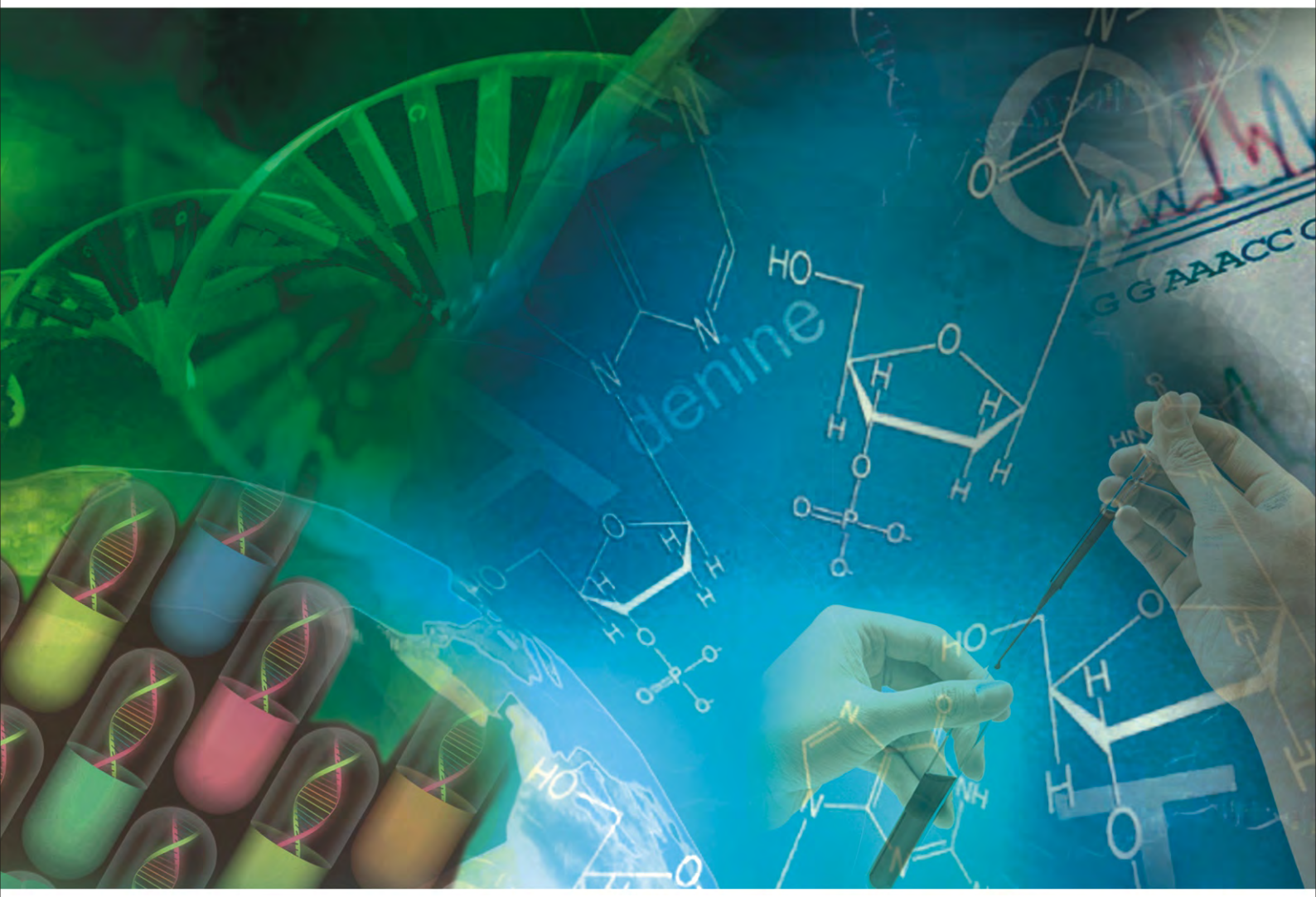
Scientific Research

JBiSE

ISSN: 1937-6871

Volume 2 Number 7 November 2009

Journal of Biomedical Science and Engineering



www.scirp.org/journal/jbise

Journal Editorial Board

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

Journal of Biomedical Science and Engineering (JBiSE)
Published Online November 2009 in SciRes (http://www.scirp.org/journal/jbise).

TABLE OF CONTENTS

Volume 2, Number 7, November 2009

Journal of Biomedical Science and Engineering (JBiSE)

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.
E-mail: service@scirp.org

Subscription Rates: Volume 2 2009

Printed: $50 per copy.

Electronic: freely available at www.scirp.org.

To subscribe, please contact Journals Subscriptions Department at service@scirp.org.

Sample Copies: If you are interested in obtaining a free sample copy, please contact Scientific Research Publishing, Inc at service@scirp.org.

SERVICES

Advertisements

Contact the Advertisement Sales Department at service@scirp.org.

Reprints (a minimum of 100 copies per order)

Contact the Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: service@scirp.org

PRODUCTION INFORMATION

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

J. Biomedical Science and Engineering, 2009, 2, 471-479 **JBiSE**
doi: 10.4236/jbise.2009.27068 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Nadi Yantra: a robust system design to capture the signals from the radial artery for assessment of the autonomic nervous system non-invasively*

Abhinav[1], Meghna Sareen[1], Mahendra Kumar[1], Jayashree Santhosh[2], Ashok Salhan[3], Sneh Anand[1]

[1]Centre for Biomedical Engineering, Indian Institute of Technology, Delhi, India; [2]Computer Services Centre, Indian Institute of Technology, Delhi, India; [3]Defence Institute of Physiology & Allied Science, Defence Research and Development Organisation, Delhi, India.
Email: [1]abhinav_cbme_iitd@hotmail.com; [2]meghnasareen@gmail.com

Received 15 June 2009; revised 30 June 2009; accepted 1 July 2009.

ABSTRACT

Ayurvedic and other alternative medical practitioners throughout the world have been using pulse diagnosis to detect disease and the organ at distress by feeling the palpations at three close yet precise positions of the radial artery. This paper presents a robust electro-mechanical system, 'Nadi Yantra' which uses piezoelectric based pressure sensors to capture the signals from the radial artery. Morphology of the waveforms obtained from our system concurs with standard physiological arterial signals. Reproducibility and stability of the system has been verified. Signal processing techniques were applied to obtain features such as amplitude, power spectral density, bandpower and spectral centroid to reflect variations in signals from the three channels. Further, wavelet based techniques were used to process the pressure signals and percussion peaks were identified. The interval between the percussion peaks was used to calculate Heart Rate Varibility (HRV), a useful tool for assessing the status of the autonomic nervous system of the human body non-invasively. Time domain indices were calculated from direct measurement of peak-peak (PP) intervals and from differences between the PP intervals. Frequency domain indices such as very low frequency (VLF) power, low frequency (LF) power, high frequency (HF) power, LF/HF ratio were also calculated. Thereafter, nonlinear Poincare analysis was carried out. A map of consecutive PP intervals was fitted to an ellipse using least squares method. Results from 7 datasets are depicted in this paper. A novel pressure pulse recording instrument is developed for the objective assessment of the ancient science of pulse diagnosis. The features calculated using multi resolution wavelet analysis show potential in the evaluation of the autonomic nervous system of the human body.

Keywords: Radial Artery; Pulse Diagnosis; Power Spectral Density; Spectral Centroid; Multi Resolution Wavelet; Autonomic Nervous System; Heart Rate Variability; Time Domain; Frequency Domain; Poincare

1. INTRODUCTION

In ancient literatures of the Ayurveda, Chinese, Unani, and Greek medicine, pulse based diagnosis has its own unparalleled importance. The organ under distress is zeroed down by feeling the palpations from the three fingers (index, middle and ring) placed on the radial artery (**Figure 1**). These pulsations dictate the physiological status of the entire human body [1]. This is a tedious and highly subjective process and takes years of practice to master this art [2].

Pulse has been ubiquitously accepted by modern clinicians as well. They examine the pulse using the method of trisection i.e. apply pressure until the pulse is maximal, and then vary pressure while concentrating on the phases of the pulse. The arterial pulse variants (for example *pulsus alternans*, *bisferiens pulse*, *bigeminal pulse*) are used in detecting cardiac disorders. However, alternative medicine practitioners carefully examine pulses at different depths, each connected with a specific part of the body and each believed to register even the slightest physiological based change.

If there can be a device that can give an objective assessment of the science of pulse diagnosis, it will assist disease diagnosis noninvasively. It will be used by alternative medicine practitioners as well as modern clinicians.

*Nadi Yantra has been applied for patent (pending approval); Nadi stands for Pulse and Yantra means Instrument.

Analysis of the R-R time series has been commonly used in electrocardiographic (ECG) signals. ECG signals are electrical signals of the heart and pressure signals from the radial artery are mechanical signals. However, pressure signals also show significant complexes just like the QRS of ECG waves. Therefore, a similar analysis on percussion peak-peak (PP) time series can be carried out to provide information regarding the status of the autonomic nervous system noninvasively.

In this paper we discuss our device *Nadi Yantra*, the three sensors of which simulate the human fingers. In Section 2 the instrument has been explained in detail. Section 3 describes the wavelet based signal processing for calculation of heart rate variability and features for evaluation of the autonomic nervous system noninvasively. Section 4 describes experiments conducted. Inferences are drawn in Section 5.

2. THE INSTRUMENT-NADI YANTRA

There is a need for the development of a quantitative system for pulse diagnosis [3]. Investigations have been attempted globally to develop a system that replicates the human **three-finger** method of pulse based diagnosis [4,5]. In previous attempts

1) Signals have been captured for a very small span of time (1-2 minutes).
2) External pressure applied over the sensors varies while recording.
3) Motion artifacts become a reasonable consideration when the recordings are for a longer duration of time.

Advances over the earlier systems are that *Nadi Yantra* allows recording for hours by an automated external pressure on the three positions thereby completely removing the potential for errors incurred when pressure is applied (**Figures 2, 3, 4**). The locking mechanism significantly resists the motion artifacts as well.

Figure 1. A practitioner evaluating the patient's pulse.

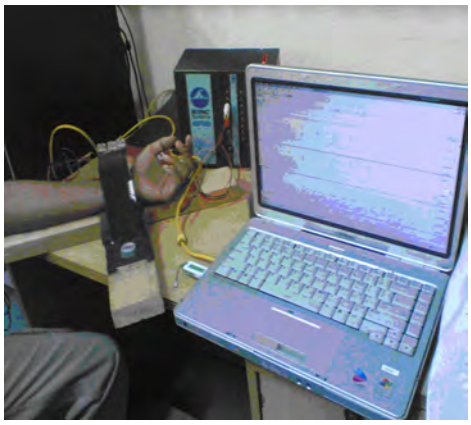


Figure 2. Recording the Pulse using *Nadi Yantra*.

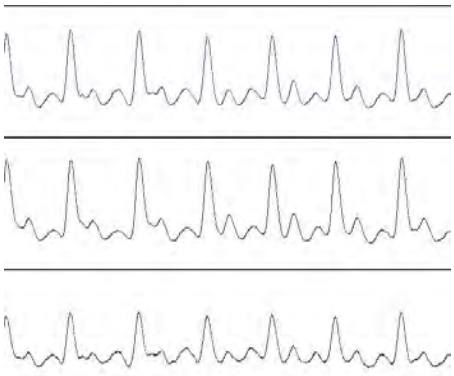


Figure 3. Signals (zoomed) as observed in the three channels.

In the mechanical design, the system has three finger like projections whose positions can be adjusted at the tip region to find out the best locations to capture the signal. Springs attached to them help in damping thus simulating the natural damping present due to muscles in the tip region of the practitioner's fingers (**Figure 5**).

Once the three best positions are found, they are locked with another hard spring fitting. This lock resists the motion artifacts as well. Discrete increments in pressure are possible by changing the lock's position towards the slant side. (**Figure 6**).

In the electrical design, we used three identical piezo film based sensors to capture the waveform. The raw signal was filtered, amplified, and transferred to the computer using BioPac 150^{TM} (Signal Acquisition System) operated at a sampling rate of 1000 samples per second.

It is found that the signal captured using Nadi Yantra corresponds well with standard pulse from the radial artery (**Figure 7**, **Figure 8**). A standard pulse from the radial artery comprises of the following waves: [5]

a) Percussion Wave
b) Tidal Wave
c) Dicrotic Wave
d) Valley

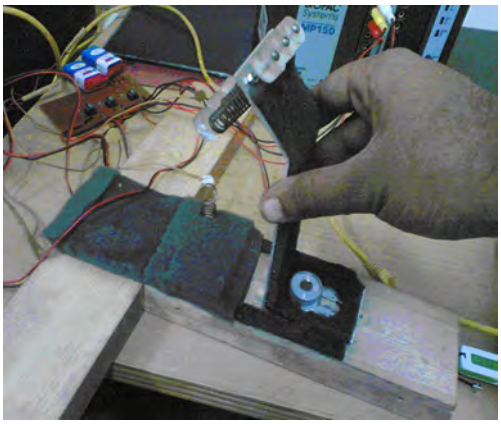


Figure 4. Flexibility of the system.

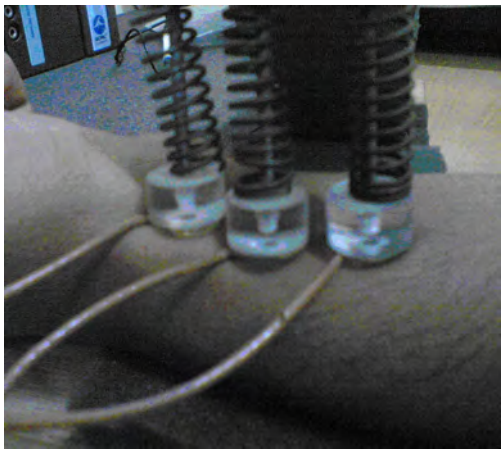


Figure 5. Simulation of fingers with the pressure sensors.

