculty's Ethics Committee. Participants and/or caregivers gave voluntary consent to their participation in the study. In order to ensure confidentiality, the personal information (*i.e.* name, address) of the participants was taken from the questionnaires and discarded, after which the other information was entered and stored for data-analysis.

2.5. Statistical

Percentages were used to provide background information on the demographic characteristics of the sample, knowledge of the medication and self-reported information. Chi-square tests were utilized to examine whether statistical associations existed between non-metric dependent and independent variables. A *P*-value of < 5% (*i.e.* 95% confidence interval) was used to determine the statistical associations between the variables.

3. RESULTS

3.1. Demographic Characteristic of Sample

There were 344 participants in the sample, of which 53.7% (n = 185) came from the Day Clinic at Bellevue Hospital and 64.2% (n = 159) came from the Hagley Park Health Centre. Most of the sample were diagnosed

with schizophrenic disorder (58.7%, n = 202), with other health conditions being depression (13.1%, n = 45), bipolar disorder (10.5%, n = 36), and drug-induced psychosis (4.7%, n = 16). Sixty-two percent (n = 212) of the sample had secondary level education; 64.8% were unemployed; 93.0% lived with either parents, partner, children, sibling or in a nursing home; 84.1% had family support; 80.5% were not in a union relationship (unmarried, 58.7%; separated, 13.4%: divorced, 4.6%; widowed, 3.8%); 10.5% were married; 44.7% complied with the instruments of the medical practitioners in regards to medication usage; and the mean age was 43.6 years (± 1.5 years).

The most common reasons among those who did not comply with the specifications of the medications (55.3%, n = 189) were 'medication makes me feel too drowsy' – 19.0%; 'I was out of medication' – 18.0%; 'I forgot' – 16.9%; and 'medications make me feel worse' – 10.1%.

The majority of the participants (76.2%, n = 262) indicated that they were able to purchase the medications. The participants indicated that if the medication was not available at the hospital or health centre, they would purchased them from a private pharmacy (65.1%), wait for the next appointment in order to see if it was now available (16.6%), and used a home remedy (0.6%).

According to **Table 1**, no significant statistical relationship was found between the type of medical facility utilized and, gender, employment status, education or age ($P > 0.05$); nor for compliance ($P > 0.05$) or insightfulness ($P > 0.05$). However, an association did exist between the type of medical facility utilized and marital status ($P < 0.05$). Twenty-four percent of those who utilized the Hagley Park Health Care Centre were in intimate unions (relationships), compared to 15.6% of those who visited the Bellevue Hospital.

Gender was the only demographic variable that was associated with diagnosed health conditions ($P < 0.05$; **Table 2**). Most of the participants were diagnosed with schizophrenic disorder, with more males than females with the illness. Females were more likely to be diagnosed with bipolar disorder and depression than males. However, there were more males diagnosed with drug-induced psychosis (5.9%) than females (2.8%).

A statistical relationship was found between gender and compliance/non-compliance of participants ($P < 0.05$). Males were more likely to comply with the specifications of their medication (51%) than females (35.9%). In addition, significant statistical associations were found between family support and compliance/non-compliance ($P < 0.05$), and family support and gender of respondents ($P < 0.05$). Males (59.7%) had more family support compared with females (40.3%), and more of those with family support complied with the specifications of the medications (53.3%).

Table 1. Demographic characteristics of sample.

Variable	Hagley Park Clinic (n = 158) n (%)	Outpatient Day clinic at Bellevue (n = 186) n (%)	*P*
Gender			NS
Male	92 (58.2)	110 (59.1)	
Female	66 (41.8)	76 (40.9)	
Marital status			< 0.05
Married	16 (10.1)	20 (10.8)	
Common-law	22 (13.9)	9 (4.8)	
Widowed	8 (5.1)	5 (2.9)	
Divorced	6 (3.8)	10 (5.4)	
Separated	16 (10.1)	30 (16.1)	
Unmarried	90 (57.0)	112 (60.0)	
Employment status			NS
Employed	48 (28.5)	54 (29.0)	
Unemployed	112 (70.9)	128 (68.8)	
Other	1 (0.6)	4 (2.2)	
Education			NS
Primary or below	36 (30.4)	66 (35.5)	
Secondary (including vocational)	100 (63.3)	112 (60.2)	
Tertiary	10 (6.3)	8 (4.3)	
Compliance			NS
Yes	65 (41.1)	88 (47.8)	
No	93 (58.9)	96 (52.2)	
Insightfulness			NS
Insight	83 (52.5)	83 (44.9)	
Undecided	3 (1.9)	15 (8.1)	
Lack insight	72 (45.6)	87 (47.0)	
Age Mean (SD)	42.1 years (1.4)	44.1 years (1.2)	NS

NS not significant

Approximate thirty seven percent (37.3%) of the participants were knowledgeable about the medication they were taking. Sixty-two percent of those who were knowledgeable about the medication indicated that they received the information from their medical doctors, 16.9% from pharmacists and 15.4% from nurses. Concurrently, when they were asked to state who was responsible for controlling their illness, one-half (50%) indicated their medical doctor was responsible for controlling their illness, nurse 10.5%, pharmacists 9.0%, God 9.0%, self 4.1%, and other 3.5%.

The majority of the participants indicated that apart from medication and other health care professionals, other factors that controlled their illness were: praying (52.4%), cigarettes (11.0%), smoking marijuana (10.5%), fasting (5.05), leaves from sour-sop tree (3.5%) and Obeah (0.3%). A cross-tabulation between gender and insightfulness revealed a significant statistical association ($P < 0.05$). Lacking insightfulness was greatest for males (68.6%) and insightfulness was greatest among females (50.0%).

Table 2. Diagnosed mental illness by demographic characteristic

Characteristic	Health condition						P
	Schizophrenia n = 202	Bipolar Disorder n = 36	Depression n = 45	Drug Induced psychosis n = 16	Other n = 7	Undiagnosed n = 38	
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Gender							< 0.05
Male	125 (61.9)	19 (9.4)	22 (10.9)	12 (5.9)	2 (1.0)	22 (10.9)	
Female	77 (54.2)	17 (12.0)	23 (16.2)	4 (2.8)	5 (3.5)	16 (11.3)	
Age group							NS
≤ 20 years	1 (0.5)	1 (2.8)	12 (26.7)	1 (6.2)	0 (0.0)	0 (0.0)	
21 – 39 years	90 (44.6)	18 (50.0)	23 (51.1)	12 (75.0)	4 (57.1)	10 (26.3)	
40 – 59 years	87 (43.1)	17 (47.2)	10 (22.2)	3 (18.8)	2 (28.5)	23 (60.5)	
60+ years	24 (11.9)	0 (0.0)	0 (0.0)	0 (0.0)	1 (14.2)	5 (13.2)	
Marital status							NS
Married	17 (8.4)	5 (13.9)	9 (20.0)	3 (18.7)	2 (28.6)	1 (2.6)	
Common-law	14 (6.9)	4 (11.1)	8 (17.8)	2 (12.5)	1 (14.3)	1 (2.6)	
Widowed	8 (4.0)	1 (2.8)	1 (2.2)	0 (0.0)	0 (0.0)	2 (5.3)	
Divorced	10 (4.9)	2 (5.6)	2 (4.4)	1 (6.3)	0 (0.0)	2 (5.3)	
Separated	28 (13.9)	4 (11.1)	5 (11.2)	2 (12.5)	1 (14.3)	6 (15.8)	
Unmarried	125 (61.9)	20 (55.5)	20 (44.4)	8 (50.0)	3 (42.8)	26 (68.4)	
Employment Status							NS
Employed	48 (48.5)	14 (14.1)	13 (13.1)	6 (6.1)	2 (2.0)	16 (16.2)	
Unemployed	150 (62.5)	21 (8.8)	32 (13.3)	10 (4.2)	5 (2.1)	22 (9.2)	
Other	4 (80.0)	1 (20.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Education							NS
Primary or below	73 (36.1)	6 (16.7)	16 (35.6)	3 (18.8)	3 (42.8)	13 (34.2)	
Secondary	119 (58.9)	27 (75.0)	27 (60.0)	11 (68.8)	4 (57.1)	24 (63.2)	
Tertiary	10 (5.0)	3 (8.3)	2 (4.4)	2 (12.5)	0 (0.0)	1 (2.6)	

NS not significant

4. DISCUSSION

In this study, the majority of the participants were diagnosed with schizophrenia, followed by depression, bipolar disorder and drug induced psychosis which is consistent with other findings in the medical literature [1,2]. The majority (84.1%) of the participants have family support, but of this amount only 54% complied with taking the medication as prescribed. This suggests that family support is a significant determinant of medication compliance. This finding corroborates that of Garcia and colleagues, that the support from family caregivers predicts the use of psychiatric medication [25]. According to Hickling [26] family members of mentally ill patients in Jamaica take on the ongoing tasks of monitoring the patient and supervising his or her medication program. Along with the patient and the community mental health service providers, family members participate in formulating and carrying out plans for treatment and vocational rehabilitation. Although such family involvement is not unique, the extent to which the Jamaican community mental health service has to rely on the extended family is noteworthy.

Most (80.5%) of the participants were not in a stable intimate partner relationship, a fact which could have influenced the amount and the quality of social support they received. Possible explanations could be the negative stigma associated with mental disorders in Jamaica and their inability to provide financially in such a relationship, since 64.8% of the participants were unemployed. However, in Jamaica, the fear and stigma associated with mental illness has been greatly reduced, mental illness is openly discussed in the media, and patients are comfortable receiving treatment for a variety of psychiatric conditions in public and private community treatment facilities [27].

Medication compliance among the participants in our study was very low with only 44.7% taking their medication. Among those participants who did not comply with the medication specifications (55.3%), 19% of this group indicated they experienced drowsiness, and 10.1% indicated that they felt worse. According to Feuertein *et al.* [28] estimates of noncompliance ranges between 4% and 92%, with average from 30 to 35 percent. The reason for non- compliance among the participants in this study could be due to a number of factors such as: discomfort resulting from treatment (side effects of medication), expense of treatment, problems of filling a prescription, decision based on personal value, judgment or religious or cultural beliefs about the advantages and

disadvantages of the proposed treatment, maladaptive personality, traits or coping style (for example the denial of illness), the stigma attached to drugs for mental health conditions and lack of understanding about the nature of the illness [29].

According to Demyttenwere, the dropout rate of psychiatric patients is attributed to various factors, such as illness and patient's characteristics, time taken to improve or poor doctor-patient relationship [30]. Mann indicated that the variance between what the doctors prescribe and what the patients take can be reduced by the doctor listening more carefully to the patients and addressing their concerns about the side effects of the medication [31]. In addition, educating the patients about their disorder, pharmacological management and addressing their concerns and fears of being addicted to the drug increases their compliance [32]. In our study, 18% among the noncompliant participants suggested that their medication was finished, and 16.9% indicated that they forgot to take their medication. These patients can be instructed about life skill techniques, cueing, about the medication data displayed on the cap, and about audiovisual feedback as this kind of instruction increases medication compliance among psychiatric patients [33]. Interventions to increase medication compliance through patient education in Jamaica are necessary because only a minority (37.3%) of the participants reported that they were knowledgeable about the medication they were taking, and the majority (76.2%) of the participants in dicated that they were able to buy their medication, so the cost of the drug was not a prohibitive factor. Medical doctors can play a major role in this regard because 62.0% of the participants who were knowledgeable

Table 3. Compliance by demographic characteristic and health condition.

Variable	Compliance n = 153 n (%)	Non-compliance n = 189 n (%)	P
Gender			< 0.05
Male	102 (66.7)	98 (51.9)	
Female	51 (33.3)	91 (48.1)	
Age group			NS
≤ 20 years	1 (0.7)	2 (1.1)	
21 – 39 years	59 (38.6)	87 (46.0)	
40 – 59 years	68 (44.4)	85 (45.0)	
60+ years	25 (16.3)	15 (7.9)	
Marital status			NS
Married	11 (7.2)	28 (14.8)	
Common-law	16 (10.5)	12 (6.4)	
Widowed	8 (5.2)	5 (2.6)	
Divorced	6 (3.9)	10 (5.3)	
Separated	17 (11.1)	28 (14.8)	
Unmarried	95 (62.1)	106 (56.1)	
Employment status			NS
Employed	40 (26.1)	59 (31.2)	
Unemployed	112 (73.2)	127 (67.2)	
Other	1 (0.7)	3 (1.6)	
Education			NS
Primary or below	56 (36.6)	56 (29.6)	
Secondary	88 (57.5)	124 (65.6)	
Tertiary	9 (5.9)	9 (4.8)	
Health condition			NS
Schizophrenia	92 (60.1)	108 (57.1)	
Bipolar disorder	14 (9.1)	22 (11.6)	
Depression	20 (13.1)	25 (13.2)	
Drug-induced psychosis	5 (3.3)	11 (5.8)	
Other	5 (3.3)	2 (1.2)	
Undiagnosed	17 (11.1)	21 (11.1)	

NS not significant

Table 4. Family support by demographic characteristic and health condition.

Variable	Family support Yes n = 290 n (%)	Family support No n = 37 n (%)	P
Gender			< 0.05
Male	173 (59.7)	15 (40.5)	
Female	117(40.3)	22 (59.5)	
Age group			NS
≤ 20 years	3 (1.0)	11 (29.7)	
21 – 39 years	129 (44.5)	20 (54.1)	
40 – 59 years	126 (43.5)	6 (16.2)	
60+ years	32 (11.0)	0 (0.0)	
Marital status			NS
Married	26 (9.0)	2 (5.4)	
Common-law	36 (12.4)	2 (5.4)	
Widowed	15 (5.2)	2 (5.4)	
Divorced	11 (3.8)	2 (5.4)	
Separated	31 (10.7)	5 (13.5)	
Never married	171 (58.9)	24 (64.9)	
Employment status			NS
Employed	74 (25.5)	18 (48.6)	
Unemployed	213 (73.5)	18 (48.7)	
Other	3 (1.0	1 (2.7)	
Education			NS
Primary or below	95 (32.8)	13 (35.1)	
Secondary	181 (62.4)	22 (59.5)	
Tertiary	14 (4.8)	2 (5.4)	
Health condition			NS
Schizophrenia	173 (59.7)	15 (40.5)	
Bipolar disorder	31 (10.7)	4 (10.8)	
Depression	36 (12.4)	8 (21.6)	
Drug-induced psychosis	13 (4.5)	3 (8.2)	
Other	7 (2.4)	0 (0.0)	
Undiagnosed	30(10.3)	7 (18.9)	
Compliance			
Yes	154 (53.3)	10 (27.8)	<0.05
No	135 (46.7)	26 (72.2)	

NS not significant

about their medication stated that they received information from their doctor.

There is a significant relationship between marital status and the clinic used, because 24.0% of the participants used the Hagley Park Health Centre Clinic compared to 15.6% for the Bellevue Clinic. This is a surprising finding which requires more research. One possible explanation for this finding is that the Bellevue Hospital is the oldest and only major psychiatric hospital in Jamaica. It is heavily stigmatized given the very negative perception Jamaicans have toward mental disorders and people with mental disorders. Therefore, it is possible that the married participants attended the Hagley Park Clinic to avoid the very negative stigma associated with the Bellevue Clinic.

Many chronically ill patients take less of their medication than has been prescribed, owing to cost concerns, especially those patients with low incomes, multiple chronic health problems, or no prescription drug coverage [34]. The consequences of cost-related medication underuse include increased emergency department visits, psychiatric admissions, nursing home admissions, as well as decreased health status [35]. In Jamaica there is government assistance for low-income patients. According to McKenzie, few patients ask for it. In his study, in one clinic, none of those who were prescribed medication (at least 35) asked for assistance, despite the fact that none of them were working. It is customary for families to pay for their relatives' medication. Caregivers considered a lack of money and the need to pay for prescriptions as a deterrent for attending the clinic [11].

There is a significant relationship between gender and psychiatric disorder which is an unexpected finding. A majority of males and females were diagnosed with schizophrenia, with a higher incidence among males. More males than females were diagnosed with drug-induced psychosis, while more females were diagnosed with depression and bipolar disorder than males. Further research is required to understand the nuances of the relationship between gender and mental disorder which is an under-researched area. In addition, there is a related significant relationship between gender and medication compliance, in which males were more likely to comply with their medication. This finding may be explained by the significant relationship between insightfulness about medication and gender, where males account for 68.0% of the participants who lacked insightfulness. This finding suggests that it is possible that there is greater medication compliance among males, because they are less insightful about their medication and its side effects than females. It is also possible that since the majority (59.7%) of participants with family support are males, the social support increases medication compliance in this group [25].

In the present study, non-compliance of medication-taking among those with bipolar disorder and depression was over 80%. Lack of insight [36] is a factor associated with poor drug compliance. Hummer and Fleischacker [37] explained that non-compliance with medication is owing to the patients' perception that the illness is not serious to enough to warrant treatment compliance. Conversely, a study by Cramer and Rosenheck found that 58% of those diagnosed with psychosis and 65% of those with depression complied with the treatment prescription [38], which is significantly higher than a similar group in Jamaica. One researcher admitted that the side effects of neuroleptics are real, but can be managed [39], which clearly is accepted by those diagnosed with drug psychosis and bipolar disorder in Jamaica

Patients' belief in their ability to control their illness is very important for medication compliance. However, the participants in our study believed in a range of external factors apart from medication that could be used to control their disorder. These external factors are: praying, use of Obeah (Jamaican witchcraft), fasting, smoking marijuana or use of sour-sop leaves, and smoking cigarettes. There are also some participants (69.5%) who believed that doctors, nurses and pharmacists were responsible for controlling their disorder. Thus the majority of the participants in our study displayed an external locus of control (rather than an internal locus of control) which downplays the belief that they themselves have major control over what happens in their lives and their wellbeing. This can clearly has an effect on medication compliance. These findings are corroborated by the findings of Volis and colleagues, that the locus of control is important in medication compliance because it mediates the relationship between medication compliance and social support [40,41].

This study has contributed to the literature by unearthing a range of factors that are significantly related to medication compliance among psychiatric patients in Jamaica, which increases our understanding of this very important health issue. There are some limitations to our study. While our sample was taken from the two public mental health clinics that treat the majority of psychiatric cases in the country, these clinics are located in the metropolitan region of Kingston. Therefore, it is possible that our sample did not capture psychiatric patients from other urban centres and rural areas. Therefore only cautious generalizations should be made. In addition, there is the possibility of social desirability bias, where some of the patients told the interviewers what they wanted to hear to get their approval. It is also possible that the patients' mental disorders affected the accuracy of their self-reports. Despite the aforementioned fact, medication non-compliance among schizophrenic patients in this research was low and in keeping with another in a non-Caribbean nation which found that it ranges between

24-90% [42]. This work also concurs with the findings of the Jamaican Ministry of Health, which found that medication compliance among schizophrenia patients in Jamaica was 69% [43].

6. CONCLUSIONS

Medication compliance among males is about average, but is extremely low among females. The majority of the participants with an average age of 43.6 years from the Bellevue and Hagley Park mental health outpatient clinics do not comply with their medication regimen for schizophrenia, bipolar disorder, depression and psychosis among other disorders. This non-compliance can be explained by three significant factors. These are gender, which is related to the factor of reduced insightfulness about taking medication and family support. Non- compliance may also be explained by the locus of control which was not tested in this study, but the majority of the noncompliant participants believe that factors external to them have greater control over their illness than they do. The current study highlights the challenge for public health practitioners and policy makers in addressing this high non-compliance of medication among the mentally ill in Jamaica. A pertinent finding of this study is the fact that the level of education did not change compliance (or non-compliance) among mentally ill patients, suggesting that there is a need for more research to unearth the range of factors that influence medication compliance, the belief system of mentally ill patients, an examination of alternative approaches to the treatment of mental illness, and a social intervention programme that is geared towards holistic education strategies in patient care. What is evident from the study is the fact that medication compliance can be explained by 1) the perception of the severity of the illness and the usefulness of the relevant medication, and 2) the perceived side-effects of prescribed medications.

Medication non-compliance places mentally ill patients at great risk of exacerbation of their symptoms, homelessness and interruptions in their daily lives, so much so that this has become a public health concern which must be addressed with urgency and care. Family support emerged as a positive determinant of medication compliance, suggesting that public health practitioners must begin to explore the role of social support in treating mentally ill individuals, as well as aiding in the drive for increased medication compliance among these individuals.

REFERENCES

[1] Royes, K. (1962) The incidence and features of psychosis in a Caribbean community. *Journal of Social Psychiatry*, **2**, 1121-1125.

[2] Burke, A.W. (1974) First admissions and planning in Jamaica. *Journal of Social Psychiatry*, **9**, 39-45.