Figure 6. Lock mechanism for discrete pressure increment.

3. ANALYSIS OF RADIAL PULSE SIGNAL

3.1. Pressure Wave Decomposition Using Wavelets

A radial pulse signal feature extraction system using wavelet-based multi-resolution analysis was developed and evaluated.

The choice of the wavelet function depends on how closely the scaling function matches the shape of the original signal [6,7]. Daubechies 9 (D9) of Daubechies is similar in shape to the radial pulse signal complex and was used in the decomposition.

Figure 9 shows the details of a signal captured by Nadi Yantra. The original signal is shown at the top of the plot. Below the signal, the details for seven wavelet scales are shown which are scaled for better illustration. Adding together all these details plus the signal approximation A7 returns the original signal.

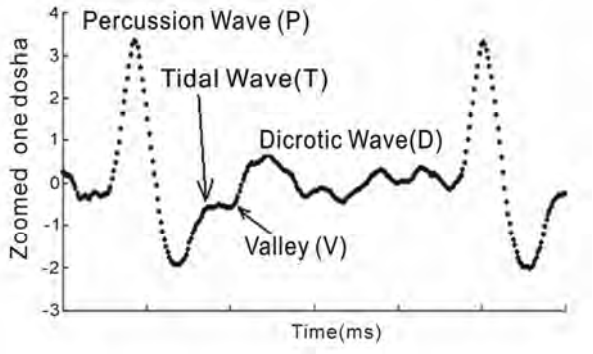


Figure 7. A standard signal from the radial artery.

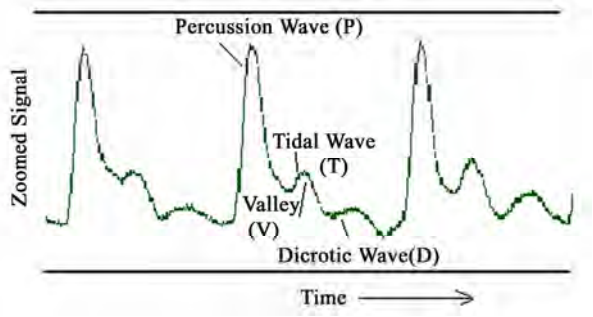


Figure 8. Zoomed signal as acquired by *Nadi Yantra*. (Note: No digital filters have been used to capture these raw signals)

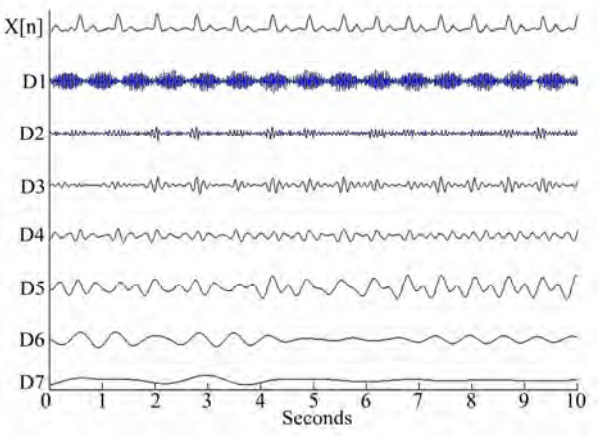


Figure 9. Wavelet decomposition of a typical radial pulse signal.

A few points can be noted from the plot.

1) The high frequency noise content is captured at the smallest scales, namely D1, D2 and D3.

2) Scales D4, D5 and D6 contain most of the information that depicts the radial pulse signal.

3) Also, scales D7 (and above) correspond to very low frequencies which do not contribute to the complex. They account for the DC component of the signal and baseline wander resulting from motion etc.

3.2. Preprocessing of Signal

Low frequencies (which constitute baseline wander) appear at high scales (details D7 and above). Removal of such details corrects baseline wander. High frequency burst of noise is captured at the smallest scales (which correspond to high frequencies), namely D1, D2 and D3. Removal of such details corrects noise. Additionally, a notch filter was applied to remove the 60 Hz power line interference. The result is shown in **Figure 10**.

A frequency domain analysis of the denoised pressure signal was carried out. A Fourier Transform of the pressure signal from the radial artery is shown below. It shows the significant frequencies present in the signal. The signal contains low frequencies from 0-10 Hz (**Figure 11**).

3.3. Percussion Peak Detection

It was observed that the details D3, D4 and D5 are the most significant and contain the information that represent the radial pulse signal complex, including percussion wave. These details are devoid of high frequency noise and low frequency baseline wander as well as mean DC component. The signal was appropriately squared to enhance percussion peaks which ensured better detection. A thresholding technique was used to detect the percussion peaks. Thereafter, the peak-peak (PP) time series was obtained.

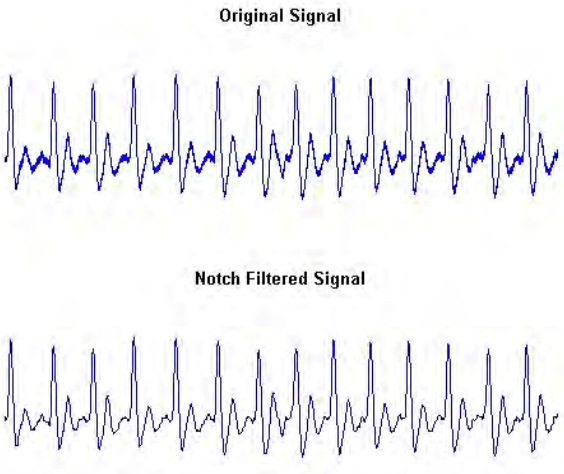


Figure 10. Clean denoised signal.

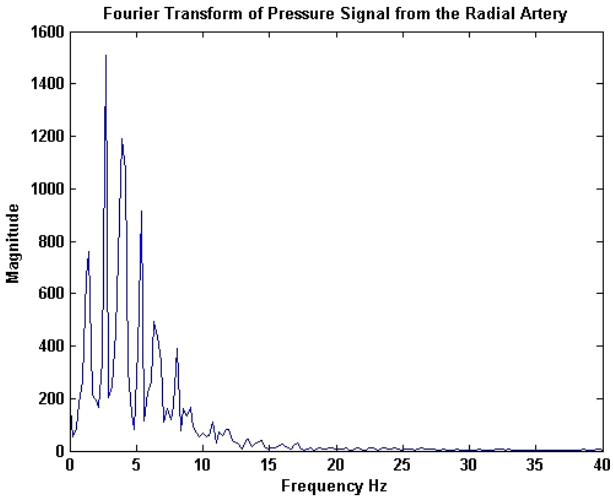


Figure 11. Fourier transform of denoised pressure signal.

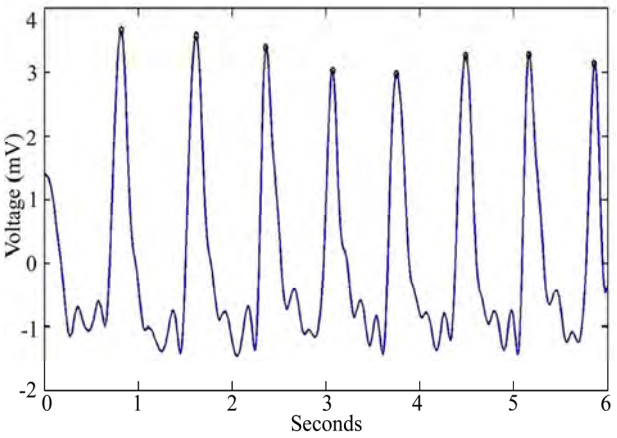


Figure 12. Peak detection in pressure signals using wavelets.

3.4. Feature Extraction from Percussion Peak-Peak (PP) Intervals

It is well known that perturbations to autonomic activity, such as respiratory sinus arrhythmia and vasomotor oscillations cause corresponding fluctuations in heart rate. The alterations of the heart beat known as heart rate variability (HRV) is a useful tool for assessing the status of the Autonomic Nervous System (ANS) non-invasively [8,9]. Our approach for HRV calculation is based on Percussion complex detection. HRV signal is computed from the time difference between two consecutive percussion complexes known as the PP time series. A HRV signal is shown in **Figure 13** below.

Changes in heart rate occur as a result of the autonomous nervous system's actions through the parasympathetic (P) and the sympathetic (S) pathways which have opposite influences. The S stimulation leads to an increase in heart rate and the P stimulation does the opposite. These different actions result in fluctuations in the

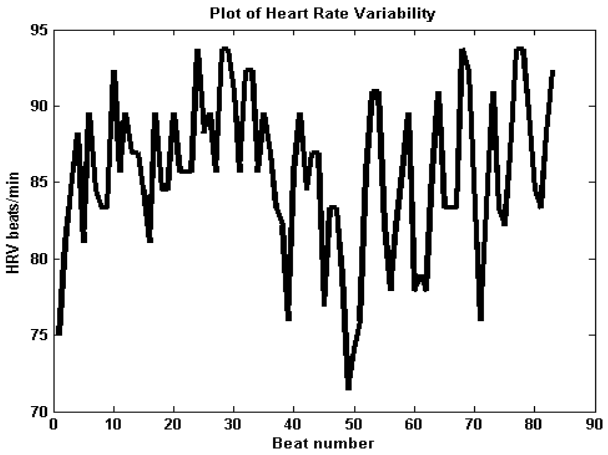


Figure 13. HRV signal.

heart rate, the best known being the sinus respiratory arrhythmia modulated by the P, and the ones related with the baroreflex action modulated by the S.

Time domain indices were derived from direct measurements of the PP intervals or from differences between the PP intervals. The former include mean and standard deviation (SDNN i.e. standard deviation of the normal to normal interval i.e. square root of variance), SDANN (standard deviation of the average NN interval calculated over short periods). The most commonly used measures derived from interval differences include RMSSD, the square root of the mean squared differences of successive PP intervals, NN50, the number of interval differences of successive NN intervals greater than 50 milliseconds and pNN50, the proportion derived by dividing NN50 by the total number of NN intervals [10]. These provide information on the short-term and long-term variability of the P-P time series.

Frequency domain indices provide information on both total variability as well as distribution as a function of frequency. The spectral analysis of this HRV signal allows to quantitatively distinguish between the different activities of the ANS such as very low frequency component (*VLF*: *0-0.04 Hz*), a low frequency component (*LF*: *0.04-0.15 Hz*) and a high frequency component (*HF*: *0.15-0.4 Hz*). HF power is supposedly a pure measure of parasympathetic activity and represents momentary respiratory influences on heart rate or respiratory sinus arrhythmia, and LF power is reflective of sympathetic modulation and parasympathetic tone. It derives from short term regulation of blood pressure.

A total of 40 signals were analyzed and results from 7 datasets are shown below in **Table 1**.

There is some evidence for the involvement of nonlinear phenomena in the genesis of HRV. It is conceived that assessment of HRV with nonlinear measures may supply information different from and additional to that derived through linear measures. A non-linear Poincare analysis was carried out for a representative P-P time series where a Poincare plot is a scatter-plot of the current P-P interval against the preceding P-P interval. The PP_i is plotted versus PP_{i-1} and the plot shown below is obtained. An ellipse was fitted to it using least squares method and values of the centre of the ellipse, major and minor axis, angle with the x axis and equation of the ellipse obtained. The plot provides summary information as well as detailed beat-to-beat information on the behavior of the heart. A distinct advantage of Poincare plots is their ability to identify beat-to-beat cycles and patterns in data that are difficult to identify with spectral analysis [11]. The Poincare plot (**Figure 14**) allows the HRV researcher to measure the variability of heart rate from different points of view such as long-term variability, overall variability, variability on basal heart rate, variability on accelerated heart rate, variability on decelerated heart rate as well as the sympathetic-parasympathetic balance.

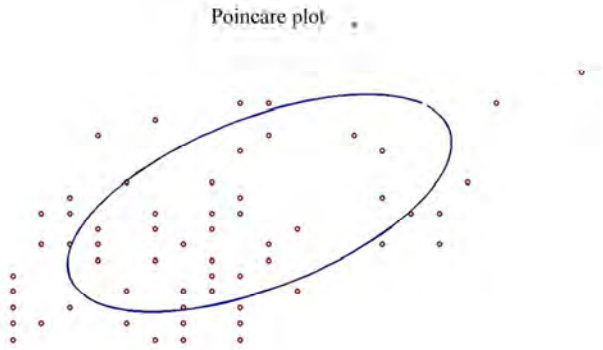


Figure 14. Poincare analysis.

Table 1. Features extracted from P-P time series.

VLF (0-0.04 Hz) $*10^5$	3.89	3.45	3.82	3.67	4.09	4.09	4.03
LF (0.04-0.15 Hz) $*10^4$	1.66	1.38	2.02	1.55	1.80	1.81	1.91
HF (0.15-0.4 Hz) $*10^3$	5.47	3.59	6.90	4.71	5.71	5.53	6.93
LF/HF	3.03	3.84	2.93	3.30	3.15	3.26	2.76
Mean	70.4	61.2	68.6	65.8	73.6	73.6	72.7
SDNN	4.40	2.61	9.24	3.21	3.73	3.72	7.07
RMSSD	4.51	2.72	9.63	3.51	3.97	3.97	7.38
SDNN/RMSSD	0.97	0.96	0.96	0.91	0.94	0.94	0.96
NN50	69	84	73	79	72	72	72
pNN50 (%)	83.1	86.6	84.8	87.7	90	90	88.8

X_c = 72.68, Y_c = 72.65, A = 8.55, B = 4.54, Phi = 0.81 where X_c and Y_c are the x and y axis centre of the ellipse, A and B are the major and minor axis of the ellipse respectively, Phi is the radian angle of the major axis with respect to the x axis. The equation of the ellipse is $0.6x^2$-0.6xy+$0.5y^2$-37.4x-33.4y+2554.7

4. EXPERIMENTS

4.1. Description of Experiment

We chose 5 healthy volunteers to carry out a whole day analysis. The dynamics of the physiological signals were examined for pre-lunch and post-lunch states.