[3] Hickling, F.W. (1991) Psychiatric hospital admission rates in Jamaica, 1971 and 1988. *British Journal of Psychiatry*, **159**, 817-821.

[4] Kane, J.M. and Marder, S.R. (1993) Psychopharmacologic treatment of schizophrenia. *Schizophr Bull*, **19**, 287-302.

[5] Crane, G.E. (1973) Clinical psychopharmacology in its 20th year: Late, unanticipated effects of neuroleptics may limit their use in psychiatry. *Science*, **181**, 124-128.

[6] Davis, J.M. and Gierl, B. (1984) Pharmacological treatment in the care of schizophrenic patients. In: Bellack, A.S., Ed. Treatment and care for schizophrenia. Grune & Stratton, Orlando.

[7] Johnson, D.A. (1988) Drug treatment of schizophrenia. In: Bebbington, P. and McGuffin, P., Eds. Schizophrenia: The major issues. Heinemann Professional, Oxford, 158-171.

[8] Ellenbroek, B.A. (1993) Treatment of schizophrenia: A clinical and preclinical evaluation of neuroleptic drugs. *Pharmacology & Therapeutics*, **57**, 1-78.

[9] Kane, J., Honigfeld, G., Singer, J., *et al*. (1988) Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry*, **45**, 789-796.

[10] Tuunainen, A., Wahlbeck, K. and Gilbody, S.M. (2002) Newer atypical antipsychotic medication in comparison to clozapine: A systematic review of randomised trials. *Schizophrenia Research*, **56**, 1-10.

[11] McKenzie, K. Jamaica: Community mental health services. http://www.publications.paho.org/english/Jamaica_CD_183. pdf

[12] Bruer, J.T. (1982) Methodological rigor and citation frequency in patient compliance literature. *American Journal of Public Health*, **72**, 911-1123.

[13] Razali, M.S. and Yahya, H. (1995) Compliance with treatment in Schizophrenia: A drug intervention program in a developing country. *Acta Psychiatrica Scandinavica*, **91**, 331-335.

[14] Montgomery, L.A. (2002) The relationship between health belief model constructs and medication compliance in the treatment of bipolar disorder. *Dissertation Abstracts International*: *Section B*: *The Sciences and Engineering*, **62(12-B)**, 5974.

[15] Cohen, N.L., Parikh, S.V. and Kennedy, S.H. (2000) Medication compliance in mood disorders: Relevance of the Health Belief Model and other determinants. *Primary Care Psychiatry*, **6(3)**, 101-110.

[16] Lund, V.E. (2000) Perceived quality of life for persons with bipolar disorder: The role of medication compliance, family and health stress, level of coping, and health locus of control. *Dissertation Abstracts International*: *Section B*: *The Sciences and Engineering*, **60(7-B)**, 3198.

[17] Buckalew, L.W. and Sallis, R.E. (1986) Patient compliance and medication perception. *Journal of Clinical Psychology*, **42(1)**, 49-53.

[18] Thompson, C., Peveler, R.C., Stephenson, D. and McKendrick, J. (2000) Compliance with antidepressant medication in the treatment of major depressive disorders in

primary care: A randomized comparison of fluoextine and a tricyclic antidepressant. *The American Journal of Psychiatry*, **157(3)**, 338-343.

[19] Diaz, E., Levine, H.B., Sullivan, M.C., *et al.* (2001) Use of the Medication Events Monitoring System to estimate medication compliance in patients with schizophrenia. *Journal of Psychiatry and Neuroscience*, **26(4),** 325-329.

[20] Janssen, B., Gaebel, W., Haerter, M., Komaharadi, F., Lindel, B. and Weinmann, S. (2006) Evaluation of factors influencing medication compliance in patient treatment of psychotic disorders. *Psychopharmacology*, **187 (2)**, 229-236.

[21] Hayward, P., Chan, N., Kemp, R. and Youle, S. (1995) Medication self-management: A preliminary report on an intervention to improve medication compliance. *Journal of Mental Health*, **4(5)**, 511-517.

[22] Seltzer, A., Roncari, I. and Garfinkel, P.E. (1980) The effect of patient education on medication compliance. *Canadian Journal of Psychiatry*, **25(8)**, 638-645.

[23] Cramer, J.A. and Rosenheck, R. (1999) Enhancing medication compliance for people with serious mental illness. *Journal of Nervous and Mental Disease*, **187(1)**, 53-55.

[24] (1993) Demographic Statistics, Jamaica 1992. Statistical Institute of Jamaica, Kingston.

[25] Garcia, J.I., Chang, C.L., Young, J.S., Lopez, S.R. and Jenkins, J.H. (2006) Family support predicts psychiatric medication use among Mexican American individuals with schizophrenia. *Social Psychiatry and Psychiatric Epidemiology*, **41(8)**, 624-631.

[26] Hickling, F.W. (1975) An analysis of the 1873 Mental Hospital Law in Jamaica and its effect on mental treatment facilities. *Jamaica Law Journal*, 64-69.

[27] Hickling, F.W. (1994) Community psychiatry and deinstitutionalization in Jamaica. *Hospital & Community Psychiatry*, **45**, 1122-1126.

[28] Feuertein, M., Labbe, E.E. and Kuegmierezyk, A.R. (1986) Health psychology: A psychobiological perspective. Pleneum Press, New York.

[29] Adewuya, A.O., Ola, B.A., Mosaku, S.K., Fatoye, F.O. and Eegunranti A.B. (2006) Attitude towards antipsychotics among outpatients with schizophrenia in Nigeria. *Acta Psychiatrica Scandinavica*, **113**, 207-211.

[30] Demyttenwere, K. (1997). Compliance during treatment with antidepressants. *Journal of Affective Disorder*, **43**, 27-39.

[31] Mann, J.J. (1986) How medication compliance affects outcome. *Psychiatric Annals*, **16(10)**, 567-570.

[32] Seltzer, A., Roncari, I. and Garfinkel, P.E. (1980) The effect of patient education on medication compliance. *Canadian Journal of Psychiatry*, **25(8)**, 638-645.

[33] Cramer, J.A. and Rosenheck, R. (1999) Enhancing medication compliance for people with serious mental illness. *Journal of Nervous and Mental Disease*, **187(1)**, 53-55.

[34] Safran, D.G., Neuman, P., Schoen, C., *et al.* (2002) Prescription drug coverage and seniors: how well are states closing the gap? *Health Affairs*, W253-W268.

[35] Tamblyn, R., Laprise, R., Hanley, J.A., *et al.* (2001) Adverse events associated with prescription drug cost-sharing among poor and elderly persons. *Journal of the American Medical Association*, **285**, 421-429.

[36] Kane, J.M. (1985) Compliance issues in outpatient treatment. *Journal of Clinical Psychopharmacology*, **5**, 22-27.

[37] Hummer, M. and Fleischhacker, W.W. (1999) Ways of improving compliance. In: Lader, M. and Naber, D., Eds., Difficult clinical problems in psychiatry. Martin Dunitz, London, 229-238.

[38] Cramer, J.A. and Rosenheck, R. (1998) Compliance with medication regimens for mental and physical disorders. *Psychiatric Services*, **49**, 196-201.

[39] Chen, A.M. (1991) Non-compliance in community psychiatry: A review of clinical interventions. *Hospital and Community Psychiatry*, **42(3)**, 282-287.

[40] Voils, C.I., Steffens, D.C., Flint, E.P. and Bosworth, H.B. (2005) Social support and locus of control as predictors of adherence to antidepressant medication in an elderly population. *The American Journal of Geriatric Psychiatry*, **13(2)**, 157-165.

[41] Meehan, A.J. (1995) From conversion to coercion: The police role in medication compliance. *Psychiatric Quarterly*, **66(2),** 163-184.

[42] Perkins, D.M. (2002) Predictors of non-compliance in patients with schizophrenia. *Journal of Clinical Psychiatry*, **63(12)**, 1121-1181.

[43] (2008) Mental health and substance abuse unit. Annual Report, Ministry of Health, Jamaica (MOH), Kingston.

Appendix

Question	Particular
Ques 1-8	Demographic characteristics Gender; age at last birth day; educational level; marital status; Employment status; source of income; whom do you live with
Ques 9	Do they assist or support you in taking care of your conditions?
Ques 10	What is your diagnosis?
Ques 11	Do you know what medication(s) you are taking?
Ques 12	How many different types of medication do you take daily? One; Two; Three; Four; Five; Six or more
Ques 13	Are you on monthly injection for psychiatric illness?
Ques 14	How often are you to take your medication? Daily, less than three times per week; more than three times per week
Ques 15	I take my medication as prescribed by my doctor Always; most times; Sometimes; Rarely; Never
Ques 16	What are the reasons for not taking your medication as is prescribed?
Ques 17	Did you take the medication this morning? Yes or no
Ques 18	I know when I will become sick? Strongly agree; agree; unsure; disagree; strongly disagree
Ques 19	Medication should only be taken if you are experiencing a symptom of mental illness? Strongly agree; agree; unsure; disagree; strongly disagree
Ques 20	Were you told about the medication that you are taking? Yes or no
Ques 21	If yes (Ques 20), whom? Doctor; Nurses; Pharmacist; Other
Ques 22-30	Questions on medication availability; health care insurance coverage; dispensary of medication; waiting time for medication dispensary
Ques 31	Will counselling on how to take medications help me to take them better?

J. Biomedical Science and Engineering, 2010, 3, 612-617 **JBiSE**
doi:10.4236/jbise.2010.36083 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Performance comparison of neural network training methods based on wavelet packet transform for classification of five mental tasks

Vijay Khare[1], Jayashree Santhosh[2], Sneh Anand[3], Manvir Bhatia[4]

[1]Jaypee Institute of Information Technology, Department of Electronics and communication Engineering, New Delhi, India;
[2]Indian Institute of Technology, Computer Services Centre, New Delhi, India;
[3]Indian Institutes of Technology, Centre for Biomedical Engineering, New Delhi, India;
[4]Department of Sleep Medicine, New Delhi, India.
Email: vijay.khare@jiit.ac.in; jayashree@cc.iitd.ac.in; sneh@iitd.ernet.in; manvirbhatia1@yahoo.com

Received 11 February 2010; revised 15March 2010; accepted 25 March 2010.

ABSTRACT

In this study, performances comparison to discriminate five mental states of five artificial neural network (ANN) training methods were investigated. Wavelet Packet Transform (WPT) was used for feature extraction of the relevant frequency bands from raw electroencephalogram (EEG) signals. The five ANN training methods used were (a) Gradient Descent Back Propagation (b) Levenberg-Marquardt (c) Resilient Back Propagation (d) Conjugate Learning Gradient Back Propagation and (e) Gradient Descent Back Propagation with movementum.

Keywords: Electroencephalogram (EEG); Wavelet Packet Transform (WPT); Artificial Neural Network (ANN)

1. INTRODUCTION

Brain computer interface use a non muscular communication channel for conveying message and command to the external words in the absence of biological channels [1-3]. Neuromuscular disorders like Amyotrophic lateral sclerosis can temporarily or permanently impair spoken and physical communication. Those most severely affected may lose all voluntary muscle control and may be completely locked in to their bodies, unable to communicate in any way. Using cognitive abilities is sometimes the only the way to restore communication and motor function [4-7]. Through training, subjects can learn to control their brain activity in predefined fashion that is classified by pattern recognition algorithms. Accuracy of classification is affected by the quality of EEG signals and the processing algorithms. The processing algorithms include preprocessing, feature extraction and classification. Previous studies investigated the effect of different feature extraction algorithms with along and the different mental tasks on classification accuracy was investigated [8,9].

In this study, wavelet packet transform (WPT) method was used to capture the information of mental tasks from eight channel EEG signals of nine subjects. The coefficients of wavelet packet transform (WPT) were used as the best fitting input vector for ANN. Five various artificial neural networks (ANN) training methods were used to compare the performance in discrimination of five mental tasks.

2. METHODOLOGY

2.1. Subjects

Nine right-handed healthy male subjects of age (mean 23yr) having no sign of any motor- neuron diseases were selected for the study. A pro-forma was filled in with detail of their age & education level**.** The participants were student volunteers for their availability and interest in the study. EEG data was collected after taking written consent for participation. Full explanation of the experiment was provided to each of the participants.

2.2. EEG Data Acquisition

EEG Data used in this study was recorded on a Grass Telefactor EEG Twin3 Machine available at Deptt. of Neurology, Sir Ganga Ram Hospital, New Delhi. EEG recording was done for five mental tasks for five days, from nine selected subjects. Data was recorded for 10 sec during each task and each task was repeated five times per session per day. Bipolar and Referential EEG was recorded using eight standard positions C3, C4, P3, P4, O1 O2, and F3, F4 by placing gold electrodes on scalp, as per the international 10-20 standard system of electrode placement as shown in **Figure 1**. The settings

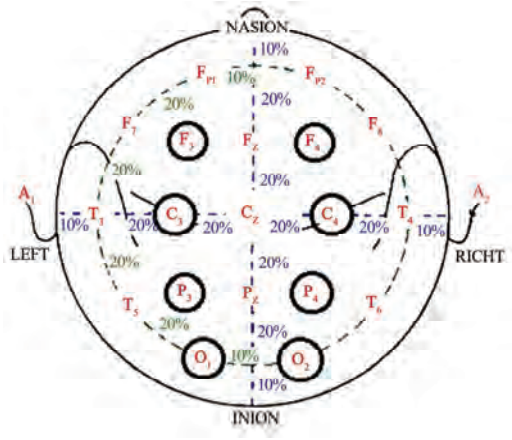


Figure 1. Montage for present study.

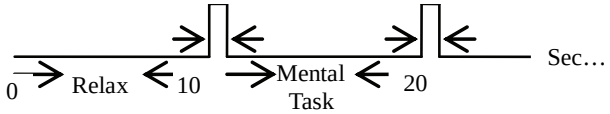


Figure 2. Timing of the Protocol.

used for data collection were: low pass filter 1Hz, high pass filter 35 Hz, sensitivity 150 micro volts/mm and sampling frequency fixed at 400 Hz. The reference electrodes were placed on ear lobes and ground electrode on the forehead. EOG (Electooculargram) being a noise artifact, was derived from two electrodes placed on outer canthus of left and right eye in order to detect and eliminate eye movement artifacts.

2.3. Experiment Paradigm

An experiment paradigm was designed for the study and the protocol was explained to each participant before the experiment. In this, the subject was asked to comfortably lie down in a relaxed position with eyes closed. After assuring the normal relaxed state by checking the status of alpha waves, the EEG was recorded for 50 sec, collecting five session of 10 sec epoch for the relaxed state. This was used as the baseline reference for further analysis of mental task. The subject was asked to perform a mental task on presentation of an audio cue. Five session of 10 sec epoch for each mental task were recorded as shown in **Figure 2**. The whole experimental lasted for about one hour including electrodes placement.

Data collected from nine subjects performing five mental tasks were analyzed. The following mental tasks were used:

Relaxed: The subject was asked to relax with their eyes closed. No mental or physical task to be performed at this stage.

Arithmetic Task: The subject was asked to perform simple arithmetic (SA) and complex arithmetic (CA). An example of a trivial calculation is to multiply 2 by 3 and nontrivial task is to multiply 49 by 78. The subject was instructed not to vocalize or make movements while solving the problem. EEG signal were recorded corresponding.

Geometric Figure Rotation (R): The subject was given 30 seconds to see complex three dimensional objects, after which the object was removed. The subject was instructed to visualize the object being rotated about an axis. The EEG signals were recorded during this period.

Movement Imagery (M): The subject was asked to plan movement of the right hand and corresponding EEG signals were recorded during this period.

2.4. Feature Extraction

The frequency spectrum of the signal was first analyzed through Fast Fourier Transform (FFT) method [10]. The FFT plot of signals from the all electrode pairs were observed and maximum average change in EEG amplitude was noted as shown in **Figures 3-7.**

For relaxed, the peaks of power spectrum almost coincide for central area in the alpha frequency range (8-13 Hz)[11]. EEG recorded with relaxed state is considered to be the base line for the subsequent analysis. Mu rhythms are generated over sensorimotor cortex during planning a movement. For movement imagery (M) of right hand, maximum up to 50% band power attenuation was observed in contralateral (C3 w.r.t C4) hemisphere in the alpha frequency range (8-13 Hz) [11,12]. For geometrical figure rotation(R), the peak of the power spectrum was increased in right hemisphere rather than left in the occipital area for the alpha frequency range (8-13 Hz) [13]. For simple arithmetic(SA), the peak of the power spectrum was increased in left hemisphere rather than right hemisphere in the frontal area for the alpha frequency range (8-13 Hz) [14].For complex arithmetic (CA), the peak of the power spectrum was increased in left hemisphere rather than right hemisphere in the parietal area for the alpha frequency range (8-13 Hz).

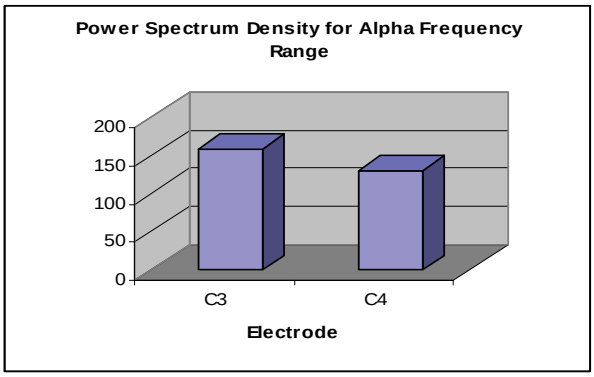


Figure 3. Power Spectra for Relax state at C3 and C4 channel.

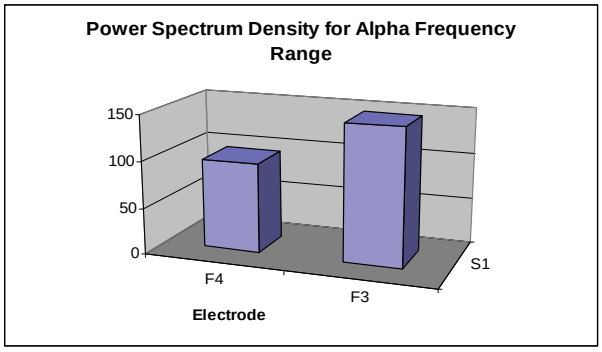


Figure 4. Power Spectra for simple arithmetic at F3 and F4 channel.

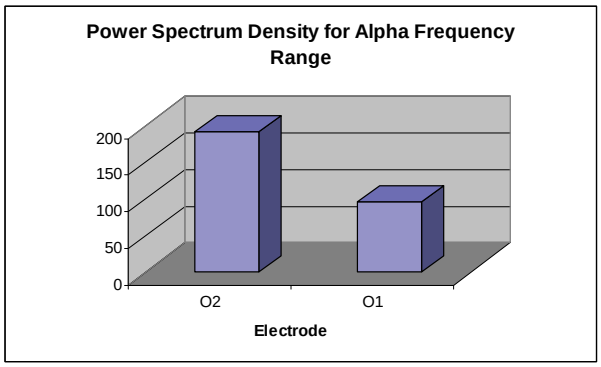


Figure 5. Power Spectra for rotation at O1 and O2 channel.

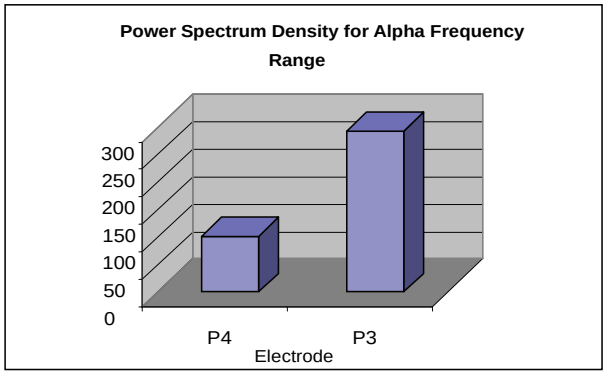


Figure 6. Power Spectra for complex Arithmetic P3 and P4 channel.

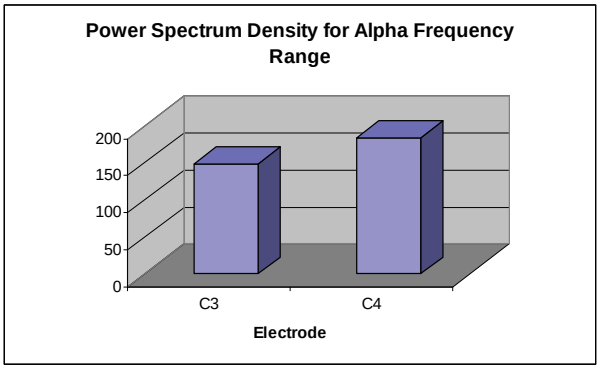


Figure 7. Power Spectra for right movement at C3 and C4 channel.