The control used in our analysis was the 20 sets of signals captured from the same subject over a period of time prior to lunch. The post lunch signals were compared with the control signals and the results are shown below (**Figures 15–17**). Signals were recorded at 1245 hrs, 1315 hrs, 1355 hrs and 1705 hrs where lunch was administered at 1420 hrs.

It can be observed that the amplitude of the 1st channel rises steadily prior to lunch and then falls post lunch. It can also be seen that the amplitude of the signals from the 2nd and 3rd channels rises post lunch (**Figures 15–20**).

4.2. Signal Processing and Results

Power spectral density of the signals was calculated and plotted in **Figures 21–23**.

It can be seen from the cumulative PSD plots that the signals from the three channels have the same frequency components in pre-lunch and post-lunch signals. The frequency peaks are more in number in the signals before lunch as compared to that after lunch. The magnitude

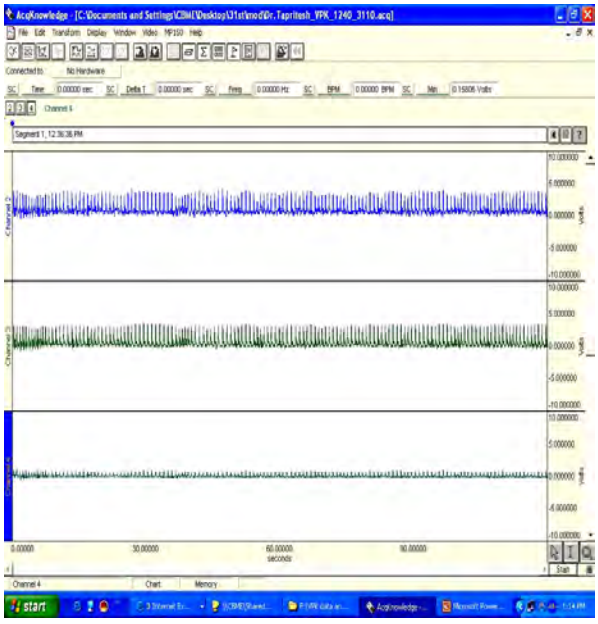


Figure 15. Signals captured at 1245 hrs.

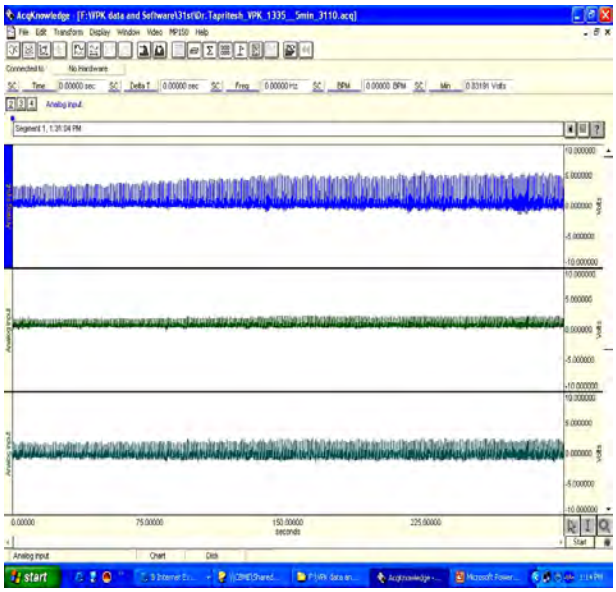


Figure 16. Signals captured at 1315 hrs.

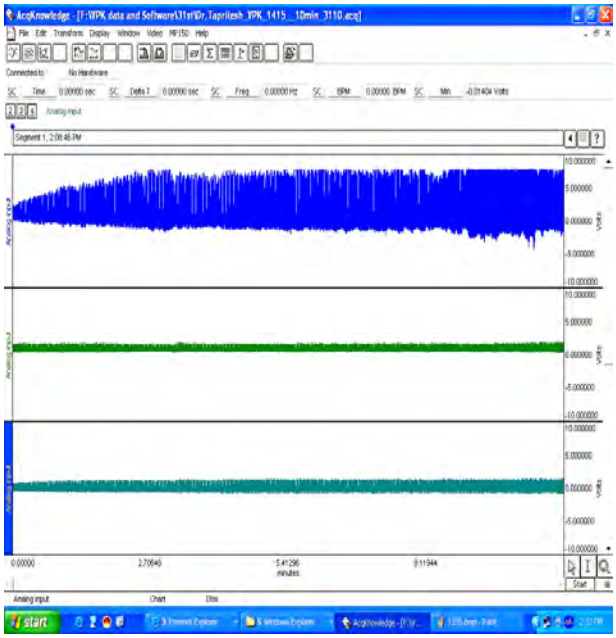


Figure 17. Signals captured at 1355 hrs.

of power in the 1st channel increases as appetite increases from 1315 to 1355 hrs and then falls post lunch at 1755 hrs. Thus, the dynamic properties of the three signals (magnitude of power and frequency information) change before lunch and after lunch. Further, a 30 minute post-lunch signal (1705 hrs) was divided into 6 segments of 5 minutes each and PSD for every segment was calculated. It can be inferred that the magnitude of power in the 2nd channel decreases and that of the 3rd channel increases over time.

A post-lunch 30 minute signal was divided into 15 segments of 2 minutes each and bandpower calculated

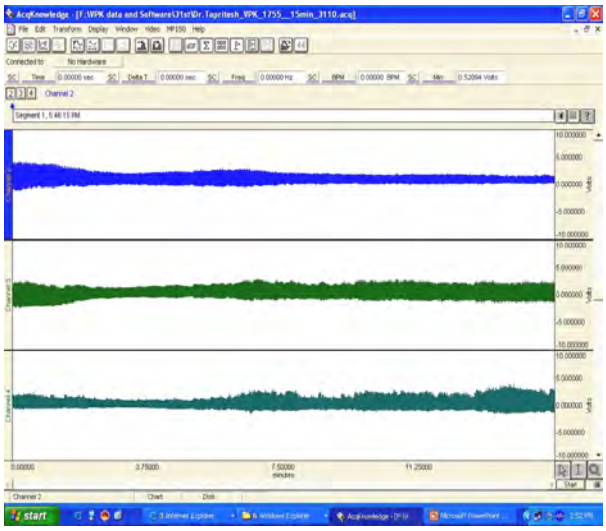


Figure 18. Signals captured at 1705 hrs.

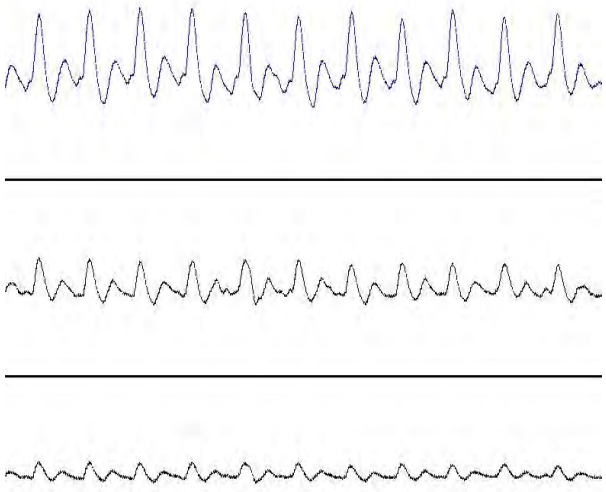


Figure 19. Zoomed view of a portion of signals from **Figure 10**.

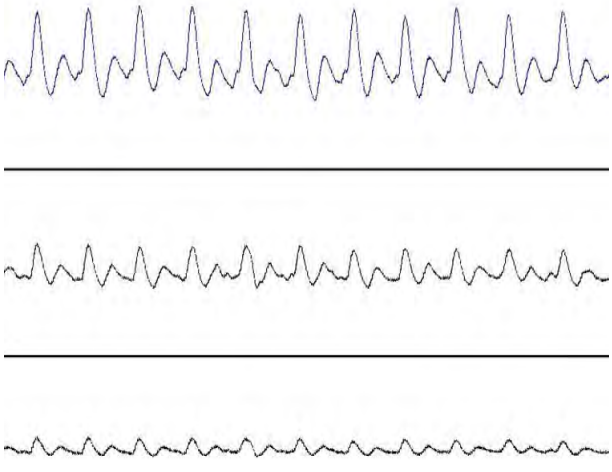


Figure 20. Zoomed view of a portion of signals from **Figure 12**.

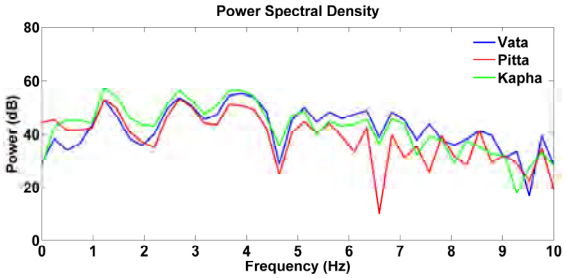


Figure 21. PSD of the signal acquired before lunch at 1315 hrs.

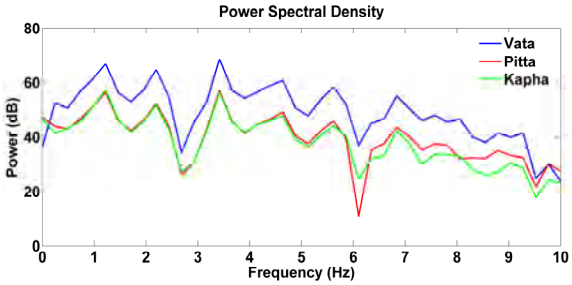


Figure 22. PSD of the signal acquired before lunch at 1355 hrs.

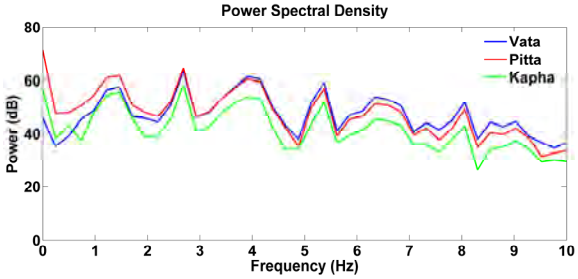


Figure 23. PSD of the signal acquired after lunch at 1705 hrs.

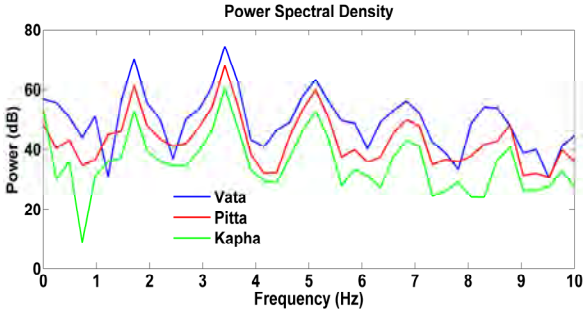


Figure 24. PSD of a post-lunch signal divided into 6 segments.

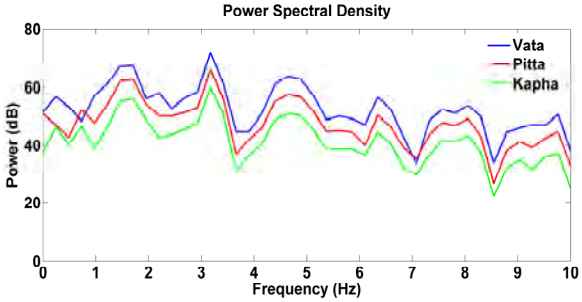


Figure 25. PSD of a post-lunch signal divided into 6 segments.

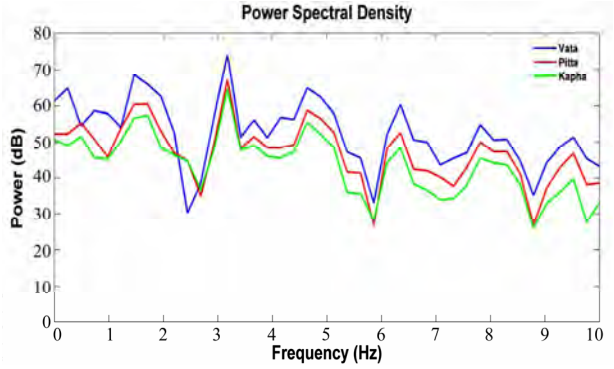


Figure 26. PSD of a post-lunch signal divided into 6 segments.

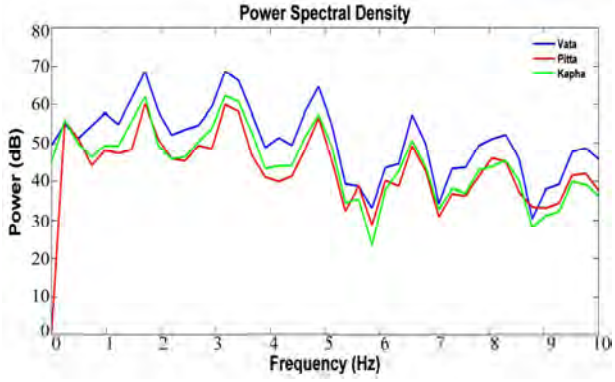


Figure 27. PSD of a post-lunch signal divided into 6 segments.

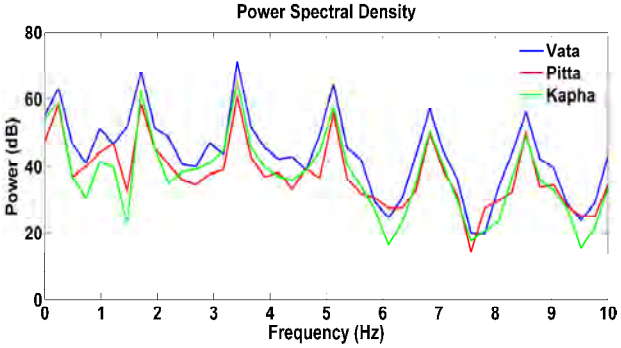


Figure 28. PSD of a post-lunch signal divided into 6 segments.

for each segment. The calculated bandpower is the area under the curve of a PSD in the bandlimits 0.5-10 Hz, appropriate for arterial pulse signals. It can be seen from **Figure 30** that bandpower for the 2nd channel falls after lunch whereas bandpower for the 3rd channel increases. Bandpower for the 1st channel remains predominant.

Spectral centroid frequency commonly referred to as median frequency was also calculated. It can be seen from **Figure 31** that the centroid frequency of the 2nd channel is higher than the 1st and the 3rd channels.

Thus, it is seen that dynamics of the radial pulse signal change on application of stimulus for example administration of lunch. These dynamics can be monitored using the device developed.

5. CONCLUSIONS

There is enough evidence in ancient literature that there is not a single disease in the human body which cannot be diagnosed by examining the pulse. However, ancient medical practitioners had to totally rely upon years of clinical experience in order to come to any conclusive diagnosis. Clinicians today have limited examination of the pulse to its rate, rhythm and volume by virtue of which they hardly come to a concrete diagnosis. If there could be a system by which the radial pulse could be critically examined, it could be one of the most useful tools in the field of non-invasive diagnosis of disease. In this paper, such a system known as Nadi Yantra has been developed. Further, wavelet based techniques were used to decompose the pressure signal from the radial artery. Multi-resolution wavelet analysis was used to detect the percussion peaks and the P-P time series was obtained. A HRV plot depicting function of the ANS was obtained

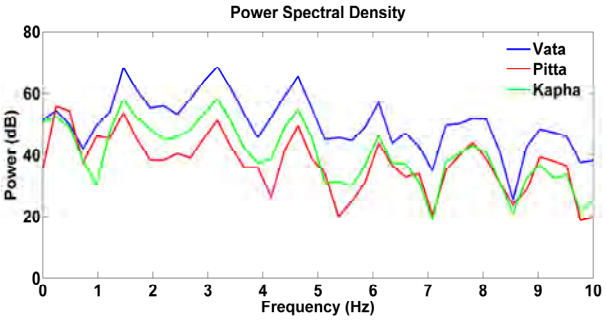


Figure 29. PSD of a post-lunch signal divided into 6 segments.

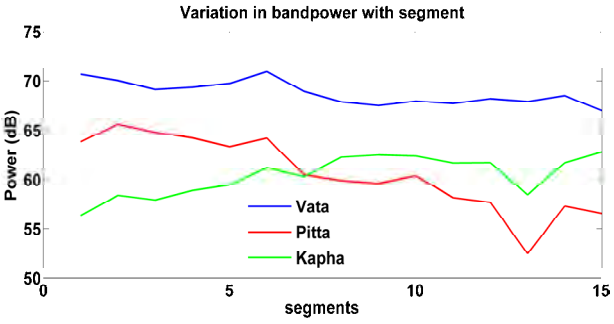


Figure 30. Variation in band power with segment.

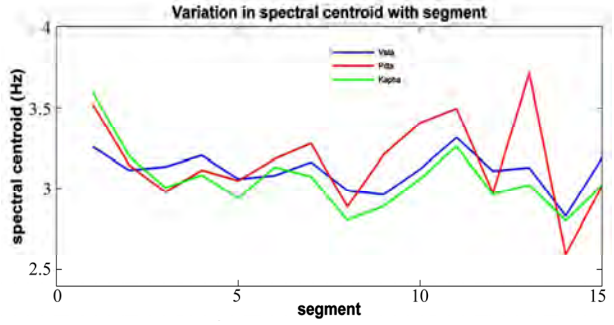


Figure 31. Variation in spectral centroid with segment.

from the P-P time series. Time domain and Spectral analysis of the P-P series gave significant features. Non-linear Poincare analysis was carried out to obtain a relationship between consecutive P-P intervals. These features hold significance in the study of short term and long term variability of the PP time series. Wavelet based method proved most effective because of its ability to filter out noise and low frequency components and retain relevant detail. The wavelet based detector has advantages of time-scale analysis as well as frequency analysis. This approach to capture the peaks from the radial artery and further, extraction of features from the P-P series is a promising tool for studying the diagnostic application of Nadi Yantra in the assessment of the autonomic nervous system. The experiments described in the paper show that Nadi Yantra has the potential to objectively measure and display the physiological changes occurring in the human body.

6. ACKNOWLEDGEMENTS

We would like to extend our thanks to Dr. Mitali Mukerji, Dr. Bhavana Prasher, Dr. Shilpi Aggarwal and Dr. Tavpritesh Sethi from the Institute of Genomics and Integrative Biology, New Delhi and Mr. Nandakumar Selvaraj from CBME–IIT Delhi for their support and encouragement. We would also like to thank Prakhar Prakash for his help with wavelet decomposition.

REFERENCES

[1] Yoon, Y., Lee, M., and Soh, K., (2000) Pulse type classification by varying contact pressure, IEEE Engineering in Medicine and Biology Magazine, **19**, 106–110.

[2] Fuzzy Theory in Traditional Chinese Pulse Diagnosis, Proceedings of 1993 International Joint Conference on Neural Networks.

[3] Wang, H. and Cheng, Y., (2005) A quantitative system for pulse diagnosis in traditional chinese medicine, IEEE EMBS, 5676–5679.

[4] Xu, L. S., Wang, K. Q., and Zhang, D., (2002) Modern researches on pulse waveform of TCPD, IEEE.

[5] Joshi, A., Kulkarni, A., Chandran, S., Jayaraman, V. K., and Kulkarni, B. D., Nadi tarangini: A pulse based diagnostic system, Proceedings of the 29th Annual International Conference of the IEEE EMBS.

[6] Grap, A., (1995) An introduction to wavelets, IEEE Comp. Sc. and Engg., **2(2)**.

[7] Torrence, C. and Compo, G. P., (1998) A practical guide to wavelet analysis, American Meteorological Society.

[8] Holger, G. and Adelmann, (1999) Design of a PC-based system for time-domain and spectral analysis of heart rate variability, Computers and Biomedical Research, **32**, 77–92.

[9] Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, (1996) Heart rate variability standards of measurement, physiological interpretation and Clinical use, European Heart Journal, **17**, 354–381.

[10] Malik, M., (1996) Heart rate variability–Standards of measurement, physiological interpretation and clinical use–Task Force–European Society of Cardiology and The North American Society of Pacing and Electrophysiologu, Eur. Heart J., **17(354)**, 81.

[11] Brennan, M., Palaniswami, M., and Kamen, P., (2002) Poincare plot interpretation using a physiological model of HRV based on a network of oscillators, Am. J. Physiol. 283: H1873-H1886.

J. Biomedical Science and Engineering, 2009, 2, 480-483 **JBiSE**
doi: 10.4236/jbise.2009.27069 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Analysis of positive feedback in the control of movement

Soroor Behbahani[1], Amir Homayoun Jafari[2]

[1,2]Biomedical Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Email: [1]soroor_behbahani@yahoo.com; [2]Amir_h_jafari@rcstim.ir

Received 1 June 2009; revised 13 July 2009; accepted 15 July 2009.

ABSTRACT

Over the past three decades, neurophysiologists studying the neural circuitry responsible for control of skeletal muscles have developed several different general theories of sensorimotor control. These have usually invoked one or more of the sources of proprioceptive signals (e.g. muscle spindle and Golgi tendon organ afferents) in positive or negative feedback loops to the homonymous alpha motoneurones. In this paper we consider to analyze the role of positive feedback in combination of negative feedback due to important role of them in stabilizing the neuromuscular system.

Keywords: Positive Feedback; Negative Feedback; Hill Model; Reflex Model

1. INTRODUCTION

The stretch reflex differs in decerebrate and intact animals and since Sherrington's time it has come to be realized that several CNS mechanisms may contribute components of different latency to the stretch reflex response. At the segmental level, muscle spindle Ia afferents activated by muscle lengthening monosynaptically excite homonymous alpha motoneurons which in turn cause the muscle to resist the stretch. In static postures Ib input generally results in homonymous inhibition, but it has been shown that this switches to longer-latency homonymous excitation during locomotion [1], at least in cat extensor muscles. Group II input from muscle spindles has also been implicated in long latency components of stretch reflexes [2,3,4]. Ia homonymous excitation represents negative displacement feedback, which augments the intrinsic stiffness of active muscles in the face of length perturbations. Ib homonymous feedback on the other hand represents positive force feedback. Positive feedback is synonymous with instability and oscillation in engineering systems, but when muscles are the actuators; their nonlinear lengthtension properties turn out to stabilize the positive feedback loop [5].

2. INTERACTIONS BETWEEN POSITIVE AND NEGATIVEFEEDBACK

According to linear systems theory, positive feedback loops alone produce instability, while negative loops alone may require very high gains to be useful. The general advantage of combining positive and negative feedback loops is that it reduces the sensitivity of system response to unpredictable variations in any of its components, for example, fatiguing of a muscle or increase of a mechanical load [5], without requiring excessive negative feedback gains.

On the other hand, Prochazka [6] demonstrated that positive force feedback could also perform this role assisted by negative length feedback.

Positive force feedback through group Ib afferents may work in parallel to the positive length feedback through the loop and/or group II afferent modulation of activity [7,8]. In the case of Ib force feedback, the loop is polysynaptic so that transmission through it is state dependent. One possible advantage of loops is that they influence and motoneurons monosynaptically [6].

In current research we analyzed the role of positive feedback in combination of negative feedback due to important role of them in stabilizing the neuromuscular system. We used a simple reflex model to show the behaviors of neuromuscular system. Whole of these models were done using MATLAB software.

2.1. Reflex Model

Figure 1 shows a highly simplified model of a reflex system. In this model, key elements of real neuromuscular systems such as tendon compliance, dynamic transfer functions of sensor, and length and velocity dependence of muscle force production are omitted or simplified. The load in this model is a mass of 1 kg.

In this model, muscle force is modeled by a first-order active component and a viscoelastic parallel stiffness. The force velocity relationship [9,10], the muscle length tension curve and tendon compliance are all neglected. 12 are used to represent the image obtained by limb leads.

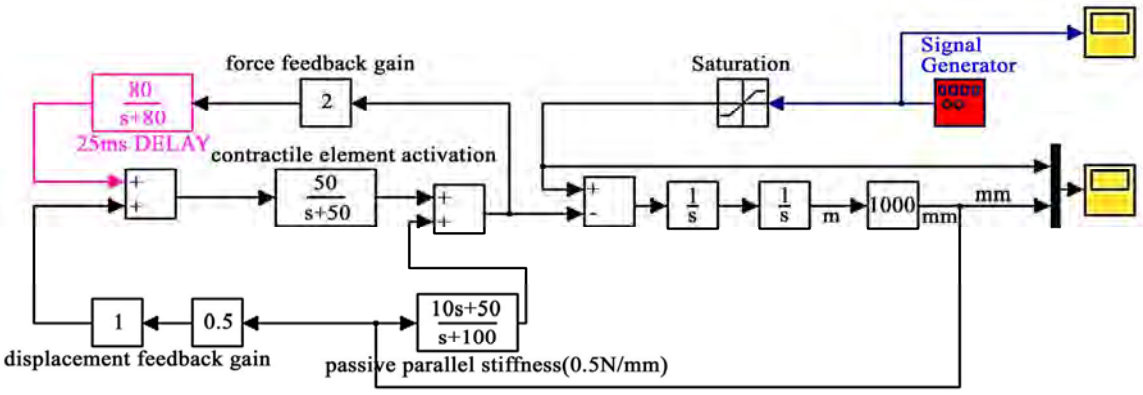


Figure 1. Reflex model.

The external force is summed with tendon force, the Resultant force acting on the inertial load. Force feedback is represented without dynamic components, but an adjustable reflex delay is included because this was shown in the empiric work to stabilize positive force feedback. Feedback from muscle spindles is likewise represented without dynamics and without delays.

Many combinations of parameters were tested in simulations. **Figure 2** shows a small set of simulations. All simulations were performed with the use of the Runge-Kutta-3 simulation algorithm in the Simulink program. Two variables are plotted: external force (perturbing input) and displacement (output).

3. RESULTS

Without displacement feedback or force feedback, the system, comprising the viscoelastic muscle parallel stiffness and the inertial load, showed a damped displacement response (**Figure 2-1**) to the external force.

Adding positive force feedback at loop gains produced stable load compensation (yield in **Figure 2B** with $G_f = 0.5$ is about half that in **Figure 2-1** without feedback).

In this model, G_f was determined by inspection as the multiple of the static gains around the force loop. Because displacement is held constant for this purpose, the only elements involved are the muscle contractile element and the force feedback element. When $G_f = 1$ in **Figure 2B**, apart from the small transients, the mass did not move, i.e., the system had infinite stiffness.

The loop remained stable up to $G_f = 1.4$ (**Figure 2C**), exhibiting an affirming reaction (i.e., instead of yielding, the mass now moved in the opposite direction of the imposed force). The reason that the system remained stable for values of G_f between 1 and 1.4 is that the parallel stiffness element provided negative displacement feedback and velocity feedback, which had a stabilizing effect (if the gain of this element was set to 0, the system went unstable at $G_f = 1$).