2.5. Wavelet Packet Transform

By applying Wavelet packet analysis on the original signal wavelet coefficients in the 8-13 HZ frequency band at the 5^{th} level node (5, 3) were obtained. The signal was reconstructed at node (5, 3). These coefficients are scaled and WPT coefficients are used as the best fitting input vector for ANN. Wavelet transform we were able to reduce 1 second of EEG data to 21 coefficients for each mental tasks [15,16].

2.6. Classification

The main advantage of choosing artificial neural network for classification was due to fact that ANN's could be used to solve problems, where description for the data is not computable. ANN could be trained using data to discriminate the feature. The five different training methods used for Classification in the present study were Gradient Descent method Resilient Back propagation, Levenberg-Marquardt, Conjugate Gradient Descent and Gradient Descent back propagation with movementum.

For classification a two layer neural networks was used for the instance a topology of {10, 1} indicate a 21 input, 10 neurons in hidden layer and one output layer. The neural network was designed to accept a 21 element input vector and give a single output. One second segments of EEG were classified. The training set was composed of 60% EEG trial per mental tasks and test set was composed 40% EEG trial per mental task. The output was designed to give 0 for baseline and 1 for task. The neural network was trained for a fixed number of epochs and the training is done using a five learning techniques [17]. Parameter used for five training methods of neural network for classification of five mental tasks as shown in the **Table 1.**

2.7. Evaluations

Performance (R_C) is defined as ratio between correctly classified patterns in the test set to the total number of patterns in the test set in percentage.

$$Rc = \frac{\text{Number of correctly classified test patterns}}{\text{Total number of patterns in the test set}}$$

With the help of above formula we calculate the performance of each method for each task [18].

3. RESULTS

For the classification of five mental tasks neural network training methods were used. Nine male right-handed subjects participated in the experiments. The subjects asked to perform five mental tasks namely relaxed, arithmetic task (simple and complex multiplication), geometrical figure rotation, and movement imagery. **Table 2** showed the comparison of the performance of five neu-

Table1. Parameter used for different back propagation algorithms with topology {10, 1}.

Gradient Descent with Momentum (GDM)	
Topology {10,1}	α = 0.01. Mu = 0.01
MSE = 1e-5	
Epoach = 5000	
Gradient Descent method (GD)	
Topology {10,1}	α = 0.01
MSE = 1exp-(5)	
Epoach = 5000	
Resilient Back propagation (RBP)	
Topology {10,1}	α = 0.01
MSE = 1exp-(5)	
Epoach = 5000	β = 0.75 and β_1 = 1.05
Conjugate Gradient descent (CG)	
Topology {10,1}	α = 0.01
MSE = 1exp-(5)	
Epoach = 5000	
Levenberg-Marquardt (LM)	
Topology {10,1}	Mu = 0.01
MSE = 1exp-(5)	
Epoach = 5000	Mu_dec = 0.1 and Mu_inc = 10

ral network (NN) training methods in classifying of five mental tasks. From this table we can say that Resilient Back Propagation method was most suitable for the classification of all five mental tasks. Because this method gave highest performance (95%) all five mental tasks. Percentage average accuracy shown in the **Figures 8-12**.

4. CONCLUSIONS

In this study, nine healthy male subjects were selected to investigate MLP classifier with various training methods to discriminate five mental tasks (relaxed state, movement imagery of right hand, geometrical figure rotation, arithmetic simple task, and arithmetic complex task) effectively. For relaxed, the peaks of power spectrum almost coincide in central area at a particular based frequency. For simple arithmetic, it was observed that the amplitude of the power spectrum for alpha frequency range (8-13 Hz) increased left hemisphere rather than right hemisphere in frontal region. For complex arithematic, it was observed that the amplitude of the power spectrum for alpha frequency range (8-13 Hz) increased

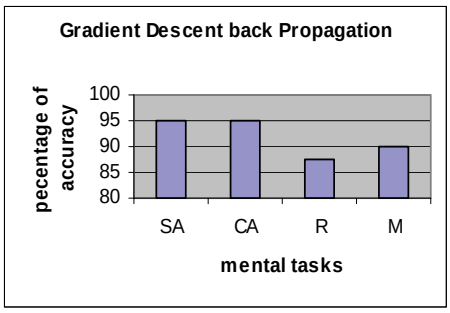


Figure 8. Classification accuracy using GDA BP training methods.

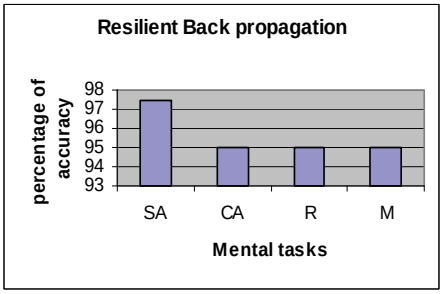


Figure 9. Classification accuracy RBP training methods.

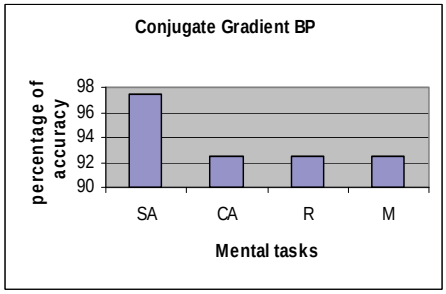


Figure 10. Classification accuracy CGBP training methods.

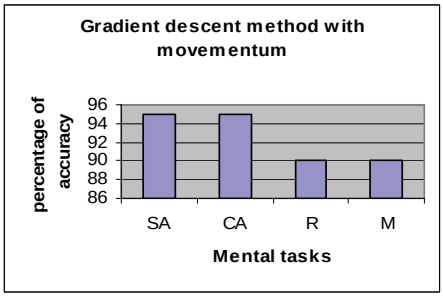


Figure 11. Classification accuracy GDM training methods.

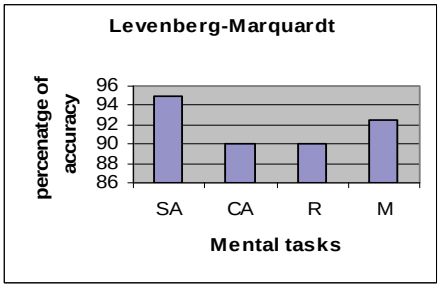


Figure 12. Classification accuracy LM training methods.

Table 2. Comparisons of Different NN training Methods.

Tasks	Baseline and Arithmetic simple		Baseline and Arithmetic complex		Baseline and Rotation		Baseline and movement	
Techniques	Correct classification	Wrong classification	Correct classification	Wrong classification	Correct classification	Wrong classification	Correct classification	Wrong classification
Gradient Descent Back Propagation	95%	5%	95%	5%	87.5%	12.5%	90%	10%
Resilient Back Propagation	97.5%	2. 5%	95%	5%	95%	5%	95%	5%
Conjugated Gradient BP	97.5%	2.5%	92.5%	7.5%	92.5%	7.5%	92.5%	7.5%
GD BP with Momentum	95%	5%	95%	5%	90%	10%	90%	10%
Levenberg-Marquardt	95%	5%	90%	10%	90%	10%	92.5%	7.5%

left hemisphere rather than right hemisphere in parietal region. For geometrical figure rotation, the peak of the power spectrum in the alpha frequency range (8-13 Hz) increased right occipital area. For movement imagery, the peak of the power spectrum in the alpha frequency range (8-13 Hz) had an attenuation central area. The result showed the performance of neural network with various training method, for classifying of mental tasks w.r.t baseline. Resilient backpropagation training method has best performance among all the training methods for classification of mental tasks w.r.t baseline.

The authors would like to extend the work with severely disabled people and to customize the device as per individual response and requirements. This kind of system can also be used in a variety of applications like– Environment control units (ECU'S), helping disable people to directly interact with hand held devices such as cell phones and PDAs.

5. ACKNOWLEDGEMENTS

The authors would like to acknowledge their gratitude to the Centre of Biomedical Engineering of IIT New Delhi and Electronics and communication department of JIIT Noida. The authors also thank the scientific and technical staff of EEG Laboratory of Sir Ganga Ram hospital, New Delhi for the help in carrying out the experiment.

REFERENCES

[1] Wolpaw, J.R., Birbaumer, N., Mc Farland, D.J., Plurtscheller, G. and Vaughan, T.M. (2002) Brain computer interfaces for communication and control. *Clinical Neurophysiology*, **113(6)**, 767-791.

[2] Pfurtschelle, G., Flotzinger, D. and Kalcher, J. (1993) Brain computer interface—A new communication device for handicapped people. *Journal of Microcomputer Applications*, **16**, 293-299.

[3] Wolpaw, J.R., Vaughan, T.M. and Donchin, E. (1996) EEG based communication prospects and problems. *IEEE Transactions on Rehabilitation Engineering*, **4**, 425-430.

[4] Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F. and Arnaldi, B. (2007) A review of classification algorithms for EEG bases brain computer interface. *Journal of neural Engineering*, **4**, 1-13.

[5] Wolpaw, J.R., Leob, G.E., Allison, B.Z., Donchin, E. and Turner, J.N. (2006) BCI Meeting 2005—Wokshop on signals and rerecording methods. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **14**, 138-141.

[6] Anderson, C.W., Stolz, E.A. and Shamsunder, S. (1998) Multivariable autoregressive model for classification of spontaneous electroencephalogram during mental tasks. *IEEE Transactions on Biomedical Engineering*, **45**, 277-278.

[7] Boostani, R., Graimann, B., Moradi, M.H. and Plurfscheller, G. (2007) Comparison approach toward finding the best feature and classifier in cue BCI. *Journal of Medical and Biological Engineering and Computing*, **45**, 403-413.

[8] Palaniappan, R. (2006) Utilizing gamma band to improve mental task based brain—Computer interface design. *IEEE Transactions on Neural and Rehabilitation Systems Engineering*, **14**, 299-303.

[9] Keirn, Z.A. and Aunon, J.I. (1990) A new mode of communication between man and his Surroundings. *IEEE Transactions on Biomedical Engineering*, **37**, 1209-1214.

[10] Palaniappan, R. (2005) Brain computer interface design using band powers extracted during mental task. *Proceeding of the* 2*nd International IEEE EMBS Conference on Neural Engineering*, Arlington, 321-324.

[11] Pfurtscheller, G., Neuper, C., Schlogl, A. and Lugger, K. (1998) Separability of EEG signals recorded during right and left motor imagery using adaptive auto regressive parameters. *IEEE Transactions on rehabilitation Engineering*, **6**, 316-325.

[12] Santhosh, J., Bhatia, M., Sahu, S. and Anand, S. (2004) Quantitative EEG analysis for assessment to plan a task in ALS patients, a study of executive function (planning) in ALS. *Journal of Cognitive Brain Research*, **22**, 59-66.

[13] Nikolaev, A.R. and Anokhin, A.P. (1998) EEG frequency ranges during reception and mental rotation of two and three dimensional objects. *Neuroscience and Bheaviour physiology*, **29**, 203-223.

[14] Osaka, M. (1984) Peak alpha frequency of EEG during a mental task: Task difficulty and hemisphere difference. *Journal of Psychophysiology*, **21**, 101-105.

[15] Ting, W., Zhenga, Y.G., Bang-huaa, Y. and Hong, S. (2008) EEG feature extraction based on wavelet packet

decom-position for brain computer interface. *Measurement*, Elsevier Journal, **41**, 618-625.

[16] Akay, M. (1995) Wavelet in biomedical engineering. *Journal of Annals of Biomedical Engineering*, **23**, 529-530.

[17] Ravi, K.V.R. and Palaniappan, R. (2006) Neural Network Classification of Late gamma band electroenceph-alogram features, *Soft Computing*, **10**, 163-169.

[18] Cheng, M., Gao, X., Gao, S. and Xu, D. (2002) Design and implementation of a brain computer interface with high transfer rates. *IEEE Transactions on Biomedical Engineering*, **49**, 1181-1186.

J. Biomedical Science and Engineering, 2010, 3, 618-624 **JBiSE**
doi:10.4236/jbise.2010.36084 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Automated neurosurgical video segmentation and retrieval system

Engin Mendi[1], Songul Cecen[2], Emre Ermisoglu[2], Coskun Bayrak[2]

[1]Department of Applied Science, University of Arkansas at Little Rock, Little Rock, USA;
[2]Department of Computer Science, University of Arkansas at Little Rock, Little Rock, USA.
Email: esmendi@ualr.edu; sxcecen@ualr.edu; exermisoglu@ualr.edu; cxbayrak@ualr.edu

Received 18 January 2010; revised 25 February 2010; accepted 6 March 2010.

ABSTRACT

Medical video repositories play important roles for many health-related issues such as medical imaging, medical research and education, medical diagnostics and training of medical professionals. Due to the increasing availability of the digital video data, indexing, annotating and the retrieval of the information are crucial. Since performing these processes are both computationally expensive and time consuming, automated systems are needed. In this paper, we present a medical video segmentation and retrieval research initiative. We describe the key components of the system including video segmentation engine, image retrieval engine and image quality assessment module. The aim of this research is to provide an online tool for indexing, browsing and retrieving the neurosurgical videotapes. This tool will allow people to retrieve the specific information in a long video tape they are interested in instead of looking through the entire content.

Keywords: Video Processing; Video Summarization; Video Segmentation; Image Retrieval; Image Quality Assessment

1. INTRODUCTION

Developing countries suffer from lack of access to the medical expertise. Due to the inadequacy of trained medical professionals, the health maintenance system of the country may face variety of problems which will directly affect individual's quality of life and also entire well-being of society. Limitations in accessing to the medical expertise may also exist in small regional hospitals & health-care centers in rural places in developed countries. Therefore, connecting as many hospitals as possible to a medical information system from regional level to the state and national levels and ultimately to the global level, which is simply illustrated in **Figure 1**, would be very beneficial in terms of improved standard of medical practice and educational aspects for medical students and staff who can not reach to the medical resources, due to resource, geographical, and time constraints.

Our research is to show not only the importance of the accommodation of the massive amount of data for educational use but also the preservation of a life long experience of pioneers in the field of neurosurgery for further use via automatically defining a logical structure of the video content. Since only a few fortunate ones get a chance to be with these experts to see how they perform such complex operations. Therefore, with this service the boundaries can be expended so that [1].

- The educational needs of residents can be complemented by allowing access in a timely efficient manner.
- The educational needs of medical students can be provided by allowing access.
- The knowledge enhancement needs of Neurosurgeons around the world with special benefits to developing countries can be supported by allowing access.
- The help in teaching of operating room (OR) nurses and physician assistants can be provided.
- The foundation for future research related to simulation technology can be constructed.

The goal of this system is summarized in **Figure 2**. The system has 3 main components which are video segmentation engine, image retrieval engine and image quality assessment module.

2. VIDEO SEGMENTATION ENGINE

Medical video libraries are dedicated to many health-related applications such as medical imaging, medical research and education, medical diagnostics and training of medical professionals. Due to the rapid development in production, storage and distribution of multimedia content, the video data of these medical repositories can be directly transmitted to the people via internet. However, due to the huge size of the videos, very large bandwidth will be required. Additionally, it will be very difficult reaching a certain portion of the video. For instance, when a medical surgeon or student wants to look through a spe-

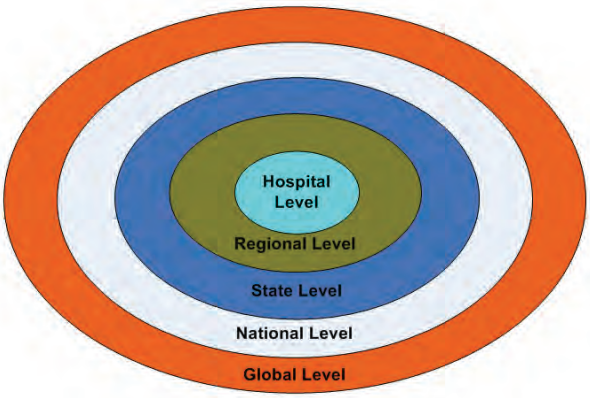


Figure 1. Work-flow of the system.

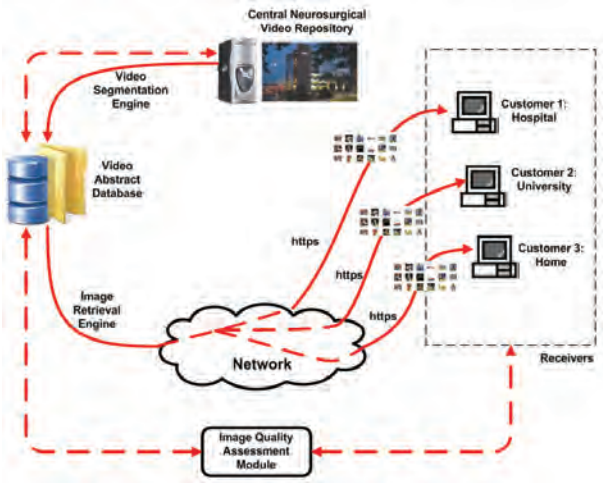


Figure 2. System architecture.

cific part of a 15-hour neurosurgery videotape, they will have to browse the entire content of the video in order to find the right part they want to see. Our video segmentation engine generates a concise summary of the semantics in the neurosurgical videotape to help them browse and search the large amount of video data. The architecture of the engine is depicted in **Figure 3**. As shown in **Figure 3**, an MPEG video source comprises a group of video shots, and a video shot is an unbroken sequence of frames captured from one perspective. The engine partitions a video sequence into a set of shots, and some key frames are extracted to represent each shot. Finally key frames are collected in the video abstract database.

Video segmentation is the central process for automatic video indexing, browsing and retrieval systems. It aims partitioning a video sequence into meaningful segments and extracting a sequence of key frames, that each key represents the content of corresponding video segment. The video sequence is divided into meaningful segments (shots) which are the basic elements of indexing, and then each shot is represented by key frames. These frames are indexed by extracting spatial and temporal features.

Video segmentation includes two major steps: 1) Shot boundary detection and 2) Key frame extraction. Shot boundary detection targets breaking up the video into meaningful sub-segments. Key frame extraction involves selecting one or multiple frames that will represent the content of each shot.

In our work, we segmented the shots by detecting the boundaries via color histogram differences and self-similarity analysis. In color histogram differences, RGB color space is converted to HSV space, and then color-quantization is applied to HSV color space. Finally the differences of HSV histograms between consecutive frames are computed to determine the peaks representing the shot boundaries [1,2]. In self-similarity analysis, HSV feature vectors of the frames within the video data are visualized with a two-dimensional matrix by applying a similarity metric [2].

Key frame extraction is the second major process of video segmentation. We used four approaches in order to select the key frames:

1) The first is the traditional k-means clustering, determining video summaries with a specific number of frames, which will represent the entire video content. The number of the key frames is specified by the user. The frames closest to the cluster centroids are selected as key frames [2].

2) The second is the dominant set clustering that automatically decide the number of key frames according to the similarity of the data without any initial decision of cluster number. The clustering is based on dominant sets which are the representation of an edge-weighted graph as a similarity matrix [2-4].

3) The third is based on salient region detection and structural similarity. Saliency maps representing the attended regions are produced from the color and luminance features of the video frames. Introducing a novel signal fidelity measurement-saliency based structural similarity index, the similarity of the maps is measured.

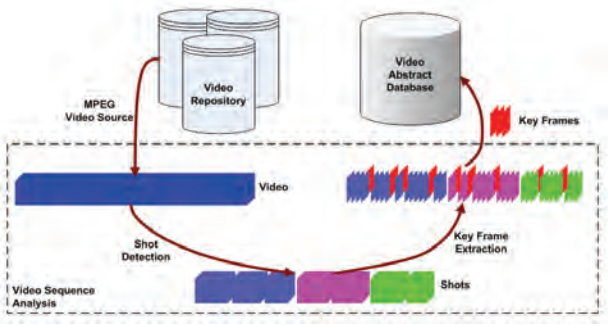


Figure 3. The architecture of the video segmentation engine.

Based on the similarities, shot boundaries and key frames are determined [5].

4) The fourth approach, called "index-based retrieval (i-Base)", is based on Discrete Cosine Transform (DCT) and Self Organizing Map (SOM). It allows user to quickly find a particular point within a certain domain and/or determine if the domain is relevant to the need. i-Base forms a hierarchy from the uniquely represented shots by using frames. User then can map only the relevant section in the source with the request issued. The idea of 'just request-to-response mapping' prevents not only the unwanted information retrieval but also saves time and bandwidth [1].

Figure 4 shows the user interfaces of the video segmentation system, allowing the user to set the parameters. Currently, we are in the process of transferring our tool over the WEB environment. A sample of key frame set of a neurosurgical training video data, presented to the user is depicted in **Figure 5**.

3. IMAGE RETRIEVAL ENGINE

Currently, all web-browser based image search engines are based on textual data, that images are associated by annotations and then searched using keywords. In terms of medical images, the most of the medical image information is not accessible or limited to one or two databases that can be searched with key words. As medical images can not be fully described by textual information, keywords are not sufficient enough to retrieve relevant medical images from large databases. For instance, for the "lung CT" key term, Google is able to retrieve only 130 image results. However, only Health Education As sets Library (HEAL, http://www.healcentral.org/) contains

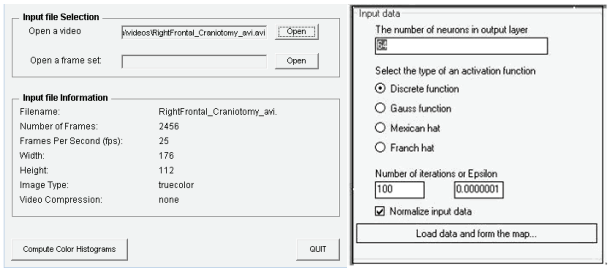


Figure 4. User interface of the tool.

more than 1,000 lung CT images [6].

Our image retrieval engine delivers the image results from our video abstract database, taking advantage of visual features of the images. **Figure 6** shows the architecture of our image retrieval engine. Several image features representing the visual content of the images in video abstract database and query image are extracted. The images in video abstract database are the key frames of the neurosurgical videotapes previously extracted by the video segmentation engine.

Based on the similarity metric, how close query image and key frames are measured. Retrieval results are then ranked according to the similarity score and delivered to make available to receivers over broadband network.

Content-based image retrieval (CBIR) is a technique using visual descriptors to search images from image databases according to users' needs. It aims effectively searching and browsing of large image digital libraries based on automatically extracted image features. In a typical CBIR system, features of every image in the database have been extracted and then compared with the query image. We have conducted a preliminary evalu-

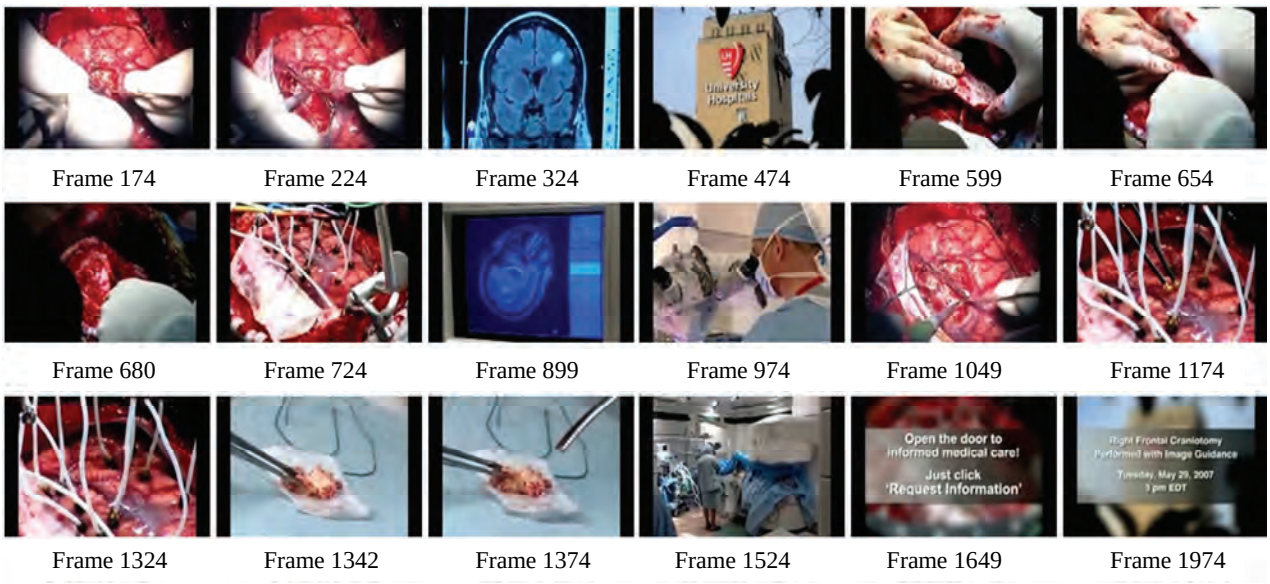


Figure 5. The 18 key frames of a neurosurgical (right frontal craniotomy) video sequence of 2500 frames presented to the user.

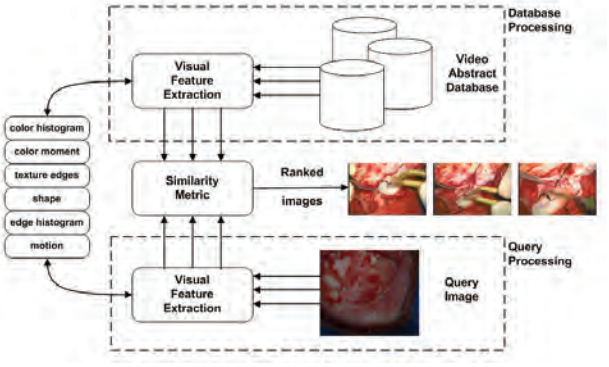


Figure 6. The architecture of the image retrieval engine.

ation on the precision performance of following two approaches:

1) The first is comparing images using color histograms. Color is one of the most important image features for CBIR. A color histogram is the representation of frequency distribution of color bins in an image. Color histograms are widely used in comparison of images, since they are robust to change in translation, rotation and angle of view. We have used two different color spaces, RGB and HSV. We quantized the RGB color space as well as HSV space to reduce the number of bins, using 256 colors (16 levels for each R, G and B channels in RGB space;16 levels for H channel, 4 levels for S channel and 4 levels for V channel in HSV space). Finally, to evaluate the similarity between query image and the image in the video abstract database, we have computed the Euclidean distance between corresponding color histograms [7].

2) The second is comparing images using two image fidelity measurements. We have used mean squared error (MSE) and structural similarity (SSIM) index [8] to quantify the similarity of images [7]. MSE compares two images on a pixel-by-pixel basis, whereas SSIM considers structural information.

We have conducted a preliminary evaluation on the precision performance of above four approaches. We have used a subset of COREL Image Database [9,10] which is available at http://wang.ist.psu.edu/docs/related.shtml. The database contains 10 image classes with 100 images each (1000 images in total). The classes are: Africa, Beach, Buildings, Buses, Dinosaurs, Flowers, Elephants, Horses, Food and Mountains.

We measured the retrieval effectiveness on the precision performance of each approach. The detailed precision results are shown in **Table 1**. Average precisions are computed by taking every image in a class as query image. As can be seen, the precision performances of the algorithms change with different classes. According to the overall results, HSV histogram is the most effective approach among others.

4. IMAGE QUALITY ASSESSMENT MODULE

Image quality assessment is an important part of content delivery over networks since network conditions vary for individual users and also digital images are subject to a wide variety of distortions during processing, storage and transmission, any of which may result in a degradation of visual quality [8,11]. Therefore quantifying the image quality degradation occurring in a system would be very beneficial, so that the quality of the images and videos produced can be controlled and adjusted. For instance, a system can examine the quality of videos and images being transmitted in order to control and allocate streaming or downloading resources. Moreover, a quality assessment module can assist in the optimal design of pre-filtering and bit assignment algorithms at the encoder and of optimal reconstruction, error concealment, and post-filtering algorithms at the decoder [8].

On the other hand, according to a recent research of Microsoft [12]; due to the difficulty of the image quality assessment problem, current web-browser based image search engines lack of user requirements, because there is no effective and practical solution to allow an understanding of image content, which fits the user needs. Image quality assessment research would greatly help improve users' browsing experiences.

Therefore we have been designing a quality assessment module automatically predicting perceived image quality. This problem is more competitive in medical imagery, because medical imagery may play crucial role in many health-related issues such as diagnostic design, patient-care, training and education of medical professionals and students. The framework of the module is depicted in **Figure 7**. We have already developed a novel objective image quality metric which is superior to the existing metrics in the literature. We have also validated our metric against a large set of subjective ratings gathered

Table 1. Precision of 10 image categories for top 30 matches.

	Histogram-based		*Similarity-based*	
Image Class	*RGB*	*HSV*	*SSIM*	*MSE*
African People	23.03	40.57	8.33	0.33
Beach	13.4	22.63	33	33.33
Buildings	18.6	24.73	14.33	3.67
Buses	31	47.03	25.33	3.67
Dinosaurs	34.83	44.6	95.33	97.67
Elephants	18.8	19.57	25.33	35
Flowers	33.3	32.7	67	90
Horses	16.83	81.93	21.33	43.67
Mountains and glaciers	13.9	38	19	31
Food	55.03	31.3	10.33	1.67
AVERAGE	25.87	38.31	31.93	34.00

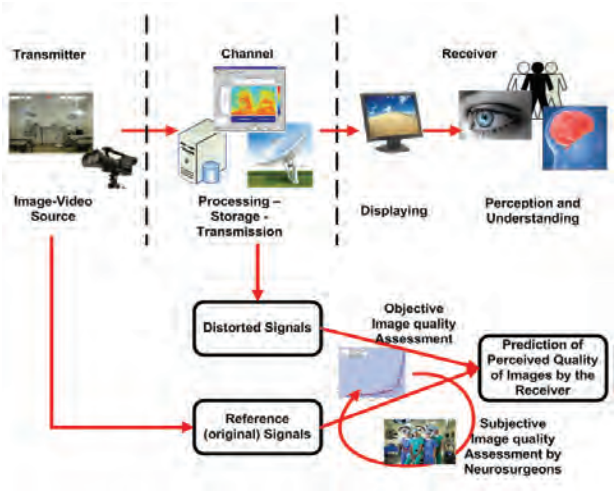


Figure 7. The framework of the image quality assessment module.

for a public image database. Currently we have been working on a subjective image quality assessment for neurosurgery imagery. This assessment will be based on expert opinions of a group of neurosurgeons from UAMS, by determining fixation points on the images while tracking their eye movements.

Image quality assessment has a great importance in several image and video processing applications such as filter design, image compression, restoration, denoising, reconstruction, and classification. The goal of image quality assessment is predicting image quality of display output perceived by the end user. Multimedia contents are subjected to the variety of artifacts during acquisition, processing, storage and delivering, which may lead to reductions in the quality. Our image quality assessment module dynamically monitor and adjust the image quality, so that the output quality of the image or video presented to the user can be maximized for available resources such as network conditions and bandwidth requirements.

Image quality metrics can be classified into 2 categories: Subjective and objective metrics. The most reliable way to measure of image quality is to look at it because human eyes are the ultimate viewer and images are evaluated by humans. Subjective evaluation by orienting on human visual system is determined by Mean Opinion Score (MOS) which relies on human perception. On the other hand, objective metrics are also very valuable to

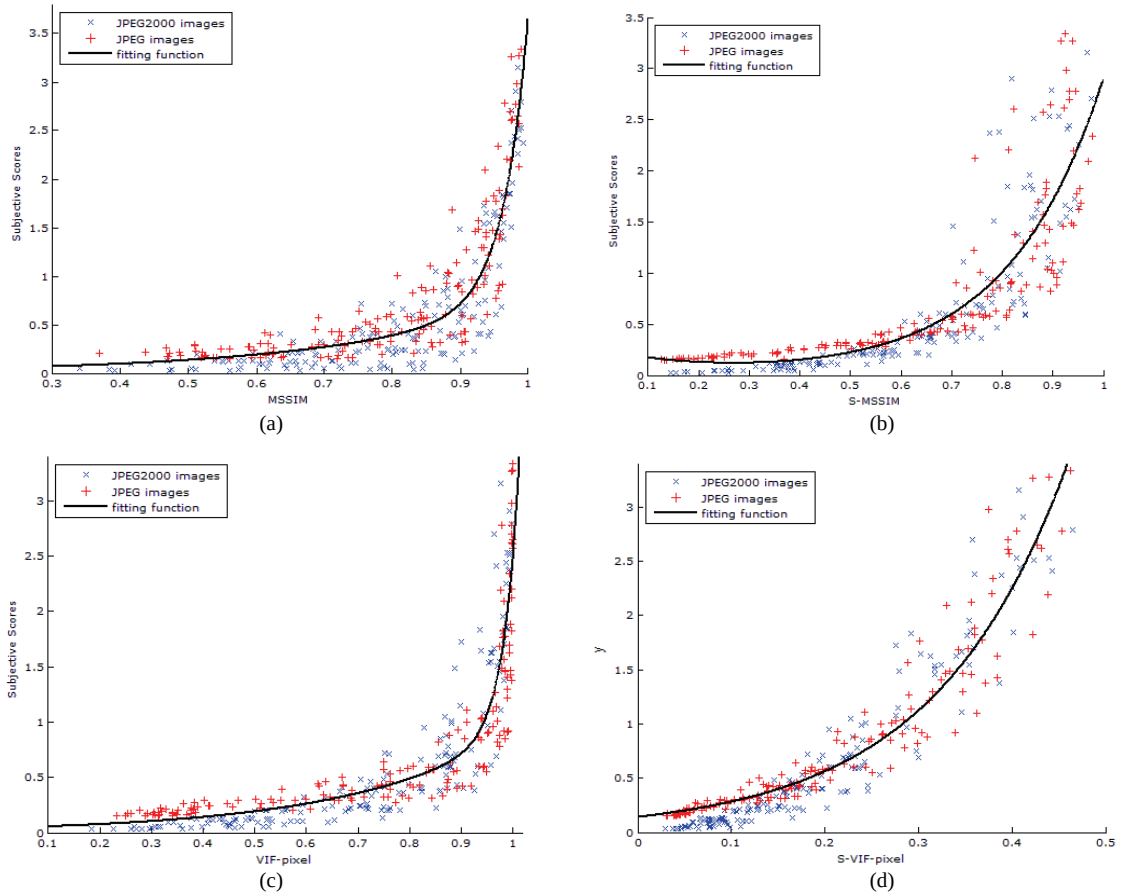


Figure 8. Scatter plots of subjective/objective scores on LIVE Database. Red points (+) and blue points (x) denote JPEG and JPEG2000 images, respectively. (a) SSIM; (b) S-SSIM; (c) VIF in pixel domain; (d) S-VIF in pixel domain.

predict perceived image quality. They are based on mathematical models that approximate results of subjective quality assessment. Amongst the objective quality metrics, full reference metrics require complete availability of original non-distorted reference image which will be compared with the corresponding distorted image, while reduced reference and no reference metrics require limited and no availability of this, respectively.

We developed a new image quality metrics, S-SSIM (saliency-based structural similarity index) and S-VIF (saliency-based visual information fidelity), based on frequency-tuned salient region detection introduced by [13]. Saliency maps are produced from the color and luminance features of the image. SSIM [8] index and visual information fidelity (VIF) in pixel domain [14] are modified by the weighting factors of the saliency maps.

We validated our approach using LIVE Image Database [15] as test bed. The database contains 29 original images and 460 distorted images (227 JPEG2000 images and 233 JPEG images) with subjective scores for each image. Non-linear regression analysis has been performed to fit the data. The Pearson correlation coefficient is used to measure the association between subjective and objective scores. Our results showed that our technique is more correlated with human subjective perception.

Figure 8 shows the results for the database. Each sample point represents the subjective/objective scores of one test image. The y axis in the figure denotes the subjective scores in the database. The x axis denotes the predicted quality of images after a nonlinear regression toward above 4 objective scores, which are SSIM, S-SSIM, VIF in pixel domain and S-VIF in pixel domain, respectively. The Pearson validation scores between assessment metrics are depicted in **Table 2** [16].

The Pearson correlation coefficient varying from –1 to 1 is widely used to measure the association between two variables. High absolute values mean that the two variables being evaluated have high correlation. As shown in **Table 2**, our metric is more correlated with human subjective perception.

5. CONCLUSIONS

We presented a medical video segmentation and retrieval research initiative. We introduced the key components of the framework including video segmentation engine, image retrieval engine and image quality assessment module. We are currently in the process of transferring our framework and software tool over the WEB environment. This will allow people to access the specific information that they are interested in among entire video. Multimedia information system, digital library, and movie industry are some of the applications work on videos. Since they are widely used, it brings out the need of processing and saving the digital video. These processes are mainly the compressing, segmenting, and indexing of the video. The neurosurgical data which is initially compressed will pass through the segmentation and indexing. Then receiver will be able to retrieve the specific section of the video that he/she is interested in with maximum quality for the available network, bandwidth and hardware resources. The overall objective is to provide convenience and easiness in accessing the relevant data without going over the whole data.

Table 2. Pearson correlation coefficients.

	SSIM	S-SSIM	VIF-pixel	S-VIF-pixel
LIVE Image Database	0.6823	0.7475	0.7126	0.9083

REFERENCES

[1] Cecen, S. (2009) Histogram based video segmentation and key frame extraction on SOM and DFT. Master's Thesis, University of Arkansas, Little Rock.

[2] Mendi, E. and Bayrak, C. (2010) Shot boundary detection and key frame extraction from video sequences. *Elsevier Information Sciences*, 2010.

[3] Pavan, M. and Pelillo, M. (2003) Dominant sets and hiera-rchical clustering. *Proceedings of the* 9*th European Conference on Computer Vision*, 362-369.

[4] Pavan, M. and Pelillo, M. (2005) Efficient out-of-sample extension of dominant-set clusters. *Advances in Neural Information Processing Systems*, **17**, 1057-1064.

[5] Mendi, E. and Bayrak, C. (2010) Shot boundary detection and key frame extraction using salient region detection and structural similarity. *The* 48*th ACM Southeast Conference*, Oxford, Mississippi, 15-17 April, 2010.

[6] Lehmann, T. M., Mller, H., Tian, Q., Galatsanos, N.P. and Mlynek, D. (2005) Augmented medical image management for integrated healthcare solutions.

[7] Mendi, E. and Bayrak, C. (2010) Performance analysis of color image retrieval. *The* 3*rd International Congress on Image and Signal Processing* (*CISP*'10), Yantai, 2010.

[8] Wang, Z. Bovik, A.C. Sheikh, H.R. and Simoncelli, E.P. (2004) Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, **13(4)**, 600-612.

[9] Li J. and Wang, J.Z. (2003) Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25(9)**, 1075-1088.

[10] Wang J.Z., Li, J. and Wiederhold G. (2001) SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23(9)**, 947-963.

[11] Chono, K., Lin, Y.-C., Varodayan, D., Miyamoto, Y. and Girod, B. (2008) Reduced-reference image quality assessment using distributed source coding. *IEEE International Conference on Multimedia and Expo*, 2008.

[12] Zhang, L., Chen, L., Jing, F. and Ma, W.-Y. (2006) Enjoy photo a vertical image search engine for enjoying high-quality photos. *The* 14*th ACM International Conference on Multimedia*, ACM Press, Santa Barbara.

[13] Achanta, R., Hemami, S., Estrada, F. and Süsstrunk, S. (2009) Frequency-tuned salient region detection. *IEEE International Conference on Computer Vision and Pattern Recognition* (*CVPR*), Miami.

[14] Sheikh, H.R. and Bovik, A.C. (2006) Image information and visual quality. *IEEE Transactions on Image Processing*, **15(2)**, 430-444.

[15] Sheikh, H.R., Wang, Z., Cormack, L. and Bovik, A.C. (2005) *Live Image Quality Assessment Database Release* 2. http:// live.ece.utexas.edu/research/quality.

[16] Mendi, E. and Milanova, M. (2010) Image quality assessment based on salient region detection. *Journal of Visual Communication and Image Representation*, Elsevier Ltd., 2010.

J. Biomedical Science and Engineering, 2010, 3, 625-632 **JBiSE**
doi:10.4236/jbise.2010.36085 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Mesoscopic relaxation time of dynamic image correlation spectroscopy

Kyongok Kang

Forschungszentrum Juelich, Weiche Materie (Soft Matter), Juelich, Germany.
Email: k.kang@fz-juelich.de

Received 16 February 2010; revised 20 March 2010; accepted 28 March 2010.

ABSTRACT

Dynamical images contain useful information of how the objects behave in time and space. When the system is in biological fluids, the motion of the object is much over-damped; the relaxation time is the characteristics in a diffusive time scale. We have found dynamical states of melting and forming of small nematic domains (10—30 μm) that are exhibited in the suspensions of fd-viruses under applied AC electric field amplitude at low frequency. Dynamic image correlation function is used for extracting the mesoscopic relaxation times of the dynamical states, which can be employed as an application to other dynamic imaging process of biologically relevant soft condensed matter and biomedical systems.