Adding a 25-ms delay to the force feedback pathway had a stabilizing effect (**Figure 2D**), allowing a higher loop gain to be attained with a larger affirming reaction (**Figure 2F**).

Note that the magnitude and sign of feedback gain are separately specified in this paper.

Negative force feedback in combination with negative displacement feedback resulted in a spring like response to external loading, the stiffness of which increased as G_d increased and G_f decreased (not illustrated). Comparison of **Figures 3E** and **3H** shows another results when the displacement feedback gain is exponential instead of constant value. The stability of system increased and the settling time will become shorter so the system will be more rapid.

Figures 3F and **3G** show that if we increase the force feedback to $G_f = 2$, the amplitude of oscillation parts will increase and transient response of the system will become more oscillating and the settling time will also increase so increasing the G_f for more than 1.4 will lead to oscillating the system, and it emphasize that max gain for obtaining optimum response is $G_f = 1.4$. **Figures 3I** and **3J** show that increasing the displacement feedback gain is effective in specific limitation. While $G_d = 0.5$ in comparison of $G_d = 1$ the system has more oscillating responses and the settling time is longer.

But in comparison of **Figures 3I** and **3J** we could infer the optimum G_d is approximately $G_d = 1$, because increasing G_d to value of 2 lead to oscillating the responses of the system. Although these oscillations are smaller than the oscillations of system when $G_d = 0.5$ and the settling time is still shorter.

4. DISCUSSIONS AND CONCLUSIONS

In the simple reflex model the muscle contractile element is modeled as a first-order low-pass filter with a cutoff frequency of 8 Hz, somewhat higher than the

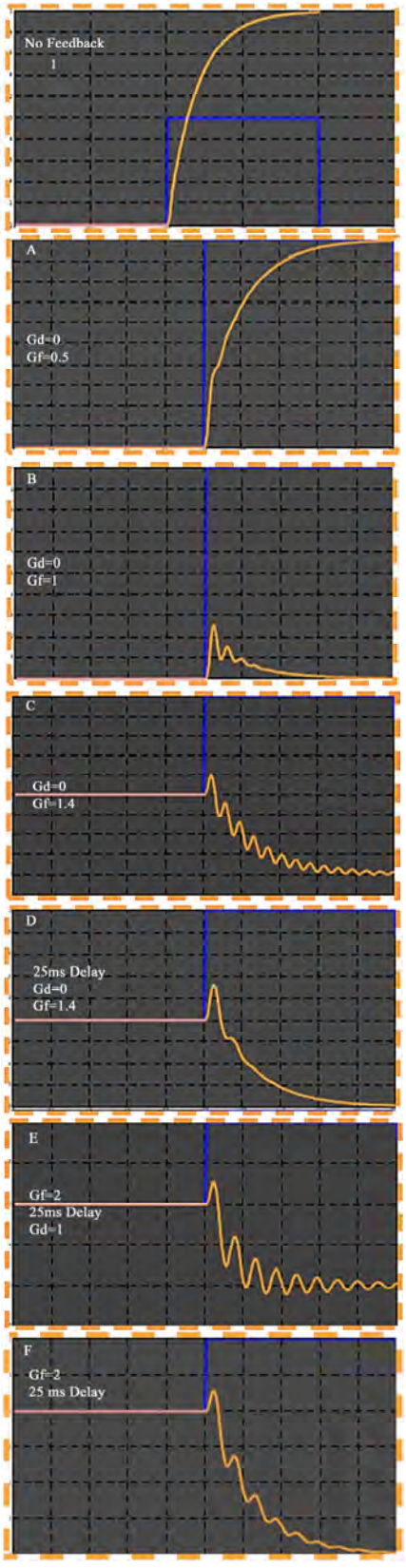


Figure 2. Simulations based on static reflex model of **Figure 1**.

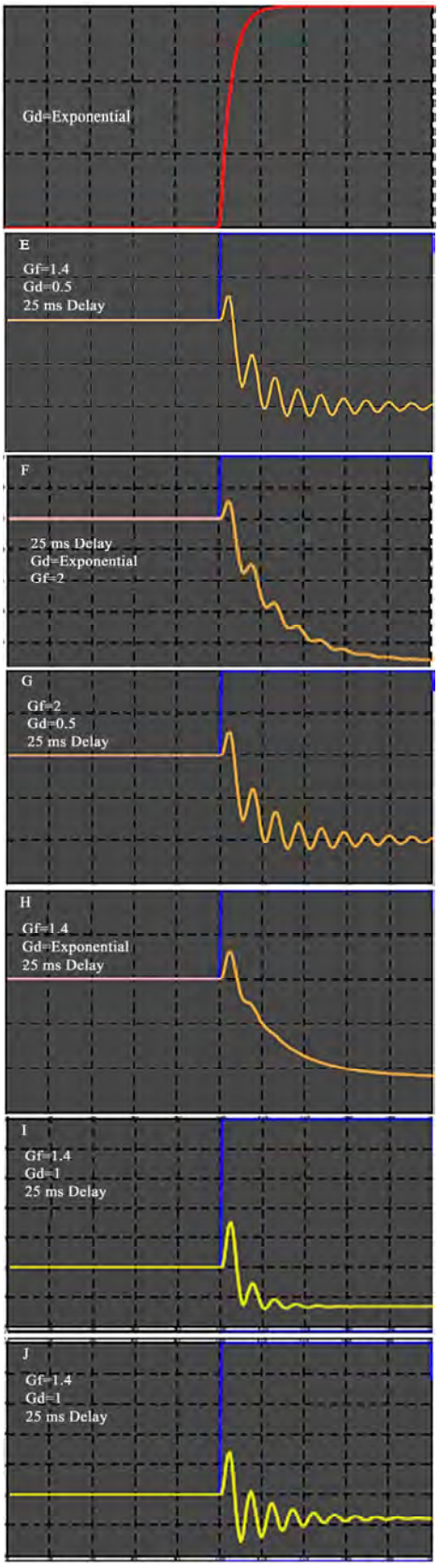


Figure 3. Exponential displacement feedback.

isometric frequency response characteristic of cat triceps surae muscles of Rosenthal that positive force feedback may be appropriate in some *et al.* 1970. A mass of 1 kg represents the inertial load borne by a cat hind limb. A force increment of 10 N represents the mean force developed by triceps surae during the stance phase of slow gait [11]. Inherent muscle properties are simplified to a linear viscoelastic element with a stiffness of 0.5 N/mm. The force feedback and displacement feedback signals are represented without dynamics. It was clear that the reflex mechanism in question represented positive feedback and this was normally associated with instability. It was tacitly assumed that the nervous system would somehow always limit positive force feedback gain within a range consistent with stability. Our results suggest that a combination of intrinsic muscle properties, concomitant negative displacement feedback, and reflex delays found in neuromuscular may provide this automatic gain control.

Our experiments and analysis verify that positive force feedback in the neuromuscular system can provide stable and effective load compensation. The analysis also shows that the conclusions regarding the stabilizing influence of muscle intrinsic properties, length feedback, and delays in positive feedback pathways were robust in the face of large parametric and structural variations in the systems considered. Stable behavior for large values of positive feedback gains was unexpected and initially quite puzzling. However, it became apparent that loop gain did not remain high, but rather it was automatically attenuated when muscles shortened and thereby reduced their force producing capability.

REFERENCES

[1] Brownstone, R. M., Gossard, J. P., and Hultborn, H., (1994) Voltage-dependent excitation of motoneurones from spinal locomotor centres in the cat, Springer-Verlag, Exp. Brain Res., **102**, 3444.

[2] Grey, M. J., Mazzaro, N., Nielsen, J. B., and Sinkjær, T., (2004) Ankle extensor proprioceptors contribute to the enhancement of the soleus EMG during the stance phase of human walking1, Can. J. Physiol. Pharmacol., **82**, 610–616.

[3] Donelan, J. M. and Pearson, K. G., (2004) Contribution of sensory feedback to ongoing ankle extensor activity during the stance phase of walking, Can. J. Physiol. Pharmacol, **82**, 589–598.

[4] Frigon, A. and Rossignol, S., (2006) Experiments and models of sensorimotor interactions during locomotion, Springer Berlin/Heidelberg, Saturday, DOI: 10.1007/s 00422-006-0129-x.

[5] Prochazka, A. and Gillard, D., (1997) Sensory control of locomotion, Proceeding of American Control Conference.

[6] Prochazka, A., Gillard, D., and Bennet, D. J., (1997) Implications of positive feedback in the control of movement, the American Physiological Society 0022-3077/97, 5.00 Copyright.

[7] Taylor, A., Ellaway, P. H., Durbaba, R., and Rawlinson, S., (2000) Distinctive patterns of static and dynamic gamma motor activity during locomotion in the decerebrate cat, Journal of Physiology, **529(3)**, pp. 825–836.

[8] Raya, J. G., Ramı´rez, A., and Mun˜oz-Martı´nez, E. J., (2004) Gamma 3 alpha linkage and persistent firing of Ia fibers by pudendal nerve stimulation in the decerebrate cat, J. Neurophysiol, **92**, 387–394.

[9] Krakauer, J. W., MD, (2006) Spinal cord physiology: Receptors and reflexes.

[10] Csercsik, D., (2006) Construction of simple dynamic models of the gamma-loop mechanism, Research Report SCL-001.

[11] Knutson, G. A., Edward, D. C., and Owens, F., MS, DC, (2003) Active and passive characteristics of muscle tone and their relationship to models of subluxation, 0008-3194/2003/269–283/$2.00/\JCCA.

J. Biomedical Science and Engineering, 2009, 2, 484-490 **JBiSE**
doi: 10.4236/jbise.2009.27070 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Calibrating volume measurements made using the dual-field conductance catheter

Simon P. McGuirk[1,2], David J. Barron[1], Dan Ewert[3], John H. Coote[2]

[1]Department of Cardiac Surgery, Birmingham Children's Hospital, Birmingham, UK; [2]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK; [3]Department of Electrical and Computer Engineering, North Dakota State University, Fargo, USA.
Email: simon.mcguirk@nhs.net

Received 1 February 2009; revised 24 June 2009; accepted 16 July 2009.

ABSTRACT

The conductance catheter technique allows real-time measurements of ventricular volume based on changes in the electrical conductance of blood within the ventricular cavity. Conductance volume measurements are corrected with a calibration coefficient, α, in order to improve accuracy. However, conductance volume measurements are also affected by parallel conductance, which may confound calibration coefficient estimation. This study was undertaken to examine the variation in α using a physical model of the left ventricle without parallel conductance. Calibration coefficients were calculated as the conductance-volume quotient ($\alpha_{V(t)}$) or the stroke conductance-stroke volume quotient (α_{SV}). Both calibration coefficients varied as a non-linear function of the ventricular volume. Conductance volume measurements calibrated with $\alpha_{V(t)}$ estimated ventricular volume to within 2.0 ± 6.9%. By contrast, calibration with α_{SV} substantially over-estimated the ventricular volume in a volume-dependent manner, increasing from 26 ± 20% at 100ml to 106 ± 36% at 500ml. The accuracy of conductance volume measurements is affected by the choice of calibration coefficient. Using a fixed or constant calibration coefficient will result in volume measurement errors. The conductance-stroke volume quotient is associated with particularly significant and volume-dependent mea- surement errors. For this reason, conductance volume measurements should ideally be calibrated with an alternative measurement of ventricular volume.

Keywords: Conductance Catheter; Calibration; Volume Measurement

1. INTRODUCTION

The conductance catheter technique is an established method to measure the ventricular volume in real-time, based on the electrical conductance of blood within the ventricular cavity [1,2,3]. Conductance volume measurements are based on the assumption that the electric field produced by the conductance catheter is homogeneously distributed within the ventricular cavity [1]. However, theoretical and experimental studies have demonstrated that this assumption is not valid [4,5,6]. As a result, the conductance catheter tends to overestimate the volume in small ventricles and underestimate the volume in larger ventricles [2,3].

The dimensionless calibration coefficient, α, was introduced by Baan *et al*. [3] in order to account for the non-uniform conductance-absolute volume relationship [3]. This calibration coefficient represents the slope of the relationship between the conductance-derived volume and the true volume. The calibration coefficient, α may also vary with ventricular volume. It is relatively high in small animals [7], lower in humans [3] and intermediate values are found in dogs [3,8,9]. Experimental studies demonstrate that α also varies during inferior vena caval occlusion [9] and may even fluctuate during the normal cardiac cycle [10,11].

The tissues and fluid surrounding the ventricular cavity also contribute to the measured conductance signal [3]. This creates a volume offset called parallel conductance. Parallel conductance also varies according to the ventricular volume [4,7,8,9,12,13]. This volume- dependent parallel conductance may confound the calibration coefficient estimation.

This study was undertaken to examine the variation in the calibration coefficient, α, over a range of volumes pertaining to clinical studies, in a physical model of the left ventricle without parallel conductance.

2. METHODS

2.1. Model Ventricle

This study used a physical model of the left ventricle previously described by this laboratory [14]. This model consisted of an ellipsoid latex balloon enclosed in a pressurised Perspex chamber (**Figure 1**). The chamber was filled with distilled water and hydraulically pressurised with an intra-aortic balloon pump (IABP; Datascope Medical Co. Ltd, Huntingdon, UK) with a paediatric volume-limiting chamber. The IABP console was connected to two 25ml intra-aortic balloon catheters in parallel. The IABP circuit was filled with helium gas to ensure a rapid pneumatic response.