Keywords: Relaxation Time; Image Correlation Spectroscopy; Dynamical states

1. INTRODUCTION

Dynamic contrast-enhanced MRI (magnetic resonance imaging) is widely used for early detection of breast tumors, where the kinetics of signal variation are monitored after an injection of a paramagnetic contrast medium to distinguish between benign and malignant breast lesions [1]. The studies of microvasculature and microcirculation are useful in the sense that the solid tumor objects can disturb the regional blood flow, and influences the lack of linkage in a basement membrane. It is still a challenge to generalize the estimation of a regional blood flow and capillary permeability as volume fraction of the intravascular flow in breast tumors [1]. According to the authors in [1], the important criterions of reliable signal contrast (or relaxation times) are dependent on two assumptions; one is a negligible relaxavity of protons in different tissue compartment, and the other is much faster diffusion of water within a tissue and the exchange rate between them. The relaxation of bulk magnetization can be found and described by a single relaxation time [2]. These assumptions are feasible for a typical diffusion coefficient for water in soft tissues, and its relevant mean-squared displacement is reported as typical value of 30-60 μm size within about 125 ms recovery time [2]. Depending on the biomedical system, for instance, gray and white matter in the brain, heart, and kidney are in a relatively slow relaxation process (relaxation time as about 1/s) within tissue and blood stream, as compared to other systems as liver and spleen [2]. Thus as far as the contrast agent is efficiently fast enough, then the slow process of water permeability can be negligible in terms of the exchange rates between different compartments for early detection of malignant tumor cells. Thus for an intermediate and slow time process, the single exponent time constant is convenient to characterize most biological relevant experimental parameters.

When the time-evolutionary information is engaged with a local space, it is somewhat complicated to distinguish between the time, and space-variable. For the system that contains self-mimicking sizes or shapes of objects (or images), but with different temporal events, it is extremely useful to predict what will happen after a certain time steps of events. Relatively large-time scale, such as astronomical or geographical-time, and macroscopic length are less relevant to predict “probable” events in mesoscopic time scale. Most physiochemical environments, *i.e.*, medical, biochemical and biophysical issues are more engaged with a shorter-time interval. In the latter case, the system of interested objects is governed by thermodynamics and hydrodynamics, commonly in between the microscopic and macroscopic approach. As far as in time-correlation aspect, this mesos-copic time scale can be influenced by globally metastable but partially (or locally) in non-equilibrium states. Thus time-correlation is still a challenge in mesoscopic length and time scale and in the speculation, nevertheless which is ambient in most soft matters. The demanding task of predicting “forward” temporal events relies on the complexity of dynamical entities (or inhomogenieties) themselves that evolve in the physiological situations (or en-

vironments). Based on an appropriate statistical sampling, one can identify the quantitatively possible "observable" physiological objects at a time. Further estimation of the next-time expected value can be governed by the relaxation rate of the dynamical variables. Recently, a luminescent conjugate polyelectrolyte (LCP) is used for monitoring the amyloid plaques that can develop at various stages in human brain as known for Alzheimer's disease [3]. To diagnose Alzheimer's disease in the earlier stage, image color maps can be monitored as a non-invasive method. Image correlation spectroscopy is also used for protein dynamics on the cell surface and live-cell imaging to extract the information of aggregation and number density [4]. The desirable feature of the imaging correlation spectroscopy is suggested to achieve the diffusion coefficients and velocity for probing proteins in living cells [5]. Employing the medical imaging for the kinetic pathways of decaying or growing of bad cellular propagation is remained as a highly valuable issue. Much efficient and conclusive biomedical performance is needed to characterize the "urgency" in a simple and an "integrated" manner in time-resolving frames. Mesoscopic time and length scale imaging correlation spectroscopy could have its applications within biomedical technology of soft-tissues. The main difficulties to overcome will probably be related to small optical density differences. In case of soft-organs or tissues that consist of fibrous structures in the presence of multivalent ionic liquids, structures can be imaged through the birefringent nature of the fibrous structures. Then the orientation contrast can be conveniently visualized between crossed polarizers, which can be combined with time-lapsed imaging processing.

We employ charged fibrous viruses (fd) suspensions to explore low frequency electric field-induced phase transitions due to their anisotropic conformations (or configurations) under optical axis. Especially, the partial ordering is found as increased with increasing field amplitude [6,7]. This kind of fibrous structure can be also found in other biological materials, such as in optical nerve cells, living polymers of F-actin and myosin fibers. Particularly optical nerve cells are directly engaged with our human brain "cognitive" central nervous system. Interesting, it is already known fact that there are object recognition activities with low frequency oscillations (called Gamma band, ~ 40 Hz) in human brain system [8,9]. It is the purpose of this paper that we introduce our working system of dynamic image correlation spectroscopy to predict the mesoscopic relaxation time of dynamical states of small nematic domains (~ 10-30 μm size) that consist of charged fd-viruses particles.

Part of this research was supported by the EU-FP7 Network "NanoDirect" (contract number CP-FP-213948-2).

2. METHOD AND DATA ANALYSIS

2.1. Sample Preparation

Sample preparation is done by following the standard biological protocols [10] to prepare the fd-virus suspensions, from the XL1blue strain of E. coli as the host bacteria. The virus particles are purified by repeated centrifugation 10^5 g for 5 hours and re-dispersed in Tris/HCl buffers with varying concentration, depending on the final ionic strength. Bacteriophage fd is a rod-like macromolecule with a length of 880 nm, a bare diameter of D = 6.6 nm, a persistence length at high salt concentration of P = 2200 nm, and a molecular weight of M.W. ~ 1.64×10^7 g/mol. Much detailed information of molecular structures and dynamics of fd coat proteins are discussed in earlier times by means of scanning transmission electron microscope and solid-state NMR method [11-14]. Hydrodynamic properties and structure of fd-viruses are measured by transient electric birefringence, and low-angle intensity fluctuations spectroscopy [15].

We prepare fd-virus suspensions at a given low ionic strength by osmotic equilibration (or dialysis) for 2 days with an analytical concentration of 1.6 mM ionic strength of Tris/HCl buffer. The same buffer is used to dilute the suspensions (by typically 20 %) to the desired fd concentration. The ionic strength of the osmotic reservoirs is the ionic strength from which the electrostatic Debye-Huekel screening length is calculated. Furthermore, osmotic equilibration assures that the pH of the suspensions is equal to those of the osmotic reservoir. The Tris/HCl buffers are prepared by adding a small volume of HCl with a concentration of 1 M to a 20 mM TRIS solution till a pH = 8.2 is attained. This buffer is then diluted with deionized water for the preparation of the lower Tris/HCl-buffer concentrations. For buffer concentrations less than about 10 mM, both the ionic strength and the pH are affected by carbon dioxide that dissolves from the air, which in turn affects the surface charges of the fd-virus particles. A detailed account of the effect of dissolved carbon dioxide on the ionic strength and pH as a function of the buffer concentration can be found in [16]. Then the fd-concentration is measured by UV spectrometer to read optical density, or the extinction coefficient at approximately 269 nm wavelength.

The charge distribution on fd-virus particles has been interpreted on the basis of different models for an fd-virus particle from titration curves. The best fit to titration curves is found when assuming that the charges are located only on the hydrophilic outer region of adsorbed coat proteins, and not on the DNA strand of fd itself [17]. The iso-electric point of native fd is equal to pH ~ 4.2. The surface charge is measured approximately negatively charged as –9000e at pH = 6.9. The buffer capac-

ity of the low-ionic strength Tris/HCl buffer that we used (for 0.16 mM salt concentration) is still sufficient to keep the pH > 6.5 [16], despite dissolving carbon dioxide from the air.

2.2. Collection of Raw Image Intensity at Finite Time Steps

The realization of morphological changes in local events at a space is much sensitive when the system is governed by dynamically variable quantities (or objects). As far as the system maintains a uniquely defined phase or state for much long-time as compared to the resolution of appropriate time binning, a decent characteristic time can be extracted for their relaxation behaviors. Also the optimization of AOI (area of interests) averaging is relevant to obtain "ensemble" average correlations, which influences the good average of background in statistics at far later time after the event occurs. Images are collected with an appropriate finite time step to develop the evolution of spatio-temporal morphological changes. Depending on the nature of dynamical states in the physical system, we have found that the resolution of time binning in sampling has to be optimized such that at least 3-10 times smaller than the actual time of dynamical events. For example, the dynamical state itself evolves in few tens of second, then appropriate time binning should be treated within a second as a time step.

Intensity-time traces for dynamical images are then constructed from the intensity map in 2-dimensional area, which normally depends on the optimization between the quality of the CCD camera and area of interest (AOI) in system. **Figure 1** shows the example of collections of dynamical images at few seconds of time steps. Here the sampling time is every 1 sec, and total duration time is 1800 sec. The dynamical morphologies are driven by external AC electric field at a low frequency (of 10 Hz) and amplitude (of 3.2 V/mm). This dynamical state is still slowly changing in time and space, so-called slow dynamical state as D_S state. The change of visual observation of these slow dynamical states is typically occurred within 30 sec. By applying higher field amplitude, the slow dynamical state, D_S transits to the fast dynamical state, *i.e.*, D_f, where we have chosen multi-runs short time collections, as time binning as 300 ms and duration time for 60 s and 5 runs averaging. Due to the speed limitation of our camera, normally the camera setting ought to be optimized between fast speed and high reso lution option. We have optimized the CCD camera set as fast camera exposure time since the dynamical state domains are quite visible in CCD pixel by pixel.

We are then able to capture the dynamical events at least 3 or 4 different time steps within a sec temporal event in the fast dynamical, D_f-state.

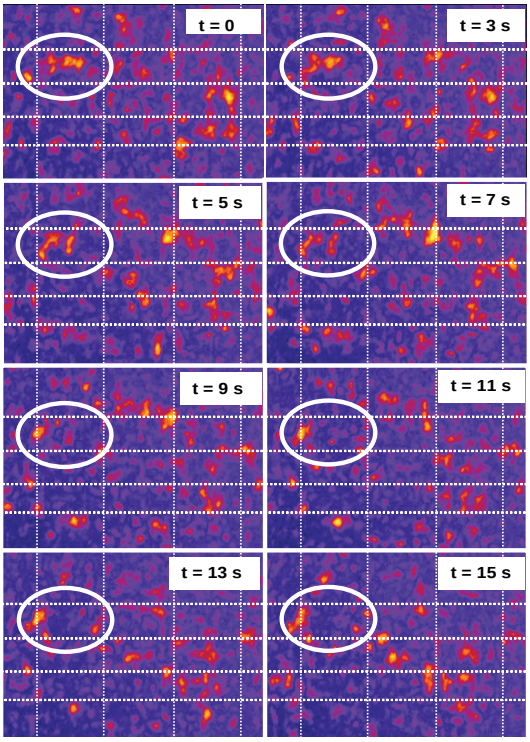


Figure 1. Temporal images of dynamical states that are induced by external electric AC field amplitude (of 3.2 V/mm) at a low frequency (of 10Hz). Field of view is of 400 μm × 300 μm, and the dotted lines are guiding eyes for detecting small neamtic domains of fd-virus particles. White circle is presented for tracing particular region of spatiotemporal images for dynamical states in each frame. Dynamical image correlation function is calculated as a resolution of single pixel averaging in 2-d over 1800 s duration time, where the flexible time resolution can be adapted as 0.3 s, 1 s, for fats and slow dynamical states, respectively.

It turned out that this time binning may not be fast enough to characterize "completely" in the fast dynamical states, where the relaxation behaviors could be realized at much microscopic time scale. As a consequence, the background value turned out to be occasionally below zero, indicating that the time step is too large to carry on the local events correlations. It would be desirable when the time binning of the sampling time has to be tested with the background value to present as the dynamical state. Also it has to be chosen as much faster exposure CCD camera to ensure the exposure time of camera setting is much faster than the actual duration time for measurement. Thus by allowing the time step is much smaller than the change of visual morphological changes in dynamical states, time-lapsed images are recorded within the chosen frame setting, and saved (or exported) as time-sequential bmp files with 8 bit gray data that is mostly common format in handling image objects.

2.3. Construction of Modified Intensity Matrix

To construct the intensity-time traces of images, the intensity matrix values of every pixel AOI (1 pixel × 1 pixel) are reconstructed for dynamic image correlation function, as a "readable" Ascii format to correlate with different time-frames of AOI. The modified intensity matrix is defined as the subtracted intensity average value of AOI pixel for every time frame from the raw intensity data. **Figure 2(b)** shows the simple sketch of the enhanced contrast between "bright" and "dark" intensity AOI pixels.

The modified intensity matrix is then reconstructed as a building block of individual time frame, such that the each time frame image is subtracted with the overall average intensity value of the initial time frame, as can be seen in **Figure 2(b)**. Then enhanced contrast is sketched as a bright (white) and dark (or black) unit pixel area, against the mean average of a gray color code. In principle, for a long time of duration, the change of the dynamical events can be completely "independent", so that at far later time after the event occurs, the time-correlation of dynamical states is expected to be " non-correlated" at the end of time event. This is briefly depicted in the long time later after an event occurs in **Figure 2(b)**, which has now shown flipped from "white" to "black" and "black" to "white" code.

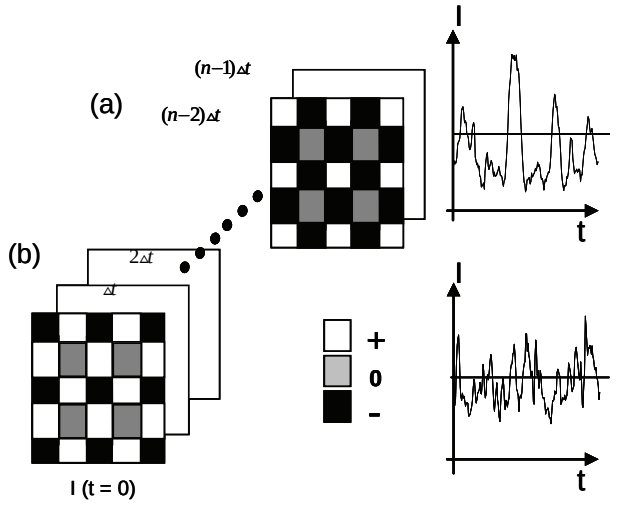


Figure 2. Simple sketches of the procedure of data analysis: (a) A time-lapsed sequence of raw images with a "readable" intensity matrix as 8 bit gray bitmap images. (b) A time-lapsed sequence of modified intensity matrix as a subtracted average intensity value at each time frame. Bright and dark intensity is contrasted in this modified intensity matrix. The resultant intensity traces over time is presented as typical slow and fast oscillations, and white, gray, and black unit cell represents the sign of the intensity values, notified in the modified intensity map in (b).

The right panels of **Figure 2(b)** are shown for the intensity—time traces for both slow and fast dynamical state, up and down figures, respectively. The apparent intensity oscillations are due to the spatio-temporal motion of dynamical states of the small nematic domains (of fd-virus particles).

2.4. Calculation of the Image Correlation Function

The correlation function is defined as

$$C(t) = \frac{\left\langle \left[I(t) - < I(t) > \right] \left[I(0) - < I(0) > \right] \right\rangle}{\left\langle \left[I(0) - < I(0) > \right]^2 \right\rangle} \qquad (1)$$

where I is the transmitted intensity at a given pixel and the brackets <…> denote averaging over all pixels. We made a program to calculate the dynamical image correlation function. We have found the best performance of the intensity averaging is 1 pixel by 1 pixel, as compared to 5 pixels by 5 pixels (or 10 pixels by 10 pixels) averaging. This may be related to a low magnification objective lens (10X objective) that is chosen for larger sampling of AOI in our system. **Figure 3(a)** shows the typical correlations functions of our dynamic imaging correlations as a function of increasing field amplitude (as indicated the arrow direction). Here, the relaxation time of dynamical states of melting and forming of small nematic-domains of fd-virus particles is presented as a function of applied field amplitude at a low frequency (of 10 Hz).

3. RESULTS AND DISCUSSION

To briefly summarize the field-induced phase/state transitions of charged fibrous viruses (fd) at a low ionic strength:

- At low driving frequency and a low field amplitude, the chirality is induced, which is seen as a "stripe" texture (as a chiral nematic, *N**-phase) in depolarized light.
- At higher field amplitude, this field-induced chirality is slowly disconnected and developed as smaller domains that contain chiral nematics, so called as N_D^*-phase.
- Further increasing of the field amplitude, there is a discontinuous transition between melting away of chirality and appearing of small nematic domains that are kinetically responding in space and time. Thus this leads to the dynamical states as slow and fast in an increase of the amplitude, as D_S and D_f -state, respectively.

Detailed electric phase/state diagrams are presented and characterized in [6].

As one can see in **Figure 3**, the relaxation time is decreased as increasing field amplitude; typical dynamical image correlation functions of a low frequency (10 Hz) are shown for varying the applied field amplitude

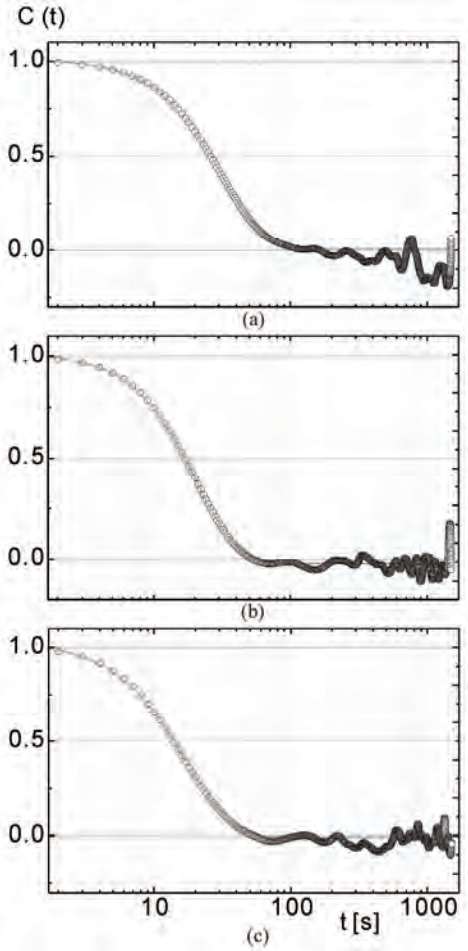


Figure 3. Typical dynamical image correlation functions of a low frequency (10 Hz) for a fd-concentration (of 2.8 mg/ml): Relaxation time of dynamical states is decreased as an increase of the applied AC field amplitude as (a) 3.1 V/mm, (b) 3.5 V/mm, and (c) 3.9 V/mm. Note that there are relatively maintaining noise distribution in the background at longer time tails due to the disturbance of the external field.

from 3.1 V/mm, 3.5 V/mm, and 3.9 V/mm. Also the microscopic discontinuity is observed in an increase of field amplitude at low frequency, at the transition of chiral-nematic phase to dynamical states (around 3.3 V/mm), by means of dynamic lights scattering [6,7]. The discontinuity may be due to the fact of "intrinsic" differences on the microscopic dynamics between chiral-nemtic phase and non-chiral nematic domains that are present between isotropic and nematic coexistent phase. At this transition, the chirality slowly disappears and at the same time, non-chiral small nematic domains are becoming dynamic, in terms of their melting and formation of the small non-chiral nematic domains (~ 20 μm). Further increasing field amplitude at a low frequency, the dynamical states of melting and forming of small nematic domains are becoming faster. Thus kinetics of dynamical states is enhanced by an increase of field amplitude, as it can be seen in **Figure 4(a)**. The mesoscopic relaxation times of these dynamical image correlation functions are indeed shown as "faster" as an increase of the field amplitude in **Figure 4(b)**. These dynamic image correlation functions of applied field amplitude at a low frequency can be fitted with an exponential decay with a stretching exponent. The solid lines are fitted with a fitting function of

$$C(t) = B + A\exp\left\{-(t/\tau)^{\beta}\right\} \quad (2)$$

where $B = 0$, $A \sim 0.5 - 0.6$, $0.5 < \beta < 1$, and $\boldsymbol{\tau}$ is the measure for a characteristic time for melting and appearing small nematic domains of fd-virus particles.

This characteristic time diverges on approach of the chiral-phase transition on lowering the field amplitude, in a mesoscopioc correlation time scale. Hoswever, the transition from the D_S- to- D_f -state is defined as the point where the sharp decrease of $\boldsymbol{\tau}$ with increasing field amplitude ceases to occur. The mesoscopic relaxation time has a tendency of saturation that the dynamics be

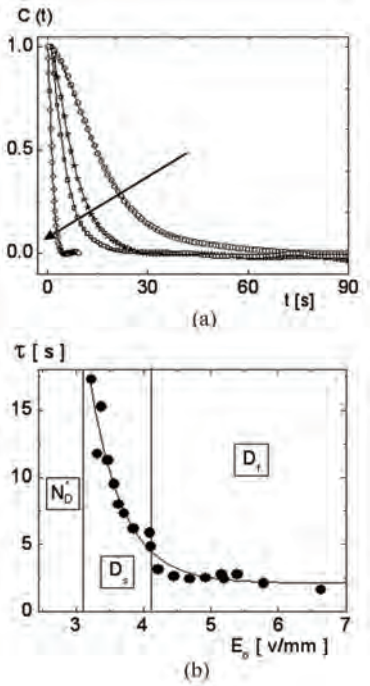


Figure 4. (a) Dynamical image correlation functions for various AC field amplitudes at low frequency (of 10 Hz) of the fd-concentration (2.8 mg/ml at 0.16 mM salt). Arrow indicates the direction of an increase of field amplitude (from 3.26, 3.78, 4.09, to 5.77 V/mm). Note that there is a very small modulation at longer time window; (b) Results of the mesoscopic relaxation time for dynamical image correlation functions for a low frequency (of 10 Hz).

comes essentially independent of the field amplitude. Whether this is due to the lack of our limited choice of time binning or the "intrinsic" feature of our dynamical system is not yet clear. We also have performed dynamic image correlation spectroscopy on approaching of the D_f -to-D_S -state transition, by increasing frequency at a high field amplitude, where "critical" slowing down behaviors of small nematic domains are found, where the critical exponents can be obtained [18].

Figure 4(a) has shown the results of dynamical correlation functions for various field amplitudes, where the arrow indicates the increase of the field amplitude. Disregard of the background oscillations, the relaxation time decreases with an increase of the field amplitude. This means that the fast dynamical states are relaxing much faster than the slow dynamical states, as it is expected in the time correlation method. It also maintains as a "reversible", which is a good candidate for "tuning" the dynamical states with external driving frequency. **Figure 4(b)** is the resultant relaxation time as a function of applied field amplitude, where the field-induced N_D^*-phase, slow, D_S, and fast dynamical D_f states are shown. Note that there are relatively "noisy" oscillations can be obtained in the long time tails (see **Figure 4(a)**), which may be due to the constant disturbance (or "feedback") of the external field. If the system is purely diffusive, then this correlation function has zero background. However, since the system is constantly feed backed by external alternating current amplitude, the intensity-time correlation function background does not go "directly" to absolute zero, but slightly oscillates around zero, although mathematically in principle should be "zero".

Thus we speculate these small contributions of white noise in terms of our low driving frequency as well. This can be useful for many biological imaging processing or biomedical electronics researches, since the signals are too much "weak" and "dark" to detect and to carry on in biological fluids and compartments (soft tissues).

Figure 5 indicates the zooming view of the background noise oscillations at larger time window for both slow (in (a)), and faster (in (b)) dynamical state. As one can see, the background oscillations are "modulated" in a complex manner, and the quality of the frequency is much "noisy" than the driving (or input) frequency.

In order to look at these possible oscillations in more detail, we propose a data analysis as follows. Prepare two independent backgrounds. One is from the measured back ground oscillation at larger times, as shown in **Figure 6(a)**. The other is a "virtual" background oscillation (the sinusoidal waveform that is chosen as comparable amplitude in **Figure 6(a)** for a frequency that is ~ 1% of driving frequency.

Typically ~ 10% signal-to-noise ratio is feed into in

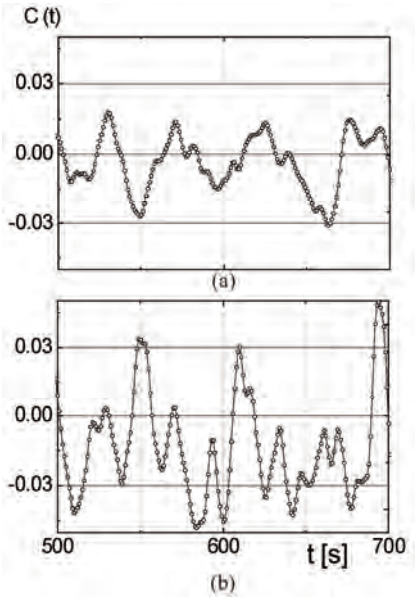


Figure 5. Zooming view of the background noise oscillation in the image correlation functions at longer time- window: (a) for slow dynamical state (3.24 V/mm), and (b) for faster dynamical state (3.86 V/mm).

the real correlation background noise. Then subtracted signals from the "virtual" oscillations are shown in **Figure 6(b)**, as given same time window, the resultant of the background signal is shown. The Fourier transform of the resulting signal gives the power spectrum given in **Figure 6(c)**. Dashed line is the estimation of the power spectrum by using a frequency (10 Hz) pass filter. Interestingly, there is a sharp major peak value found as 1% of input driving frequency, which is good agreement of cut-off envelop signal. Whether this analysis is sufficient to contribute to the interpretation of the dynamics of the melting and forming of nematic domains needs further investigation. However, the weak background oscillations are not influenced by the choice of the temporal-sensitivity, and indicate the intrinsic feature of field-induced dynamical states. It would be quite intriguing to know whether the smaller peaks in the power spectrum (between the major peaks) in **Figure 6(c)** can be amplified or modulated to tune the low frequency spectrum in the slowly varying dynamical systems. However this is beyond the scope of this paper.

4. CONCLUSIONS

For a dynamical image processing, we have used an experimentally working system for the time-lapsed dynamic image correlation spectroscopy. Series of collected time stacks of region of interests are chosen such that they maintain enough collective motion of single particles. In our experiment, typical average size of small nematic

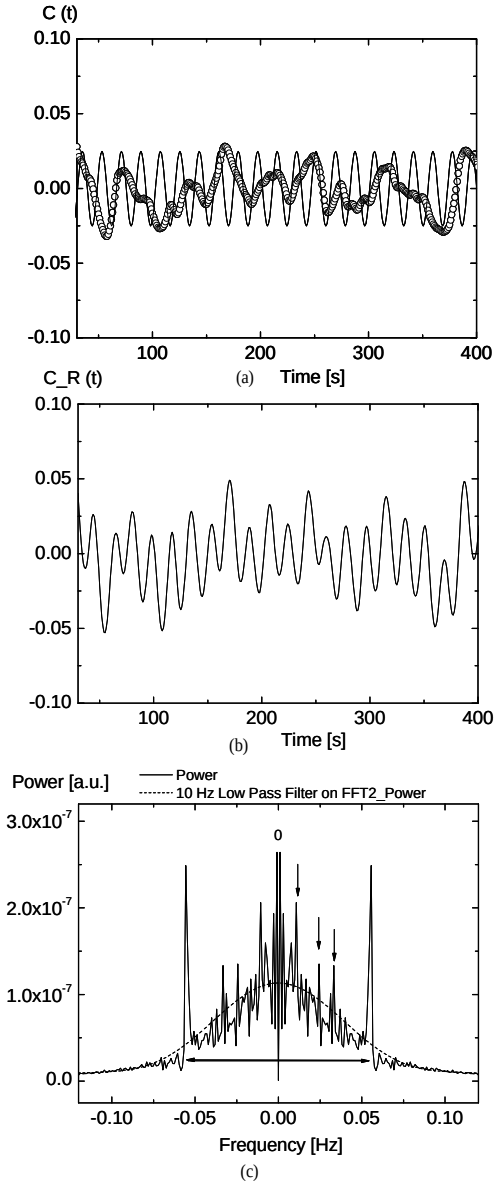


Figure 6. Further analysis of the background noise in the dynamical image correlation function: (a) Circle data points present a raw background signal, which is less than 5 % of total S/N ratio, and the thin solid line data is for an input signal with an optimized amplitude of an input frequency (of 10 Hz). (b) Resultant of the background signal, defined as input signal in (b) subtracted from raw signal in (a). (c) FFT of the resultant signal in (c). Notice that small peaks are shown within between the centered beam and the major peak around 0.1 Hz (around 1% of dark signal). Thus the modulation in a background noise of a dynamical image correlation function can be used as additional information to characterize the dynamical states in much weak signal-oriented physical systems.

domain is approximately averaged value of 20 μm, where the dimensions of an individual fd-virus particle is about ~1 um long and its diameter is ~ 0.03 μm (for a low ionic strength). Thus the transmitted intensity at a given pixel, which is calibrated as ~ 0.96 μm/pixel conversion rate, is calculated for the image correlation function, as normalized with averaging over all pixels. The image speckle patterns of the dynamical states are obtained from the modified image sequence at a certain time step, subtracted from the raw image, and image correlation function with an initial intensity profile (at a time t = 0 s). It turned out to be important to choose the initial frame of the dynamical event as much as in quasi-equilibrium process to carry on the meso-scopic relaxation dynamics. For instance, we perform to recording after 3-5 min. of waiting after applying field condition to ensure that more or less the same dynamical events are occurred during the measurement. Also depending on the characters of the dynamical events, either slow or fast-process, the time binning should be at least ~ 3-10 times faster for extracting the appropriate time resolution. The resulting correlation function in time of the speckles (or our small nematic domains) is found to be as an exponential decay function with a slightly stretched exponent. The mesoscopic relaxation time of melting and forming of small nematic domains is then obtained systematically by varying the field amplitude at a low frequency. However, the mesoscopic relaxation time constant in "fast" dynamical states has tendency of saturation with the field amplitude at higher field amplitude. This may be related to either the lack of the resolution in our experiments, or the genuine behavior in the dynamical system. Thus it would be worthy to test the prediction of the relaxation time with an image correlation spectroscopy for other similar dynamical events of biological interest or medical environments.

Interestingly, we have also found that the time constant for the relaxation behavior diverges on approaching to the chiral-phase regime from the transition of dynamical states (close to ~ 3.2 V/mm). This may imply the dynamical states are the consequence of the interaction of charged fd-rods via induced polarization effect in thick electric double-layers. We are currently developing the possible theory for an explanation.

To conclude, mesoscopic time and length scale imaging correlation spectroscopy needs a challenge to customize the biomedical technology of soft-tissues. The main difficulties are not only related to the density differences, but also the "weak" and "reliability" of signals/images of temporal evolution in micro-structural contrast. However as far as any soft-organs or tissues consist of fibrous structures in the presence of multivalent ionic liquids, any types of the elastic envelops can be monitored by their orientations in biological fluids. Then the orientation contrast can be conveniently visualized by depolarized light that can be combined with time-lapsed

imaging processing, as we have introduced in this report. If one could come up with an integrated package for fast time-lapsed imaging processing to pixel-to-pixel intensity auto correlation signal progressing, then this could be an application to characterize the mesoscopic relaxation behaviors of dynamical situations in many relevant soft biological and biomedical front matters. It can be adapted to early detection or "progressive" medicare issues. For instance, biological fluids (as blood, saliva, urine) can be accessed by in-vitro method under a low external electric field to recognize the "abnormality" related the growth/decay of "foreign" cells in terms of the change of microscopic "viscosity" or "resistance".

In this aspect, the "controllable" non-equilibrium approach of the dynamical image correlation spectroscopy is much more informative than a steady-state or quasi-equilibrium image processing.

5. ACKNOWLEDGEMENTS

Author thanks to J.K.G. Dhont for his scientific insight and introduce the flexible-time resolving image correlation functions to characterize the relaxation time in slow and fast dynamical states.

REFERENCES

[1] Eason, G., Brix, G., Kiessling, F., Lucht, R., Darai, S., Wasser, K., Delorme, S. and Griebel, J. (2004) Microcirculation and microvasculature in breast tumors: Pharmacokinetic analysis of dynamic MR image series. *Magnetic Resonance in Medicine*, **52**, 420-429.

[2] Donahue, K.M., Weisskoff, R.M. and Burstein, D. (1997) Water diffusion and exchange as they influence contrast enhancement. *Journal of Magnetic Resonance Imaging*, **7**, 102-110.

[3] Peter, K., Nilsson, R., Aslund, A., Berg, I., Nystrom, S., Konradsson, P., Herland, A., Inganas, O., Stabo-Eeg, F., Lindgren, M., Westermark, G.T., Lannfelt, L., Nilsson, L.N.G. and Hammarstrom, P. (2007) Imaging distinct conformational states of amyloid-β fibrils in Alzheimer's Diseases using novel luminescent probes. *ACS Chemical Biology*, **2**, 553-560.

[4] Wiseman, P. and Peterson, N.O. (1999) Image correlation spectroscopy II, optimization for ultrasensitive detection of preexisting platelet-derived growth factor-βreceptor oligomers on intact cells. *Biophysical Journal*, **76**, 963-977.

[5] Costantino, S., Comeau, J.W.D., Kolin, D.L. and Wisemann, P. (2005) Accuracy and dynamic range of spatial image correlation and cross-correlation spectroscopy. *Biophysical Journal*, **89**, 1251-1260.

[6] Kang, K. and Dhont, J.K.G. (2010) Electric-field induced transitions in suspensions of charged colloidal rods. *Soft Matter*, **6**, 273-286.

[7] Kang, K. (2009) Image and signal correlation of electric-field induced charged fibrous viruses (fd). *Proceedings of the* 31st *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Francisco, 1294-1297.

[8] Csibra, G., Davis, G., Spratling, M.W. and Johnson, M.H. (2000) Gamma oscillations and object processing in the infant brain. *Science*, **290(5496)**, 1582-1585.

[9] Baudry, C. and Bertrand, T.O. (1999) Oscillatory gamma activity in humans and its roles in object representation. *Trends in Cognitive Sciences*, **3(4)**, 151-162.

[10] Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor, New York.

[11] Lipka, J.J., Lippard, S.J. and Wall, J.S. (1979) Visualization of polymercurimethane-labeled fd-bacteriophage in the scanning transmission electron microscope. *Science*, **206(4425)**, 1419-1421.

[12] Cross, T.A. and Opella, S.J. (1979) NMR of fd coat protein. *Journal of Supramolecular Structure*, **11(2)**, 139-145.

[13] Gall, G.M., Cross, T.A., Diverdi, J.A. and Opella, S.J. (1982) Protein dynamics by solid state NMR: Aromatic rings of the coat protein in fd bacteriophase. *Proceedings of the National Academy of Sciences*, **79(1)**, 101-105.

[14] Schwind, P., Kramer, H., Kremser, A., Ramsberger, U. and Rasched, I. (1992) Subtilisin removes the surface layer of the phage fd coat. *European Journal of Biochemistry*, **210(2)**, 431-436.

[15] Newman, J., Swinney, H.L. and Day, L.A. (1977) Hydrodynamic properties and structure of fd-virus. *Journal of Molecular Biology*, **116(3)**, pp. 593-606.

[16] Kang, K., Wilk, A., Patkowski, A. and Dhont, J.K.G. (2007) Diffusion of spheres in isotropic and nematic networks of rods: Electrostatic interactions and hydrodynamic screening. *Journal of Chemical Physics*, **126(21)**, 214501-(1-17).

[17] Zimmermann, K., Hagedorn, J. Heuck, C.C., Hinrichsen, M. and Ludwig, J. (1986) The ionic properties of the filamentous bacteriophage Pf1 and fd. *The Journal of Biological Chemistry*, **261,** 1653-1655.

[18] Kang, K. and Dhont, J.K.G. (2009) Criticality in a non-equilibrium, driven system: Charged colloidal rods (fd-viruses) in electric fields. *The European Physical Journal E*, **30(3)**, 333-340.

J. Biomedical Science and Engineering, 2010, 3, 633-637 **JBiSE**
doi:10.4236/jbise.2010.36086 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

The application of hidden markov model in building genetic regulatory network

Rui-Rui Ji, Ding Liu, Wen Zhang

Department of Automation and Information Engineering, Xi'an University of Technology, Xi'an, China.
Email: jirui@xaut.edu.cn; liud@xaut.edu.cn; zhangwendec@yahoo.com.cn

Received 6 March 2010; revised 15 April 2009; accepted 25 April 2009.

ABSTRACT

The research hotspot in post-genomic era is from sequence to function. Building genetic regulatory network (GRN) can help to understand the regulatory mechanism between genes and the function of organisms. Probabilistic GRN has been paid more attention recently. This paper discusses the Hidden Markov Model (HMM) approach served as a tool to build GRN. Different genes with similar expression levels are considered as different states during training HMM. The probable regulatory genes of target genes can be found out through the resulting states transition matrix and the determinate regulatory functions can be predicted using nonlinear regression algorithm. The experiments on artificial and real-life datasets show the effectiveness of HMM in building GRN.

Keywords: Genetic Regulatory Network; Hidden Markov Model; States Transition; Gene Expression Data

1. INTRODUCTION

In order to understand the functioning in living organisms, it should be known which genes are expressed, when and where, and to which extent. The regulation of gene expression is achieved through the interactions between DNA, RNA, proteins, and small molecules. This regulatory system can be described by the structure of network called genetic regulatory network (GRN). Building GRN is a reverse-engineering of real gene expression data to help studying the relationship between genes systematically, understanding the essential rule of biological phenomena and provide valuable idea to treat some complex diseases [1,2]. Many mathematics models have been proposed during GRN research, such as Boolean model, Linear combination model, Weighted matrix model, Differential equations model, and so on [3,4]. Resent years, probabilistic GRN model has been paid more attention for the real biological system is stochastic and the determinate model cannot infer the complex process. Bayesian network and Markov Chain have been studied [5-7]. The final result of probabilistic GRN is represented using a graph consisting vertices (genes) and edges (relationships). The relationships between genes are described through probability. Considering the current probabilistic GRN is simple and cannot give the dynamics behavior. This paper studies the application of Hidden Markov Model (HMM) and nonlinear regression algorithm to GRN building.

This paper is organized as follows: Section 2 gives a theoretical representation of HMM. Section 3 focuses on the application of HMM in building GRN. Section 4 shows the experimental result and discussion. Section 5 is the conclusion.

2. HIDDEN MARKOV MODEL

2.1. Basic Theory

Classical Hidden Markov Model is a kind of stochastic state machine based on statistical signal, which is a double stochastic process where the sequences of states can not to be observed directly. The fundamental stochastic process in this model is a Markov chain which describes the state transition; another stochastic process describes the statistical corresponding relationship between states and the observed value. Different states generate particular sequences of observation according to different probability. The observer can know the feature of the state only through the stochastic process. So this model is called Hidden Markov Model.

A parameter set $\lambda = (\pi, A, B)$ is used to describe a HMM. Where π means the initial state probability, A is the states transition matrix and B is the observed probability. Suppose the number of state is N, $S = [S_1, S_2, ..., S_N]$ is the state set. q_t means the state of the model at time t, $1 \le t \le T$, $O = (O_1, O_2, ..., O_T)$ is the observed value sequence and v_k is the observed value, the three parameters in a HMM model are defined as follows,

$$\pi = [\pi_1, \pi_2, \ldots, \pi_N], \quad \pi_i = P_r(q_1 = S_i) \tag{1}$$

where π_i is the probability of the initial state q_1 equals S_i.

$$A=(a_{ij}),\ a_{ij}=P_r(S_j/S_i),\ \sum_{j=1}^{N}a_{ij}=1 \quad (2)$$

where a_{ij} means the transition probability from state S_i to state S_j. A is a $N\times N$ matrix.

$$B=(b_j(k)),\ b_j(k)=P_r(v_k/s_j),\ \sum_{k=1}^{M}b_j(k)=1 \quad (3)$$

$b_j(k)$ means the observed probability that generating v_k under state s_j. B is a M × N matrix.

2.2. Fundamental Problems in HMM

There are three problems needed to be solved when building a HMM.

1) Evaluation

For a given HMM $\lambda=(\pi,A,B)$ and the observed sequence $O=(O_1,O_2,...,O_T)$, evaluation means calculating the probability $P(O|\lambda)$ corresponding the observed sequence generated by the model, which can evaluate the similarity of the observed sequence with the given model.

Forward-Backward Algorithm proposed by Baum is used to solve the evaluation problem. In practical application, the results are very small and usually normalized or carried on logarithmic operation in calculation process.

2) Decoding

Ascertaining an optimal state transition sequence $Q^*=(q_1^*,q_2^*,...q_T^*)$ on the given HMM $\lambda=(\pi,A,B)$ and observed sequence $O=(O_1,O_2,...,O_T)$ is called decoding.

Viterbi Algorithm can solve the decoding problem. Here, the optimal state transition sequence $Q^*=(q_1^*,q_2^*,...q_T^*)$ is the Q which can make the value of $P(Q,O|\lambda)$ maximum.

3) Learning

Learning can obtain the optimal parameters λ^* of HMM through training algorithm, where λ^* make the value of $P(O|\lambda)$ maximum, for the given HMM $\lambda=(\pi,A,B)$ and observed sequence $O=(O_1,O_2,...,O_T)$.

There are two solutions to solve the learning problem. One uses gradient technique, another is based on iteration or recursion like Baum-Welch algorithm which is often used to train parameters of HMM.