The electrocardiogram (ECG) from a patient simulator (Bioma Research Inc, Quebec, Canada) was used to trigger the IABP console at a predetermined rate (60 beats·min^{-1}). Inflation of the IABP balloon catheters caused a rise in the ventricular pressure. This displaced saline from the latex balloon through a 2/2-way solenoid valve (Bürkert GmbH, Ingelfingen, Germany) into a calibrated measuring cylinder at the top of the model ventricle. Deflation of the IABP balloon catheters caused ventricular pressure to fall, which allowed the latex balloon to refill. Electronic circuitry was used to control the opening and closure times of the solenoid valve in order to simulate different contraction patterns.

For the purposes of this study, the solenoid valve remained open throughout the model ventricular cycle to produce isobaric contractions. Each inflation of the IABP balloon catheters displaced a fixed stroke volume (*SV*) of 50ml. The stroke volume was independently

quantified before each experiment. The IABP balloon catheters were manually inflated and the displaced volume was measured in the measuring cylinder.

Three separate latex balloons were used to simulate the change in volume that may be observed during inferior vena caval occlusion. Each balloon was 13 cm in length and had a maximal volume at atmospheric pressure of 125ml, 215ml and 500ml, respectively. The shape of the balloons was ellipsoid at low volumes and became increasingly spheroid as the volume increased. The length-diameter ratio at maximal filling decreased from 3.03 for the smallest balloon to 1.52 for largest balloon.

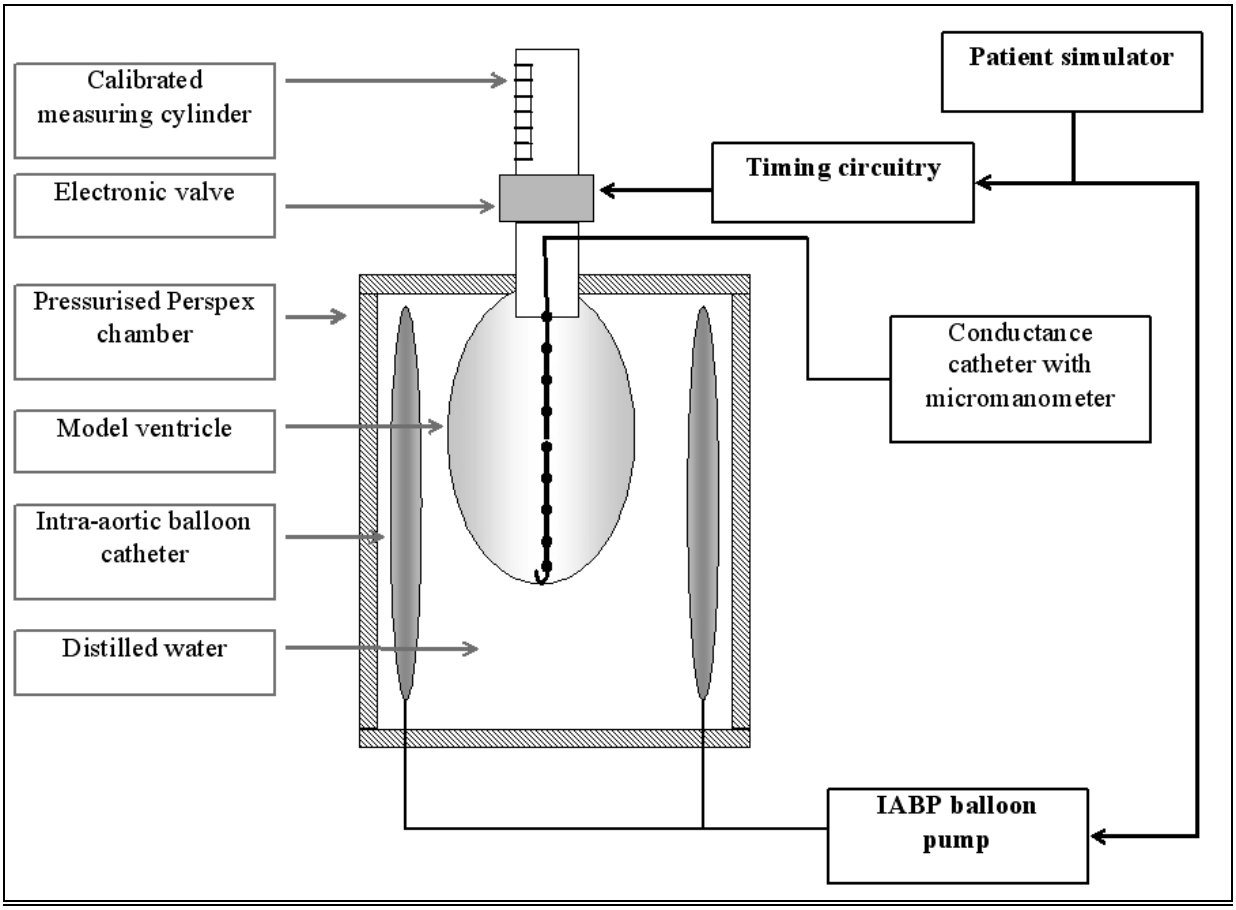


Figure 1. Schematic diagram of the model heart consisting of a model ventricle and outflow tract. The position of the conductance catheter is also illustrated.

During each experiment, the model ventricle was filled with a known volume of 0.9% normal saline at room temperature. The end-diastolic volume (V_{ED}) was varied between 100–125ml for the 125ml balloon; 160–215ml for the 215ml balloon; and 410–500ml for the 500ml balloon. The resistivity (conductivity^{-1}) of this solution was measured before each test using a dedicated measuring cuvette (CD Leycom, Zoetermeer, The Netherlands). The mean ± SD resistivity was 68.3 ± 0.5 Ω·cm.

2.2. Conductance Catheter

The principle of the conductance catheter technique for measuring LV volume has been described elsewhere [3]. In this study, a 7-French 12-electrode high-fidelity dual pressure-volume conductance catheter (Millar Instruments, Houston, TX, USA) was used. The electrodes were mounted at 10 mm intervals near the tip of the catheter. One alternating current (20 kHz, 30 µA RMS) was applied between the two outermost electrodes (electrodes 1 and 12) and a second alternating current (20 kHz, -10 µA RMS) was applied between the two adjacent electrodes (electrodes 2 and 11). This dual-field configuration was used in all studies [15]. A signal processor unit (CFL-512; CD Leycom) was used to measure the potential difference between seven consecutive pairs of the remaining eight electrodes (electrodes 3–10). These measured voltages were converted into seven time-varying segmental conductance signals, $G_i(t)$.

The conductance volume, $Q(t)$, was measured using the following formula:

$$Q(t) = L^2 \cdot \rho \cdot ([\sum_{i=1}^{7} G_i(t)] - G_P) \qquad (1)$$

where ρ is the blood resistivity, L is the inter-electrode distance, and G_P is parallel conductance.
By assuming that parallel conductance is negligible (i.e. $G_P = 0$), this formula can be simplified to:

$$Q(t) = L^2 \cdot \rho \cdot [\sum_{i=1}^{7} G_i(t)] \qquad (2)$$

The volume estimated using the conductance catheter technique was then calculated:

$$V_g(t) = \frac{1}{\alpha} \cdot Q(t) \qquad (3)$$

where α is the dimensionless calibration coefficient [3].

2.3. Experimental Data

Analogue signals representing 7 segmental conductance volumes within the model ventricle and the ECG were digitised at 12-bit accuracy and a sample frequency of 250 Hz. End-diastole and end-systole were retrospectively identified. End-diastole was defined as the R wave on the ECG and end-systole was defined as the point immediately prior to IABP circuit deflation.

Each experiment was conducted three times and, in each, data from 10 consecutive cycles were analysed. The estimated within-experiment standard deviation was 0.45 ml. The standard deviation was not significantly correlated with the mean conductance volume (Kendall's τ coefficient = -0.72), and so subsequent analyses were based on the average data from each experiment [16].

Two separate calibration coefficients were calculated. Firstly, the calibration coefficients, $\alpha_{V(t)}$ was calculated by dividing the conductance-derived volume measurement by the absolute ventricular volume at either end-diastole or end-systole:

$$\alpha_{V(t)} = \frac{Q_{ED}}{V_{ED}} \text{ or } \frac{Q_{ES}}{V_{ES}} \qquad (4)$$

where $V_{ES} = V_{ED} - SV$.

Including a "phase of measurement" (i.e. *ED* or *ES*) in this equation did not improve the goodness of fit of the model. The effect of "phase of measurement" was not significant (P=0.38) and was therefore not included in the model used to predict $\alpha_{V(t)}$.

The conductance-stroke volume quotient, α_{SV}, was calculated by dividing the conductance-derived stroke volume by the absolute stroke volume:

$$\alpha_{SV} = \frac{(Q_{ED} - Q_{ES})}{SV} \qquad (5)$$

Finally, the measurement error, ζ was calculated as a percentage of the absolute ventricular volume:

$$\zeta = \frac{V_g(t)}{V(t)} \cdot 100 \qquad (6)$$

The terms $\zeta_{V(t)}$ and ζ_{SV} were used to denote the measurement error associated with $\alpha_{V(t)}$ and α_{SV}, respectively.

2.4. Data Analysis

Data were analysed using SPSS for Windows (v12, SPSS Inc., Chicago, IL, USA). Conductance volumes, calibration coefficients and measurement errors were in turn examined as functions of the known volume of the model ventricle. These relationships were evaluated by least squares linear regression based on fractional polynomials of the ventricular volume. The standard error for the regression coefficients and intercept were also calculated. A probability, P<0.05, was taken to represent statistical significance in these analyses. Correlation coefficients between the variables were also calculated.

3. RESULTS

3.1. Relationship between Conductance and Absolute Volume

The simultaneous conductance volume, $Q(t)$, and absolute volume measurements, $V(t)$ for the three balloons

are illustrated in **Figure 2**. There was a non-linear relationship between these two volume measurements. When analysed as a whole, the conductance-absolute volume relationship was best approximated by the following relationship (**Eq. 7**).

$$Q(t) = \left[(14.2 \pm 0.39) \cdot \sqrt{volume}\right] - \left[(0.17 \pm 0.02) \cdot volume\right], (r^2 = 0.90, P < 0.001) \quad (7)$$

Compared with the line of identity (i.e. $x = y$), conductance volume measurements predicted using this model were equal to ventricular volume at approximately 150 ml; slightly overestimated ventricular volume when the absolute volume was less than 150 ml; but underestimated ventricular volume at volumes over 150 ml.

The volume measurements formed two distinct subsets. Data from the small and medium balloons were clustered on the left-hand side of the plot while data from the large balloon was clustered on the right-hand side. When considered individually, the conductance-absolute volume relationship was well approximated by linear regression (**Eq.s 8 and 9**). In neither case was the quadratic term significant. For the small and medium-sized balloons, the linear regression was:

$$Q(t) = \left[(0.63 \pm 0.06) \cdot V(t)\right] + (55 \pm 9), \quad (r^2 = 0.90, P < 0.001) \quad (8)$$

while for the large balloon, the corresponding equation was:

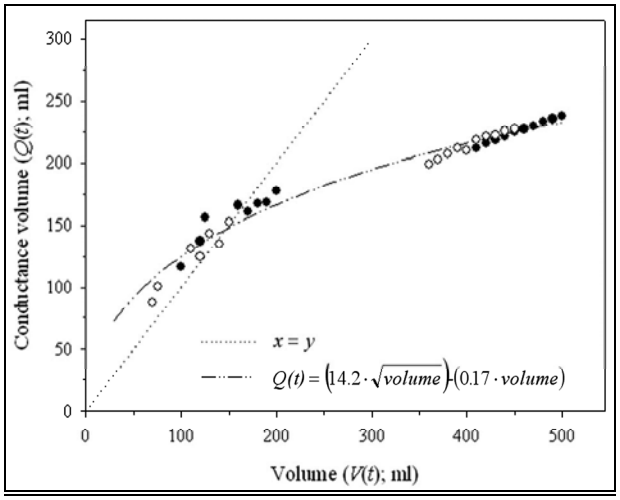


Figure 2. Conductance volume measurements, $Q(t)$ versus absolute volume, $V(t)$ at end-diastole (●) and end-systole (○) for each of the three latex balloons. The curve (dashed line) represents the regression analysis of the conductance-absolute volume relationship. The curve (dotted line) representing the line of identity ($x=y$) is also illustrated.

$$Q(t) = \left[(0.27 \pm 0.01) \cdot V(t)\right] + (104 \pm 5), \quad (r^2 = 0.96, P < 0.001) \quad (9)$$

The slope of these two mathematical models are significantly different from one another ($P<0.05$), further illustrating that a single, common linear relationship does not hold over the entire range of absolute volumes. In addition, the intercepts differ significantly from one another and from zero ($P<0.05$). Therefore, neither of these two lines passes through the origin. This introduces a "volume offset", which in this apparatus cannot be due to parallel conductance.

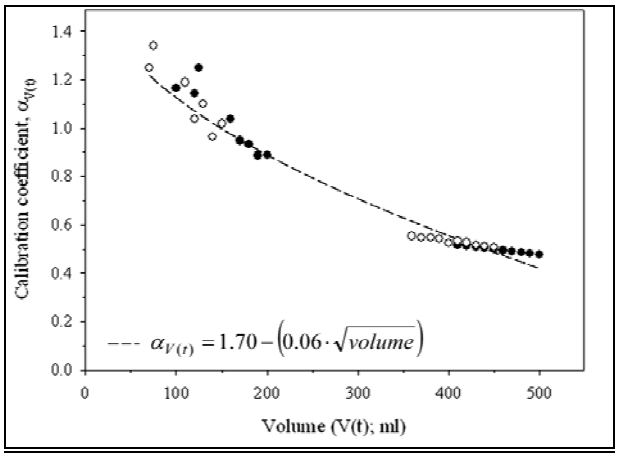
3.2. Relationship between Calibration Coefficients and Absolute Volume

The calibration coefficients, $\alpha_{V(t)}$ and α_{SV} were calculated using **Eq.s 4** and **5**, respectively. Neither calibration coefficient was constant or linearly related to the ventricular volume. Instead, both calibration coefficients decreased progressively as the ventricular volume increased (**Figures 3A and 3B**). The calibration coefficient, $\alpha_{V(t)}$ varied as a function of the square root of the absolute volume (r^2=0.97, $P<0.001$). By contrast, α_{SV} varied as a function of the inverse absolute volume (r^2=0.98, $P<0.001$; and r^2=0.93, $P<0.001$ for end-diastolic and end-systolic measurements respectively).