3. MODELING GRN BASED ON HMM

3.1. Constructing HMM

A fundamental assumption is that genes sharing similar expression levels are commonly regulated, and the genes are involved in related biological functions. Most of GRNs are built on the basis of clustering. The process of constructing HMM in this paper is carried on the genes clustered into the same class.

1) States:

Considering the genes clustered into same class, different genes are considered as different states. So the size of state transition matrix is just the same as the number of genes. State transition probability corresponds with the regulatory probability between genes. One gene may be regulated by any other genes, even itself, so the wholly connected connection structure is used.

2) Observed sequence:

The expression profiles of genes are considered as the observed sequence. Since these data is easily affected by noise, smoothing is used firstly to reduce the influence of noise.

3) Training steps:

Step 1: initializing parameters λ_0 of HMM. The number of states is equal to the number of genes and each value of state transition matrix is initialized as average value $1/N$, N is the number of states, and $P(O|\lambda_0)$ can be computed;

Step 2: revaluating HMM's parameter λ_0. Baum-Welch algorithm is used to train HMM model to acquire λ;

Step 3: computing $P(O|\lambda)$ under the obtained model λ using Forward-Backward algorithm;

Step 4: judging the convergence criterion. If $|P(O|\lambda)-P(O|\lambda_0)|\le\varepsilon$ is not satisfied, then $\lambda_0=\lambda$ and return to step 2. Else, training process is finished and final HMM model close to the observed sequence can be acquired.

3.2. Building Probabilistic GRN

Regulatory genes for each target gene can be found out based on the state transition matrix after training HMM. Then the structure of regulatory network can be built according to the following steps:

Step 1: For one target gene $x_i(i=1,2,...,N)$, the genes whose transition probability in the trained state transition matrix A is bigger than the initial average probability are found out and these genes are regarded as the parental regulatory genes of each target gene;

Step 2: repeating step 1 until finding out the global information for each target gene;

Step 3: predicting the determinate regulatory function f_i between target gene and its parental regulatory genes using multiple nonlinear regression and least squares algorithm.

4. EXPERIMENTS AND DISCUSSIONS

4.1. Experiment with Artificial Data

In order to evaluate the efficiency of our algorithms, a group of networks are required whose structure had been known. However, the real structures of GRN are unknown completely because the research about GRN is

still in an early stage. So artificial data reported in paper [8] are used in this study. Here, the adopted ALARM network contains 37 discrete variables, 46 edges and the value of every variable ranges from 2 to 4.

The network with known structure is called target network N_t and the result of our algorithm is called deduced network N_d. Three index sensitivity, specificity and F-factor are used to evaluate our algorithm. Sensitivity is used to test the inference ability, specificity reflects the degree of accuracy and F-factor is the balance of above two indicators. The bigger F-factor means the higher accuracy.

$$Sensitivity = \frac{s_1}{s_t}, \quad Specificity = \frac{s_1}{s_d} \tag{4}$$

where, s_1 means the number of the same edges in both and N_d and N_t; s_t means the total number of the edges in N_t; s_t means the total number of the edges in N_d.

Table 1 shows the comparison of standard Simulated Annealing algorithm, called BANJO developed by Hartemink and our algorithm. It can be seen that our result is better than general Simulated Annealing algorithm in all items.

4.2. Experiment with Real-Life Data

The real-life experimental data in this paper comes from yeast cell cycle expression datasets created by Spellman [9], which imply the regulatory information about genetic property in yeast cell cycle. However, the above mentioned index cannot be used to evaluate the result because the real biological regulatory network is unknown completely. So the existing regulatory relationships which had been already proved are used to evaluate our results.

As listed in Futcher's paper [10], 15 main transcriptional factors: MBP1, SWI4, SWI6, FKS1, MCM1, FKH1, NDD1, SWI5, ACE2, CDC28, CLN3, CLB2, SIC1, CLN2 and HHT1 are discussed. It had been verified that there exist interactions among these genes. For convenience, these genes are marked No. 1, 2,…15.

Table 1. The comparison of BANJO and HMM.

algorithm	BANJO	HMM
right side number	22	31
reverse side number	2	5
sensitivity	0.48	0.67
specificity	0.29	0.76
F-factor	0.36	0.71

1) Experimental results

Initial elements of state transition matrix are assigned as 1/15, **Table 2** is the final state transition matrix trained, and each row shows the transition probability corresponding with target gene, through which the probable regulatory genes of target genes can be found out.

Considering the transition probability of MBP1, SWI6, MCM1 and CLN3 to gene SWI4 are bigger than initial probability, so these genes can be regarded as the regulatory genes of target gene SWI4. **Table 3** lists the regulatory genes of target genes SWI4, SWI5, CLN2, and CLB2.

2) Determinate regulatory relationships

Predicting the determinate regulatory relationships can offset the drawbacks of the probabilistic GRN, which can not describe the specific dynamics behavior. The obtained regulatory relationships between target gene CLB2 (x_{12}) and its regulatory genes MCM1 (x_5) and SIC1 (x_{13}) is:

$$x_{12} = -0.1503 + 0.0943x_5 + 0.0075x_5^2 + 0.0039x_{13} - 0.0026x_{13}^2 - 0.17x_5x_{13}$$

Figure 1 compares the real expression profile of gene CLB2 marked in blue and its predicted value marked in black.

Figure 2 gives a local structure of our resulting GRN. Where, black connecting lines represent the verified existing edges [11]. Blue connecting lines represent the reversed direction with known relationships. Red con-necting lines represent the regulatory relationships pre-

Table 2. The states transition matrix.

Gene	MBP1	SWI4	SWI6	MCM1	SWI5	ACE2	CLN3	CLB2	SIC1	CLN2
MBP1	0.0542	0.0832	0.0623	0.0583	0.1014	0.0501	0.1357	0.0602	0.0532	0.0536
SWI4	0.0348	0.0523	0.0401	0.0347	0.0434	0.0482	0.1523	0.0531	0.0379	0.1768
SWI6	0.0629	0.0961	0.0658	0.0659	0.0653	0.0662	0.0631	0.0661	0.0598	0.0652
MCM1	0.0435	0.0759	0.0403	0.0431	0.1108	0.0463	0.103	0.1204	0.0446	0.1462
SWI5	0.0622	0.0635	0.0608	0.0592	0.0642	0.0631	0.0629	0.0633	0.1261	0.0642
ACE2	0.0582	0.0596	0.0574	0.1137	0.1302	0.0602	0.0588	0.058	0.0569	0.0611
CLN3	0.0578	0.1138	0.0584	0.0603	0.0536	0.0548	0.0637	0.0581	0.0592	0.0514
CLB2	0.0571	0.0562	0.0583	0.0615	0.0592	0.1583	0.0583	0.0633	0.0544	0.1019
SIC1	0.0585	0.0599	0.0592	0.0623	0.0582	0.0592	0.0552	0.0538	0.0653	0.0916
CLN2	0.0598	0.0635	0.066	0.0576	0.0658	0.0649	0.0538	0.0498	0.0531	0.0653

Table 3. The regulatory genes of several target genes.

target gene	regulatory genes
SWI4	MBP1、CLN3、MCM1、SWI6
SWI5	ACE2、MCM1、MBP1
CLN2	MCM1、SWI4、CLB2、SIC1
CLB2	MCM1、SIC1

dicted by our algorithm which is remained to be verified further.

3) Discussions

It had been verified in Reference [12] that gene CLN3 activates gene SWI4, gene SWI4 regulates gene CLN2 meanwhile, gene CLN1 and CLN2 are both influenced by gene CLB2, and CLB2 regulates CLN2. The expression pattern of gene SWI5 is similar with SIC1 and it has been proved that SWI5 regulates SIC1. These conclusions confirm that our algorithm is effective.

Moreover, it had been verified that CLN3 is regulated by gene SWI4 and CLB2 is regulated by gene SIC1, which is identical with the predicted results by our algorithm.

5. CONCLUSIONS

This paper discusses the application of HMM in building GRN. The regulatory genes for each target gene can be found out through the state transition matrix and then the global structure of GRN can be determined. Simulative experiment proves that this algorithm is more effective. The results in real-life data also show its rationality. Compared with the determinate model, HMM is more scientific because it describes the transcriptional regulatory degree between genes through probability. Especially, the present algorithms can find out self-regulatory relationships of genes.

There are still many problems should also be considered during the research of GRN using HMM, for example, how to choose the initial model. Since the biological GRN is a time-continuous and complicated dynamic system and haven't been completely known, as a result, how to evaluate the GRN integrated with biological meanings effectively is the next research.

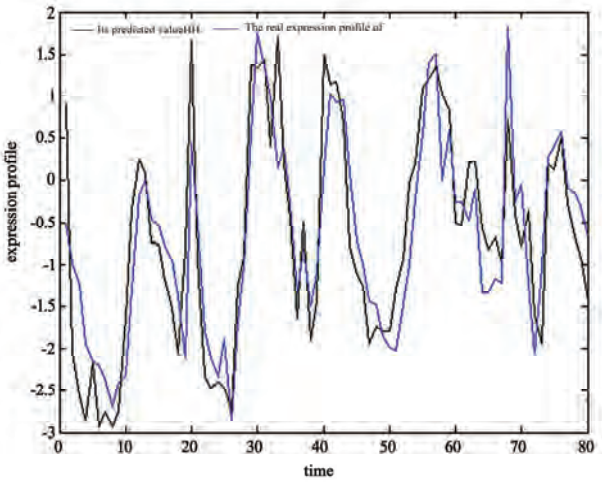


Figure 1. The real expression profile and regression result of gene CLB2.

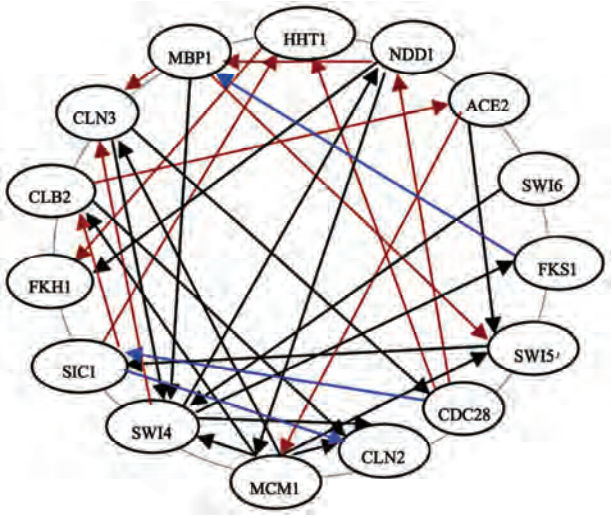


Figure 2. The local structure of GRN.

REFERENCES

[1] Zhao, G.P. (2002) Bioinformatics. Science Press, Beijing.

[2] Hidde, D.J. (2002) Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, **9(9)**, 67-103.

[3] Akustu, T., Miyano, S. and Kuhara, S. (2000) Inferring Qualitative relations in genetic networks and metabolic arrays. *Bioinformatics*, **16(8)**, 727-734.

[4] Bower, J. (2001) Computational modeling of genetic and biochemical networks. MIT Press, Cambridge.

[5] Hartemink, A., Gifford, D., Jaakkola, T., *et al*. (2002) Bayesian methods for elucidating genetic regulatory networks. *IEEE Intelligent Systems*, **17(2)**, 37-43.

[6] Ching, W., Fung, E., Ng, M. and Akustu, T. (2005) On construction of stochastic genetic networks based on gene expression sequences. *International Journal of Neural Systems*, **15(4)**, 297-310.

[7] Zhang, S.-Q., Ching, W.-K. and Yue, J. (2008) Construction and control og genetic regulatory networks: A multivariate Markov chain approach. *Journal of Biomedical Science and Engineering*, **1**, 15-21.

[8] Tsamardinos, I., Brown, L.E. and Aliferis, C.F. (2006) The max-min hill-climbing bayesian network structure learning algorithm. *Machine Learning*, **65(1)**, 31-78.

[9] Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Amders, K., Eisen, M.B., Brown, P.O., Botstein, D. and Futcher, B. (1997) Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Molecular Biology of the Cell*, **9(12)**, 3273-3297.

[10] Futcher B. (2002) Transcriptional regulatory networks and yeast cell cycle. *Current Opinion in Cell Biology*,

14(6), 676-683.

[11] Paul A.F. M.D. (2004) IncyteDB/OL. http://www.i-ncyte.com/proteome/YPD

[12] Zhang, Z.-F. (2004) Constructing and predicting gene regulatory network using micro-array data. National Central University, Taiwan.

J. Biomedical Science and Engineering, 2010, 3, 638-643 **JBiSE**
doi:10.4236/jbise.2010.36087 Published Online June 2010 (http://www.SciRP.org/journal/jbise/).

Classification of femoral neck fractures according to pauwels: interpretation and confusion

——Reinterpretation: a simplified classification based on mechanical considerations

Andrej Maria Nowakowski[1], Peter Emil Ochsner[2], Martin Majewski[1]

[1]Orthopedic Department, University of Basel, Basel, Switzerland;
[2]Clinic for Orthopedics Lucerne, Klinik St. Anna, Luzern, Switzerland.
Email: andrej.nowakowski@unibas.ch

Received 29 March 2010; revised 3 May 2010; accepted 9 May 2010.

ABSTRACT

The Pauwels Classification of femoral neck fractures, initially published in 1935, is used world-wide. Unfortunately, modern textbooks give varying angle and anatomic specifications between the classified fracture grades. This inconsistency is perpetuated in the literature, so that it is difficult to compare conclusions made by different authors. Pauwels himself left room for interpretation. He published two studies, one in 1935 and one in 1973, each including 3 diagrams. The 1935 version cited an angle of 8° representing the vector of static forces acting on the femoral head. The 1973 diagrams, however, cited an angle of 16° to represent dynamic forces, without changing the angle from horizontal. This already complex scheme is complicated by the fact that it depends on other factors such as femoral neck shaft (CCD) angle, femoral neck and head diameter, and/or distance of the fracture from the center of the femoral head. The multitude of factors argues against a rigid classification based on fixed angles from horizontal. Pauwels himself did not establish fixed critical angles between the fracture grades. In his own explanation of the system, he placed more value on mechanical considerations such as compression stress, shear stress, tensile force, shearing force, and torque. We propose therefore a simplified version of the Pauwels Classification: Grade I for fractures impacted in valgus, Grade II for fractures without free torque, and Grade III for fractures with free torque.

Keywords: Pauwels; Garden; Classification; Femoral Neck Fracture

1. INTRODUCTION

There is a pervasive wish for a simple classification system corresponding to the treatment of femoral neck fractures. One frequently used classification is that of Pauwels, originally published in 1935 [1]. Unfortunately, modern textbooks give varying angles and anatomic specifications to differentiate the three fracture grades. These inconsistencies are perpetuated in the scientific literature, so that it is difficult to make a clear correlation between fracture type, treatment modality and course of healing (**Table 1**).

The ambiguity of this system has doubtless contributed to the lack of consensus regarding its use. However, the classification per se is not at fault, instead, its interpretation. For example, some authors ignore entire segments (50° to 70°) of injury in their interpretations. In addition, there are numerous publications that make comparison impossible, since they do not indicate which interpretation of the classification they are using [2-7].

Bartoniček [8] remarked on this problem in 2001 in his study, “Pauwels’ Classification of Femoral Neck Fractures: Correct Interpretation of the Original,” and stipulated the following categories:

Grade I < 30°, Grade II 30-50°, and Grade III > 50°.

However, Pauwels considers the fracture in his Figure 40 (page 39), to be a Grade I injury despite an angle of 35°. This, he explains, is because impaction negates the shearing forces [1]. (**Figure 1**).

Redefining the distinction between Grade I and II fractures as 35° instead of 30° is counterproductive. Instead, the original intent of Pauwels should be respected. This was to classify these fractures based on mechanical considerations.

2. MATERIAL AND METHODS

We examined the classification of femoral neck fractures based on pre-determined angles, using the two original publications from Pauwels [1,9]. In addition, theoretic considerations and modifiable variables were compiled

Table 1. Diverse interpretations of the pauwels classification.

Grade Ⅰ	Grade Ⅱ	Grade Ⅲ	Literature
< 30º	30-50º	> 50º	[80,10,14,19,21,22]
< 30º	30-60º	> 60º	[23]
< 30º	30-70º	> 70º	[13,20,24,25]
30º	50º	70º	[12,26,27]
< 50º	50-65º	> 65º	[28]
< 30º	30-50º	> 70º	[29]

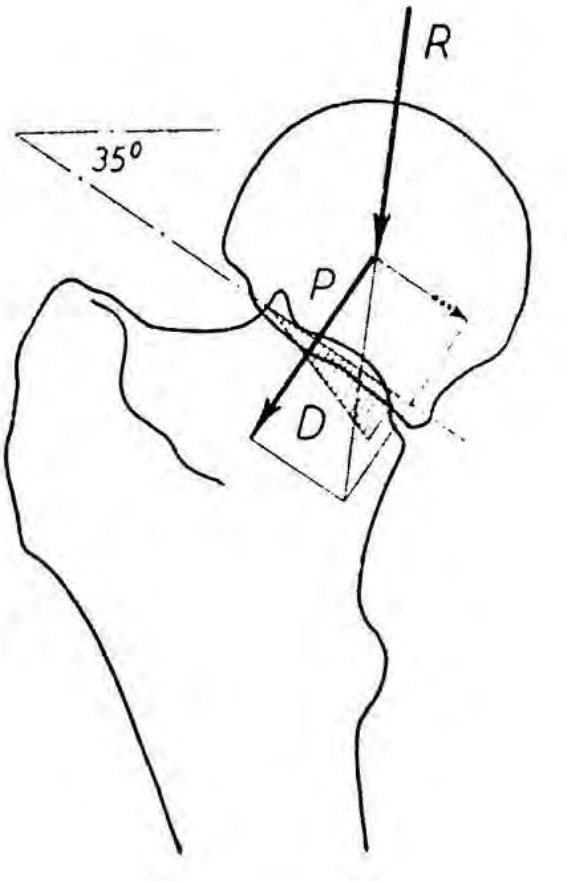


Figure 1. Original illustration by Pauwels 1935 [1] showing a fracture impacted in valgus without free shearing force (Grade I). Reprinted by courtesy of Georg Thieme Verlag KG, Stuttgart, after transfer of Ferdinand Enke Verlag, Stuttgart, from 1971.

and incorporated into a simplified version of the Pauwels' Classification.

3. RESULTS

The most important criteria for the Pauwels Classification are found in the legend describing his **Figure 2** [1].

Grade I: The partial compressive force component P acts only as its magnitude approaches the value of R. The partial force of pure displacement S is countered by friction at the fracture site. P generates compression forces distributed evenly over the entire fracture surface, because R intersects the fracture surface adjacent to the mid-line.

Grade II: Here the free shearing force Ks will act on the zone of fracture healing. The compression component P will not act on this area, instead will generate latent compressive forces in only one portion of the fracture surface.

Grade III: The free shearing force Ks and tensile force Z are generated by free torque (R, OA).

Only the lowest corner of the trochanter fragment (O), the fulcrum for the head fragment, will be acted upon by latent compressive forces.

Differentiation between Grade I and Grade II fractures, then, is made based on the presence or absence of the free shearing force Ks. When shearing force is present, the fracture is a Grade II. This explains why some fractures impacted in valgus, despite steep fracture angles, are considered Grade I fractures. Impaction neutralizes Ks. (**Figure 1**) According to Pauwels [1]: "Under certain circumstances, impaction may have a decided influence on the course of healing, because it neutralizes a certain measure of the displacement force S and therefore reduces the extent to which the free shearing force Ks acts on the zone of fracture healing." In cases of more vertical fractures, however, the compression achieved from impaction can no longer neutralize the dynamic shearing forces [1].

The differentiation between Grade II and III fractures is made based on the free torque (R, OA). When free torque is present, the fracture is Grade III.

Interestingly, almost 40 years after the initial paper, Pauwels published modified diagrams [9]. These stipulate a dynamic force vector R with an angle of 16° instead of the static vector of 8° found in the original version. (**Figure 3**)

It is possible that improved operative techniques and implant devices allowing partial weight bearing and early mobilization of the patient led to this alteration.

Pauwels wrote even in 1935 [1] that Grade II fractures are those with a fracture inclination up to 50°. His own diagram shows a fracture with a distinctly steeper fracture angle than 50°, however, without free torque. (**Figure 3(c)**) Manipulation of the vector angles alone, dependent on other factors, can blur the distinction between categories, *i.e.*, the values of the angles of inclination.

Figure 4 demonstrates this for the border between Grade II and III fractures. (**Figure 4**)

The location of the fracture along the femoral neck medially or laterally has a definite influence on the presence or absence of free torque. (**Figure 5**)

Other factors influencing the critical angles are: the femoral neck shaft (CCD) angle, the femoral neck and head diameters, and/or distance from the fracture to the

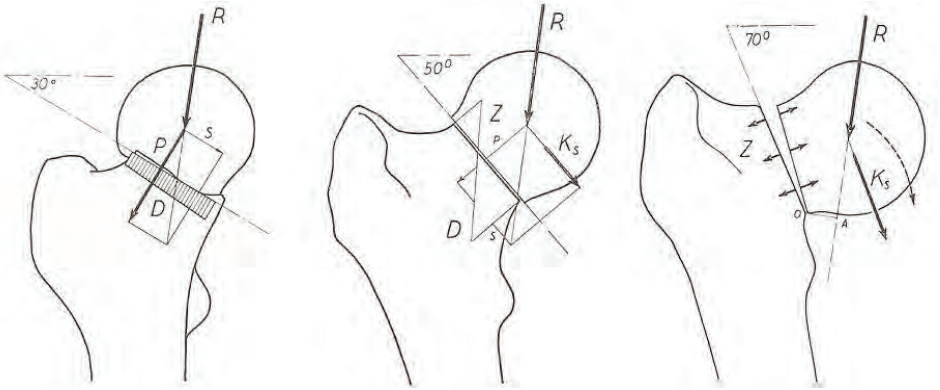


Figure 2. Original illustration by Pauwels 1935 [1]: Classification of femoral neck fractures. Reprinted by courtesy of Georg Thieme Verlag KG, Stuttgart, after transfer of Ferdinand Enke Verlag, Stuttgart, from 1971.

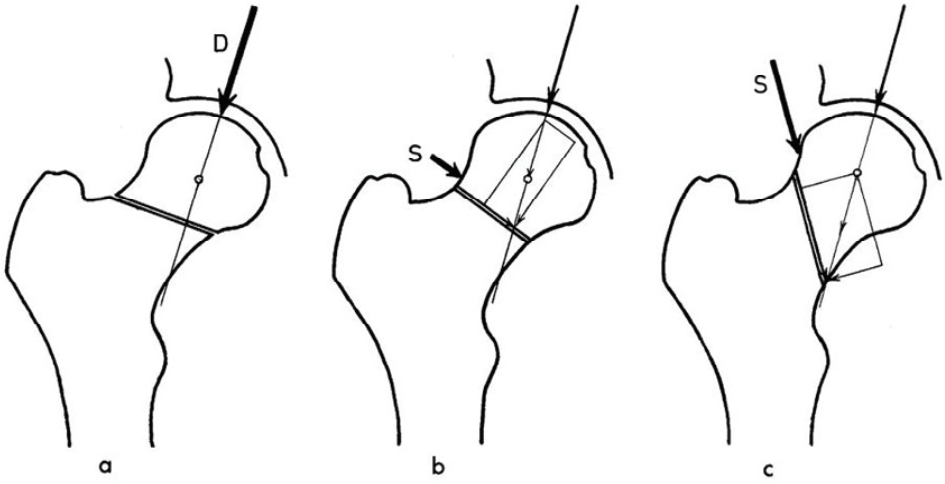


Figure 3. Original illustration by Pauwels 1973 [9]: Classification of femoral neck fractures. Reprinted by courtesy of Springer-Verlag Berlin/Heidelberg.

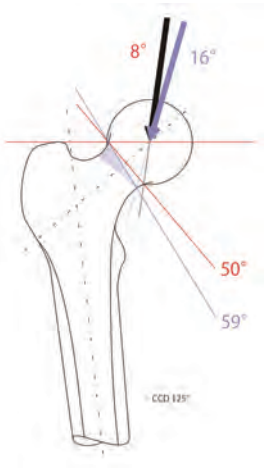


Figure 4. Impact of different vectors on the classification of femoral neck fractures. Relationship demonstrated for the border between Grade II and III fractures.

center of the femoral head. For example, then, fractures in varus hips will reach a higher fracture grade when the angles are smaller than those of fractures in valgus hips. (**Figure 6**)

These issues clearly demonstrate that a rigid classification system using defined angle measurements is not practical. The Pauwels study [1] assigns fractures in valgus an exceptional position. We propose, therefore, a simplified version of Pauwels' Classification, based on mechanical considerations:

Grade I: fractures impacted in valgus,
Grade II: fractures without free torque,
Grade III: fractures with free torque.

4. DISCUSSION

Parker and Dynan [10] declared the Pauwels Classification clinically irrelevant due to the multitude of studies

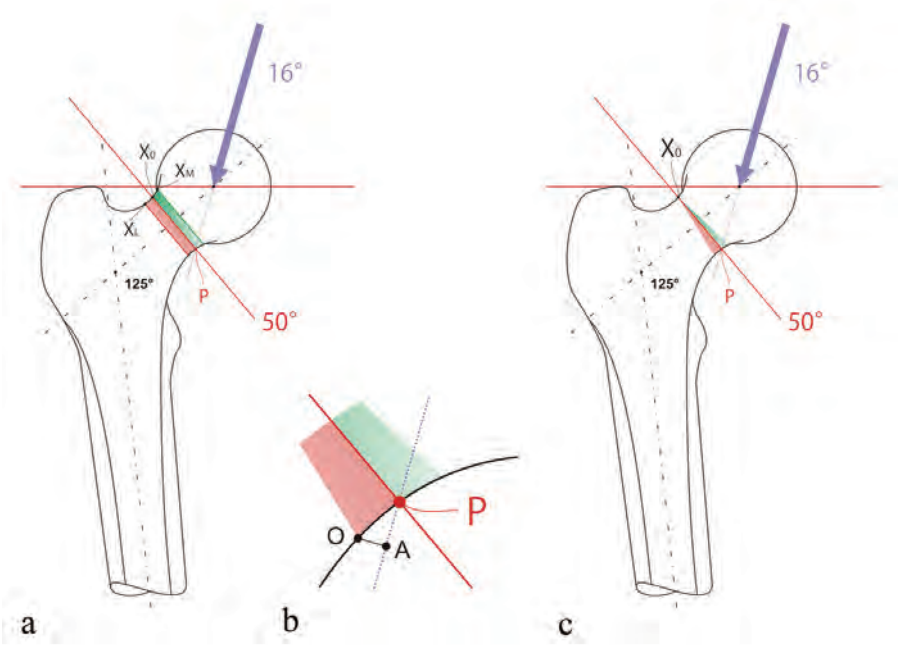


Figure 5. Demonstration of boundary zone for femoral neck fractures Grades II and III at a theoretical fracture starting point X_0 and a fracture angle of 50° to horizontal; (a) More lateral fractures (X_L) lead to the presence of free torque, more medial fractures (X_M) do not. (b) Magnification of (a) and (c): more lateral fractures or more vertical fracture angles at critical point P lead to the presence of free torque (R, OA). (c) More vertical angles (red zone) lead to the presence of free torque, more horizontal fractures (green zone) do not.

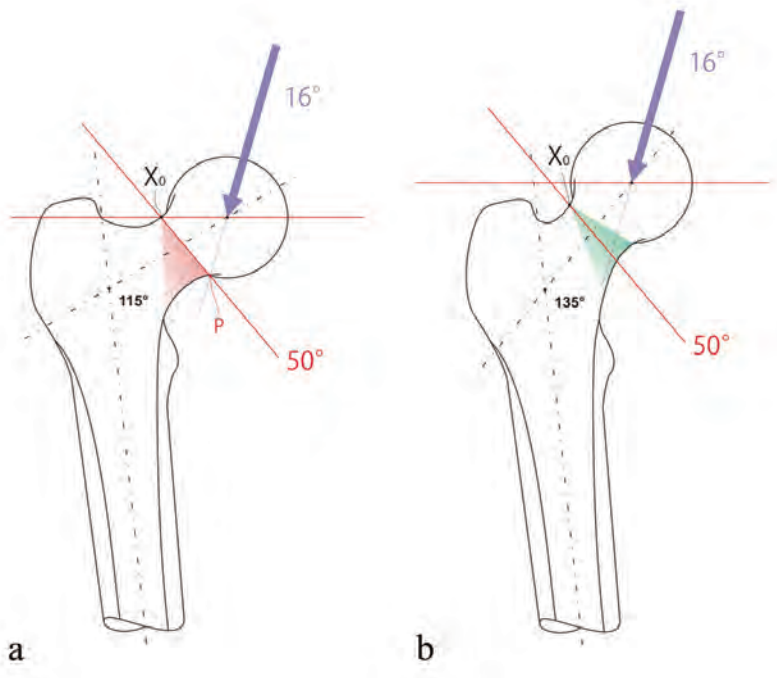


Figure 6. Influence of the CCD-angle on boundary zones for the Pauwels Classification 1973 [9], demonstration of Grades II and III, starting at a theoretical fracture point X_0: (a) More varus CCD-angles lead to the presence of free torque (R, OA) at critical point P in many cases. (b) More valgus CCD-angles seldom lead to the presence of free torque (R, OA).

showing no correlation between Pauwels' angles and non-union. This statement is based most notably on two large studies, from Otremski, *et al*. [3] and Raaymakers and Marti [11], which showed no correlation between fracture angle and secondary displacement and/or non-union in impacted femoral neck fractures. Verheyen *et al*. [12] determined that the critical angle between Grade I and Grade II injuries is impractical, since almost all cases of impacted femoral neck fractures lie in the gray area and it is too difficult to differentiate them.

However, in his original 1935 work, Pauwels [1] emphasized the exceptional nature of impacted femoral neck fractures (**Figure 1**).

The conclusions from individual studies are difficult to compare, but this is due to varying interpretations of the critical angles. The problem does not lie with the Classification itself. For instance, Krastman, *et al*. [13] established a Pauwels Grade II as 30° to 70°, and allocated 71% of 112 cases as Grade II injuries. In contrast, Prinčič, *et al*. [5] identified only 7.1% of 351 cases as Pauwels Grade II injuries. They did not indicate which critical angles they used.

The Pauwels Classification has also been criticized regarding the difficulty to determine angles preoperatively because of rotational error [3,11]. Projection errors are a generalized problem for fracture evaluation, however, and are ubiquitous in all two-dimensionally imaged fractures that are not taken orthograde. Graphic techniques, *i.e*., comparison of the contralateral side, can be of assistance. As well, intraoperative radiographs post-reposition, as suggested by Raaymakers [14,15] can be used to combat this problem.

Another criticism of femoral neck fracture classifications is that they don't consider the second plane. This is true of the Pauwels Classification [1], which doesn't use the lateral plane, as well as the Garden Classification [16], for which in clinical practice it is often neglected. Because of this, and because of variability in interpretations, Frandsen, *et al*. [17] and Zlowodszki, *et al*. [18] suggested that the Garden Classification [16] is no longer useful in clinical practice.

Individual studies have identified correlations between fracture angle and the occurrence of secondary displacement and/or non-union [4-6]. Many authors use the Pauwels [1] as well as the Garden [16] Classifications to evaluate the respective risks of non-union or femoral head necrosis and to determine the best mode of treatment [2,7,13,19,20].

On the one hand, there is the need to predict post-operative course based on measurable pre-operative parameters, and to thereby choose the best method of treatment. On the other hand, these parameters should not be too complicated for regular clinical use, since that can lead to confusion and sub-optimal therapeutic choices. Therefore, the newer classification developed by Caviglia, *et al*. [21], with its six grades of fracture and ten subdivisions, appears too complex for routine clinical use in evaluating femoral neck fractures. The inclination of the fracture continues to be an important consideration to estimate the course of healing and determine the best method of treatment. Fractures impacted in valgus remain a special category. Although in these cases more vertical fractures can be more mechanically favorable, it is important to pay close attention for the presence of dorsal tilting.

At the time of Pauwels' writing, it was important to distinguish non-impacted femoral neck fractures having a moderate fracture angle but no free torque, since this influenced the choice of therapy. However, simple nailing procedures are no longer performed. Instead, more reliable techniques such as minimally-invasive fixation with two to three gliding or pressure screws or larger implants such as a dynamic hip screw (DHS) are used.
We believe that the indications for internal fixation versus hemi-arthroplasty should be based not only on the risk of femoral head necrosis (*i.e.*, using the Garden Classification), but also on the mechanical constellation of injury. This is generally accepted for the treatment of femoral neck non-unions. In the area of femoral neck fractures, however, no classifications using pure mechanical analysis can be implemented, either because they are too complex, or because they have not been comparably interpreted by the literature.

For all of these reasons, we propose the following simplified version of the Pauwels Classification system for femoral neck fractures based on mechanical considerations. Grade I fractures are those impacted in valgus, Grade II are fractures without free torque, and Grade III are fractures with free torque. To differentiate between Grade II and III fractures, the dynamic resulting vector R (16°) should be projected from the center of the femoral head onto the level of fracture. If the vector intersects the fracture surface, the injury is a more mechanically favorable Grade II injury without free torque. If, however, the vector projection intersects medial to the fracture plane, free torque is present and the injury is an unfavorable Grade III type, with increased risks of secondary displacement, femoral head necrosis, and/or non-union.

This extrapolation, strictly interpreted, is a simplification of a complex mechanical relationship. The angle of incidence of the resulting vector R is dependent on the leverage variables (CCD angle, femoral neck length, etc.) and multidimensional. However, overall it is better defined and more reproducible as a measure of the presence of free torque than the fracture angle from horizontal.

This classification can be used for borderline cases, definitively, in the operating room. Moreover, an enhancement of the mechanical constellation can be obtained using valgisation of the fracture.

Whether this modified classification is sufficient to predict the development of mal/non-union and/or femoral head necrosis, and therefore the need for fixation versus hemi-arthroplasty, remains to be seen. Further studies will be necessary to evaluate its use in the clinical setting.

REFERENCES

[1] Pauwels, F. (1935) Der schenkelhalsbruch ein mechanisches problem. *Z Orthop Ihre Grenzgeb*, **63**, 1-135.

[2] Angly, B. and Ochsner, P.E. (2004) Sind Gleitsc- hrauben bei medialen Schenkelhalsfrakturen älterer Men-schen eine taugliche Alternative zu Kopfprot-hesen? *Akt Taumatol*, **34**, 123-127.

[3] Otremski, I., Katz, A., Dekel, S., Salama, R. and Newman R.J. (1990) Natural history of impacted subcapital femoral fractures and its relevance to treatment options. *Injury*, **21(6)**, 379-381.

[4] Penschuk, C., Zilch, H. and Brenner, M. (1982) Langzeitergebnisse der druckosteosynthese mit drei ao-spongiosaschrauben bei schenkelhalsfrakturen. *Unfallchirurgie*, **8**, 33-40.

[5] Princic, J., Zupancic, F. and Cimerman, M. (1993) Late results of 351 femoral neck fractures (1986-1987 treatment period). *Unfallchirurg*, **96(9)**, 468-472.

[6] Roth, W. (1988) Complications following osteosynthesis management of femoral neck fractures with spongiosa locking screws. *Aktuelle Traumatol*, **18(1)**, 21-32.

[7] Scharf, W., Wagner, M. and Trojan, E. (1985) Differentiated procedure in fractures of the coxal femur end in adults. *Aktuelle Traumatol*, **15(5)**, 10-16.

[8] Bartonicek, J. (2001) Pauwels' classification of femoral neck fractures: Correct interpretation of the original. *Journal of Orthopaedic Trauma*, **15(5)**, 358-360.

[9] Pauwels, F. (1973) Atlas zur Biomechanik der gesunden und kranken Hüfte. Heidelberg, Berlin, Springer, New York.

[10] Parker M.J. and Dynan, Y. (1998) Is pauwels classification still valid? *Injury*, **29(7)**, 521-523.

[11] Raaymakers, E.L.F.B. and Marti, R.K. (1991) Non-operative treatment of impacted femoral neck fractures. A prospective study of 170 cases. *Journal of Bone & Joint Surgery* (*Br*), **73(6)**, 950-954.

[12] Verheyen, C.C.P.M., Smulders, T. C. and van Walsum, A. D. (2005) High secondary displacement rate in the conservative treatment of impacted femoral neck fractures in 105 patients. *Arch Orthop Trauma Surg*, **125**, 166-168.

[13] Krastman, P., van den Bent, R. P., Krijnen, P. and Schipper, I. B. (2006) Two cannulated hip screws for femoral neck fractures: treatment of choice or asking for trouble? *Arch Orthop Trauma Surg*, **126(5)**, 297-303.

[14] Raaymakers, E.L.F.B. (2006) Fractures of the femoral neck: A review and personal statement. *Acta Chir Orthop Traumatol Cech*, **73(1)**, 45-59.

[15] Raaymakers, E.L.F.B. and Schafroth, M. (2002) Medial femoral neck fracture. Controversies in treatment. *Unfallchirurg*, **105(2)**, 178-186.

[16] Garden, R.S. (1961) Low-angle fixation in fractures of the femoral neck. *Journal of Bone & Joint Surgery* (*Br*), **43**, 647-663.

[17] Frandsen, P.A., Andersen, E., Madsen, F. and Skjodt, T. (1988) Garden's classification of femoral neck fractures. An assessment of inter-observer variation. *Journal of Bone & Joint Surgery* (*Br*), **70**, 588-590.

[18] Zlowodzki, M., Bhandari, M., Keel, M., Hanson, B. P. and Schemitsch, E. (2005) Perception of Garden's classification for femoral neck fractures: An international survey of 298 orthopaedic trauma surgeons. *Arch Orthop Trauma Surg*, **125(7)**, 503-505.

[19] Füchtmeier, B., Hente, R. Maghsudi, M. and Nerlich, M. (2001) Repositioning femoral neck fracture in younger patients, valgus or anatomic reposition? *Unfallchirurg*, **104(11)**, 1055-1060.

[20] van Meeteren, M.C., de Vries, L.S., Hammacher, E.R. and van der Werken, C. (1996) Reasonably good results of Hansson's pins in elderly patients with a medial femoral neck fracture. *Ned Tijdschr Geneeskd*, **140**, 1552-1557.

[21] Caviglia, H.A., Osorio, P.Q. and Comando, D. (2002) Classification and diagnosis of intracapsular fractures of the proximal femur. *Clinical Orthopaedics and Related Research*, **399**, 17-27.

[22] Gerber, P. and Wicki, O. (1995) Stadien und Einteilungen in der Medizin. Stuttgart, Thieme Verlag, New York.

[23] Bout, C.A., Cannegieter, D.M. and Juttmann, J.W. (1997) Percutaneous cannulated screw fixation of femoral neck fractures: The three point principle. *Injury*, **28(2)**, 135-139.

[24] Bray, T. J. (1997) Femoral neck fracture fixation, Clinical decision making. *Clinical Orthopaedics and Related Research*, **339**, 20-31.

[25] Krämer, K.L., Stock, M. and Winter, M. (1997) Klinikleitfaden Orthopädie. 2th Edition, Gustav Fischer Verlag, Stuttgart.

[26] Blauvelt, T.C. and Nelson, F.R.T. (1998) A manual of orthopaedic terminology. 6th Edition, C.V. Mosby Company, St. Louis.

[27] Hasse, F.M. and Müller, L.P. (2002) Klinikleitfaden Chirurgie. 3th Edition, Urban und Fischer Verlag, München.

[28] Brown, J.T. and Abrami, G. (1964) Transcervical femoral fracture: A review of 195 patients treated by sliding nail-plate fixation. *Journal of Bone & Joint Surgery* (*Br*), **46**, 648-663.

[29] Kuner, K. and Schlosser, V. (1995) Traumatologie, 5th Edition, Thieme Verlag, Stuttgart, New York.

Call for Papers

Journal of Biomedical Science and Engineering (JBiSE)

www.scirp.org/journal/jbise

JBiSE, an international journal, publishes research and review articles in all important aspects of biology, medicine, engineering, and their intersection. Both experimental and theoretical papers are acceptable provided they report important findings, novel insights, or useful techniques in these areas. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed in hard copy.

Subject Coverage

- Bioelectrical and neural engineering
- Bioinformatics and Computational Biology
- Biomedical modeling
- Biomedical imaging, image processing and visualization
- Clinical engineering, wearable and real-time health monitoring systems
- Biomechanics and biotransport
- Software, tools and application in medical engineering
- Biomaterials
- Physiological signal processing
- Biomedical devices, sensors, artificial organs and nano technologies
- NMR/CT/ECG technologies and electromagnetic field simulation
- Structure-based drug design

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

www. scirp. org/journal/jbise

Email: jbise@scirp.org

Editor-in-Chief

Editorial Board

ISSN 1937-6871 (Print), 1937-688X (Online)

Journal of Biomedical Science and Engineering, 2010, 3, 543-643.
Published Online June 2010 in SciRes (http://www.scirp.org/journal/jbise).

TABLE OF CONTENTS

Volume 3 Number 5 June 2010