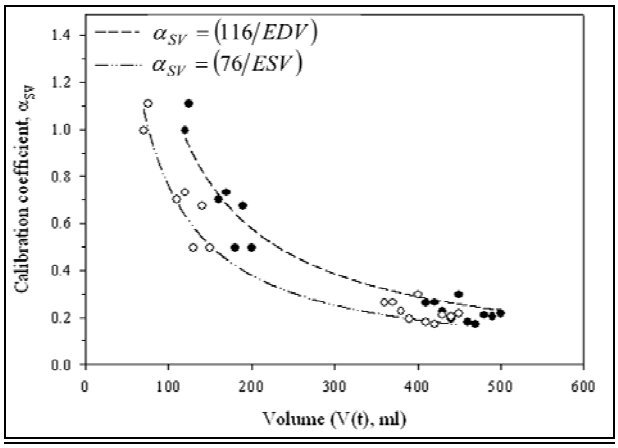
The relative ratio of the two calibration coefficients (i.e. α_{SV} / $\alpha_{V(t)}$) was also examined over the volume range (**Figure 3C**). The stroke volume quotient, α_{SV} was lower than $\alpha_{V(t)}$ at each volume measurement. However, the slope calibration coefficient ratio was not constant but became progressively smaller as the absolute volume was increased. For example, the α_{SV} / $\alpha_{V(t)}$ ratio decreased from 0.98 ± 0.52 to 0.43 ± 0.23 as the end-diastolic volume increased from 100ml to 500ml.The best approximation for this relationship was that the α_{SV} / $\alpha_{V(t)}$ ratio varied as a function of the inverse ventricular volume (r^2=0.81, $P<0.001$; r^2=0.75, $P<0.001$ for end-diastolic and end-systolic measurements, respectively).

3.3. Relationship between Calibrated Conductance and Absolute Volume and Measurement Error

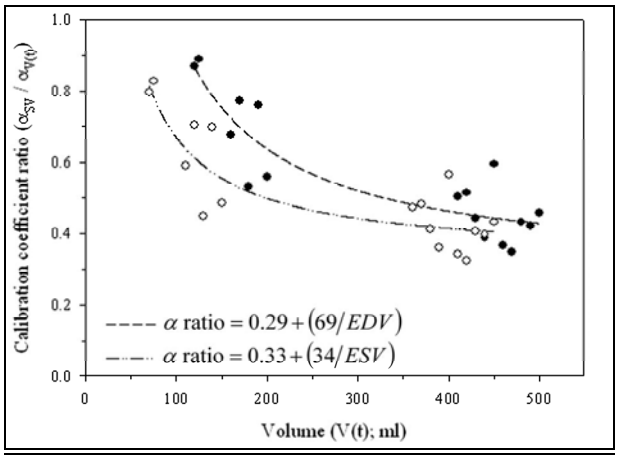
Calibrated conductance volume measurements, $V_g(t)$ were calculated using **Eq. 3**, in which α was either $\alpha_{V(t)}$-volume or α_{SV}-volume relationship, as previously described (**Figure 4**). Conductance volume measurements calibrated with the $\alpha_{V(t)}$-volume relation increased linearly with the absolute volume (r^2 =0.99, $P<0.001$). Conductance volume measurements calibrated with the α_{SV}-volume relation also increased linearly with absolute volume (r^2=0.99, $P<0.001$), but overestimated the absolute volume.



(a)



(b)



(c)

Figure 3. Calibration coefficient, α versus absolute volume, $V(t)$ at end-diastole (●) and end-systole (○) for each of the three latex balloons. The calibration coefficient was calculated as the conductance-absolute volume quotient ($\alpha_{V(t);}$ A) or the conductance-stroke volume quotient ($\alpha_{SV;}$ B). The calibration coefficient ratio (C) represented α_{SV} as a proportion of $\alpha_{V(t)}$. Curves (dashed lines) representing regression analyses of the calibration coefficient-conductance volume relations are also illustrated.

The volume measurement error, expressed as a percentage of the absolute volume, was calculated for each calibration coefficient (**Figure 5**). Conductance volume measurements calibrated with the $\alpha_{V(t)}$-volume relation had a measurement error of 2.0 ± 6.9%, which was relatively constant across the span of volume measurements. On the other hand, the measurement error associated with the α_{SV}-volume relation was volume-dependent, increasing linearly from 26 ± 20% to 106 ± 36% as the end-diastolic volume increased from 100ml to 500ml (r^2=0.96, P<0.001).

4. DISCUSSION

The non-homogeneous electric field generated by the conductance catheter results in a non-linear relationship between conductance and true ventricular volume measurements. Conductance measurements are, therefore, calibrated using the dimensionless calibration coefficient, α in order to obtain absolute volume measurements. This study investigated the accuracy of conductance volume measurements, and the effect of calibration, in a series of *in vitro* experiments that spanned the volume range observed in clinical studies [17]. This model design also avoided the potential problems associated with parallel conductance.

This study has confirmed that there is a non-linear relationship between dual-field conductance and absolute volume measurements, such that the conductance- volume relation is concave towards the true volume. The conductance volume measurements underestimated ventricular volume as the volume was increased above 150ml. This finding is in accord with previous experimental studies [4,6].

Conductance volume measurements are generally calibrated with the dimensionless calibration coefficient in order to improve the accuracy of ventricular volume measurements made using this technique [3]. In the isolated post-mortem canine heart, Mur & Baan reported that the conductance-volume relation was virtually linear over a finite volume range [2] and the authors predicted that a similar, virtually linear conductance-volume relation would be observed for the human left ventricle up to a volume of 200ml [2].

However, this study has shown that the calibration coefficient, α is not constant, but varies as a non-linear function of the ventricular volume. Using a fixed or constant α during acute volume change will inevitably result in volume measurement errors. These measurement errors may be relatively small during the normal cardiac cycle. However, procedures that produce an acute and substantial change in volume load, like vena caval occlusion, will potentially result in significant measurement errors.

Kornet *et al*. suggested that the variation in α observed *in vivo* might reflect synchronous volume-dependent

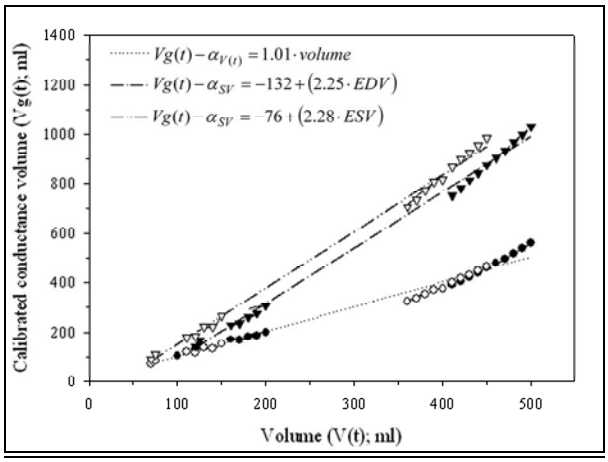


Figure 4. Calibrated conductance volume measurements, $V_g(t)$ versus absolute volume, $V(t)$ at end-diastole (●) and end-systole (○) for each of the three latex balloons. Conductance volume measurements were calibrated using either $\alpha_{V(t)}$ (●, ○) or α_{SV} (▼, ∇).Curves (dashed lines) representing regression analyses of the calibrated conductance volume-absolute volume relations are also illustrated.

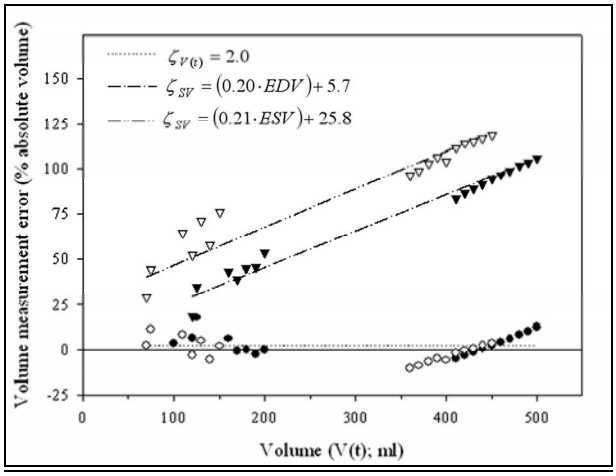


Figure 5. Calibrated conductance volume measurement error, ζ versus absolute volume, $V(t)$ at end-diastole (●) and end-systole (○).Conductance volume measurements calibrated using either $\alpha_{V(t)}$ (●, ○) or α_{SV} (▼, ∇). Volume measurement error was calculated as a percentage of the absolute volume. Curves (dotted and dashed lines) representing regression analyses of the measurement error-absolute volume relations are also illustrated.

changes in parallel conductance during the cardiac cycle [13]. However, this study used a physical model of the left ventricle without parallel conductance. Our findings suggest that parallel conductance is not a comprehensive explanation for why the calibration coefficient varies as a function of the ventricular volume.

Although conductance volume measurements can be calibrated against another synchronous estimate of ventricular volume [3,10,18], they are generally calibrated with an alternate measurement of stroke volume [6,10]. Calibration with either an alternate measurement of ventricular volume or stroke volume are generally considered equivalent. This study also identified that the $\alpha_{V(t)}$ and α_{SV} are different. The stroke volume quotient, α_{SV} was lower than $\alpha_{V(t)}$ and the calibration coefficient ratio, $\alpha_{SV} / \alpha_{V(t)}$ became progressively smaller as the ventricular volume increased. Calibrating conductance volume measurements with α_{SV} resulted in the significant overestimation of end-diastolic and end-systolic volume, and the degree of overestimation was even more pronounced at higher volumes. This demonstrates that conductance volume measurements should ideally be calibrating using an independent measure of ventricular volume like, for example, contrast cineangiography, echocardiography or magnetic resonance imaging.

The difference between the $\alpha_{V(t)}$-volume and α_{SV}-volume relations may be attributed to the non-homogeneous electrical field distribution established by the conductance catheter [4,12]. Conductance volume measurements are disproportionately influenced by the areas around the longitudinal axis of the ventricle where the electrical field is strongest. By contrast, changes in volume that occur during ejection or when loading conditions are varied primarily affect the myocardial boundary where the electrical field density is weakest. Consequently, α_{SV} and $\alpha_{V(t)}$ are not equivalent and cannot be used interchangeably.

4.1. Study Limitations

The physical model of the isolated left ventricle used in this study did not allow changes in the volume loading conditions could not be modelled continuously. Instead, a series of steady-state experiments were made at incremental end-diastolic volumes in three separate models, which carries with it the risk of repeated measurement errors. The changes in ventricular volume resulted from changes in the short-axis dimension of the model only. It was not possible to examine any effect of changing ventricular length, such as those that may occur during the cardiac cycle [19].

Ideally, the volumes within the model ventricle would have been acquired simultaneously using the conductance catheter and another independent method. In the original design, the ejected volume was measured according to the pressure generated within the fluid column [14]. However, preliminary experiments demonstrated that the fluid-filled micromanometer was not sufficiently accurate to enable instantaneous volume measurements throughout the cardiac cycle and, therefore, measurements were made with the calibrated measuring cylinder.

5. CONCLUSIONS

The conductance catheter technique incorporates a calibration coefficient, α, in order to obtain accurate volume

measurements. This study has demonstrated that this calibration coefficient varies as a function of absolute volume, independent of parallel conductance. Assuming that the calibration coefficient is fixed or constant will introduce measurement errors. The conductance-stroke volume quotient, α_{SV} is associated with particularly significant and volume-dependent measurement errors. This limits the value of volume measurements calibrated using α_{SV}. Conductance volume measurements should ideally be calibrated with an alternative measurement of ventricular volume, using any one of the techniques that are now available.

6. ACKNOWLEDGEMENTS

Simon McGuirk was supported by a British Heart Foundation Junior Research Fellowship (FS/03/102).

REFERENCES

[1] Baan, J., Aouw Jong, T. T., Kerkhof, P. L., Moene, R. J., van Dijk, A. D., van der Velde, E. T., and Koops, J., (1981) Continuous stroke volume and cardiac output from intra-ventricular dimensions obtained with an impedance catheter. Cardiovascular Research, **15**, 328–334.

[2] Mur, G. and Baan, J., (1984) Computation of the input impedances of a catheter for cardiac volumetry, IEEE Transactions on Biomedical Engineering, **31**, 448–453.

[3] Baan, J., van der Velde, E. T., de Bruin, H. G., Smeenk, G. J., Koops, J., van Dijk, A. D., Temmerman, D., Senden, J. and Buis, B., (1984) Continuous measurement of left ventricular volume in animals and humans by conductance catheter, Circulation, 70, 812–823.

[4] Wu, C. C., Skalak, T. C., Schwenk, T. R., Mahler, C. M., Anne, A., Finnerty, P. W., Haber, H. L., Weikle II, R. M. and Feldman, M. D., (1997) Accuracy of the conductance catheter for measurement of ventricular volumes seen clinically: Effects of electric field homogeneity and parallel conductance, IEEE Transactions on Biomedical Engineering, **44**, 266–277.

[5] Salo, R. W., (1992) Improvements in intracardiac impedance volumes by field extrapolation, European Heart Journal, **13(Suppl E)**, 35–39.

[6] Wei, C. L., Valvano, J. W., Feldman, M. D. and Pearce, J. A., (2005) Nonlinear conductance-volume relationship for murine conductance catheter measurement system. IEEE Transactions on Biomedical Engineering, **52**, 654–661.

[7] Cassidy, S. C. and Teitel, D. F., (1992) The conductance volume catheter technique for measurement of left ventricular volume in young piglets, Pediatric Research, **31**, 85–90.

[8] Boltwood, C. M., Appleyard, R. F., and Glantz, S. A. (1989) Left ventricular volume measurement by conductance catheter in intact dogs: parallel conductance volume depends on left ventricular size, Circulation, **80**, 1360–1377.

[9] Applegate, R. J., Cheng, C. P. and Little, W. C., (1990) Simultaneous conductance catheter and dimension assessment of left ventricular volume in the intact animal, Circulation, **81**, 638–648.

[10] Szwarc, R. S., Laurent, D., Allegrini, P. R., and Ball, H. A., (1995) Conductance catheter measurement of left ventricular volume; evidence for nonlinearity within cardiac cycle, American Journal of Physiology-Heart and Circulatory Physiology, **268**, H1490–H1498.

[11] Danton, M.H., Greil, G.F., Byrne, J.G., Hsin, M. Cohn, L. and Maier, S.E. (2003) Right ventricular volume measurement by conductance catheter. American Journal of Physiology-Heart and Circulatory Physiology, **285**, H1774–H1785.

[12] Kun, S. and Peura, R. A., (1994) Analysis of conductance volumetric measurement error sources, Medical and Biological Engineering and Computing, **32**, 94–100.

[13] Kornet, L., Schreuder, J. J., van der Velde, E. T., and Jansen, J. R., (2001) The volume-dependency of parallel conductance throughout the cardiac cycle and its consequence for volume estimation of the left ventricle in patients, Cardiovascular Research, **51**, 729–735.

[14] Al-Khalidi, A. H., Townend, J. N., Bonser, R. S., and Coote, J. H., (1998) Validation of the conductance catheter method for measurement of ventricular volumes under varying conditions relevant to cardiac surgery, American Journal of Cardiology, **82**, 1248–1252.

[15] Steendijk, P., van der Velde, E. T., and Baan, J., (1992) Single and dual excitation of the conductance-volume catheter analysed in a spheroidal mathematical model of the canine left ventricle, European Heart Journal, **13 (Suppl E)**, 28–34.

[16] Bland, J. M. and Altman, D. G., (1996) Statistics notes: Measurement error, British Medical Journal, **313**, 744.

[17] Tkacova, R., Hall, M. J., Liu, P. P., Fitzgerald, F. S., and Bradley, T. D., (1997) Left ventricular volume in patients with heart failure and Cheyne-Stokes respiration during sleep, American Journal of Respiratory Care Medicine, **156**, 1549–1555.

[18] Kass, D. A., (1992) Clinical evaluation of left heart function by conductance catheter technique, European Heart Journal, **13(Suppl E)**, 57–64.

[19] Rushmer, R. F., Crystal, D. K., and Wagner, C., (1953) The functional anatomy of ventricular contraction, Circulation Research, **1**, 162–70.

J. Biomedical Science and Engineering, 2009, 2, 491-498 **JBiSE**
doi: 10.4236/jbise.2009.27071 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Electrocardiographic interference and conductance volume measurements

Simon P. McGuirk[1,2], Dan Ewert[3], David J. Barron[1], John H. Coote[2]

[1]Department of Cardiac Surgery, Birmingham Children's Hospital, Birmingham, UK; [2]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK; [3]Department of Electrical and Computer Engineering, North Dakota State University, Fargo, USA.
Email: simon.mcguirk@nhs.net

Received 1 February 2009; revised 1 July 2009; accepted 16 July 2009.

ABSTRACT

The conductance catheter technique enables continuous ventricular volume measurements based on the electrical conductance of blood within the ventricular cavity. However, ventricular excitation also produces a measurable electrical signal within the ventricular cavity. This study was undertaken to investigate the relationship between the ventricular electrogram and conductance volume measurements in a physical model of the left ventricle without parallel conductance. The ventricular electrogram was simulated with an ECG signal, ECG_{input} connected to two ring electrodes within the model ventricle. Conductance volume measurements were made with and without ECG_{input}. The difference between these measurements, $G_{ECG}(t)$, represented the conductance volume due to ECG_{input}. $G_{ECG}(t)$ varied as a function of the first-derivative of ECG_{input} with respect to time (r^2=0.92, P<0.001). $G_{ECG}(t)$, This primarily affected volume measurements during ventricular depolarisation; during this phase the volume measurement error varied widely between –12% and +9%. As a result, end-diastole could not be reliably identified on the pressure-volume loop. The accuracy of conductance volume measurements during late diastole and early isovolumic contraction are substantially affected by the ventricular electrogram. This may result in a significant error in end-diastolic volume estimates, which has important implications for the quantitative assessment of ventricular function including, in particular, the assessment of chamber compliance.

Keywords: Volume Measurement; Conductance Catheter; Electrocardiogram; Ventricular Electrogram

1. INTRODUCTION

The assessment of ventricular function is fundamentally important for the evaluation of patients with known or suspected heart disease. Analysis of left ventricular (LV) volume in the time and pressure domains allows systolic and diastolic function to be separately quantified. The conductance catheter technique was developed to continuously measure ventricular volume in real-time [1,2]. These measurements are recorded simultaneously with intraventricular pressure measurements to provide instantaneous pressure-volume data [3].

The conductance catheter technique is associated with two, well known sources of error. Firstly, the current density generated by the conductance catheter is not uniformly distributed throughout the ventricular cavity [4,5,6]. This results in a non-linear conductance-absolute volume relationship [2,4,7]. Conductance volume measurements must be corrected with a calibration coefficient, α [8]. Secondly, the tissues and fluid surrounding the ventricular cavity also contribute to the conductance signal [2,8]. This results in an offset in the conductance-absolute volume relationship, called parallel conductance. In practice, conductance volume measurements are usually calibrated for parallel conductance using the hypertonic saline method in order to derive accurate ventricular volume measurements [8].

In our paediatric clinical experience, we have observed that the pattern of LV volume measurements is frequently abnormal. The conductance volume measurements are characterised by a narrow upward spike followed by a narrow downward spike during late diastole without any commensurate change in LV pressure (**Figures 1A and 1B**). This alters the shape of the pressure-volume loop, with the loss of the normal lower right-hand corner (**Figure 1D**). To our knowledge, this abnormal conductance volume pattern has not previously been described. However, we understand that similar

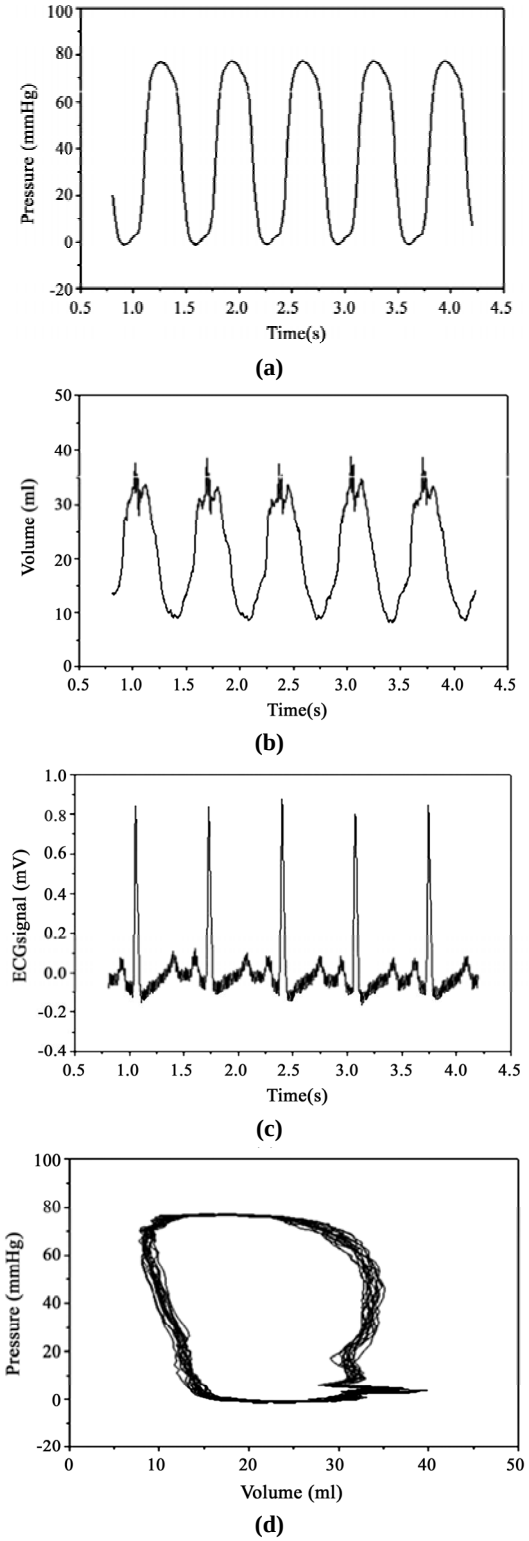


Figure 1. Time-varying pressure (A), conductance volume (B) and surface electrogram (C) signals obtained in a 6 year-old child with tricuspid atresia. The corresponding pressure- conductance volume loop for the same child (D) was developed by plotting the instantaneous pressure against the corresponding conductance volume.

findings have been observed in other patient groups, particularly in patients with a permanent pacemaker [personal communication, Professor M.P. Frenneaux, Department of Cardiovascular Sciences, University of Birmingham, UK; Dr P. Steendijk, Department of Cardiology, Leiden University Medical Center, The Netherlands].

We observed that the abnormal LV conductance volume measurements occurred synchronously with the QRS complex on the surface electrocardiogram (ECG; **Figure 1C**). We hypothesised that this abnormal LV conductance measurement may represent the effect of ventricular excitation on conductance volume measurements, which is superimposed on the normal ventricular volume cycle. This study was undertaken to examine the relationship between the ventricular electrogram and conductance volume measurements in a physical model of the left ventricle without parallel conductance. In addition, we sought to determine how this relationship was influenced by changes in electrical resistance across the model ventricle.

2. METHODS

2.1. Model Ventricle

This study used the physical model of the left ventricle previously developed by this aboratory [9] and described in the accompanying paper [10]. In summary, this consisted of an ellipsoid latex balloon enclosed in a pressurised Perspex chamber. The chamber was filled with distilled water and hydraulically pressurised with an intra-aortic balloon pump (IABP; Datascope Medical Co. Ltd, Huntingdon, UK) connected to two 25ml intra-aortic balloon catheters in parallel. A patient simulator (Bioma Research Inc, Quebec, Canada) was used to trigger the IABP console at a predetermined rate (60 beats·min^{-1}).

Inflation of the intra-aortic balloon catheters displaced the stroke volume (SV, 50ml) from the model ventricle through a 2/2-way solenoid valve (Bürkert GmbH, Ingelfingen, Germany) into a calibrated measuring cylinder at the top of the model ventricle. Deflation of the IABP balloon catheters caused ventricular pressure to fall, which allowed the latex balloon to refill. Electronic circuitry was used to control the opening and closure times of the solenoid valve in order to simulate different contraction patterns.

The latex balloon was 13 cm in length and had a maximal volume at atmospheric pressure of 500ml. The balloon was filled with 385–500ml of buffered saline solution (V) at room temperature. The saline concentration was varied between 0.18 – 1.57%. The resistivity (ρ; conductivity^{-1}) of these solutions was measured before each test using a dedicated measuring cuvette (CD Leycom, Zoetermeer, The Netherlands). The resistivity ranged between 37 ± 0.8 to 330 ± 0.5 Ω cm.

Ventricular depolarisation was simulated using a fixed output ECG signal, ECG_{input}, from the patient simulator (maximum 150.4 ± 0.1 mV, minimum –35.8 ± 0.2 mV). This was connected, via a resistor (200 Ω), to a dipole within the latex balloon. This dipole consisted of two copper ring electrodes (diameter 30mm, depth 5mm and thickness 1mm) that were positioned perpendicular to the long-axis of the balloon. The in-series resistor was adjusted so that the signal range of the intracavitary electrogram (*ECGc*) in the model ventricle was equivalent to that observed in vivo (~1mV).

In vivo, the intracavitary electrogram primarily reflects the pattern of ventricular depolarisation and repolarisation in the endocardium of the ventricle [11]. The position of the endocardium relative to the long-axis of the ventricle will vary as the ventricular volume changes during the cardiac cycle. This effect was simulated by altering the distance between the two electrodes of the dipole. The distance between the two electrodes (*D*) was varied between 3 cm and 11.5 cm such that both electrodes remained equidistant from the centre of the balloon.

3. CONDUCTANCE CATHETER

The principle of the conductance catheter technique for measuring LV volume has been described elsewhere [8]. The details of the conductance catheter used in this study are described in the accompanying paper [10]. The catheter measured seven time-varying segmental conductance signals, $G_i(t)$. As parallel conductance is negligible in our model, the total conductance volume, $Q(t)$, was determined using the following formula:

$$Q(t) = L^2 \cdot \rho \cdot [\sum_{i=1}^{7} G_i(t)] \quad (1)$$

where ρ is the blood resistivity and L is the inter-electrode distance. The dimensionless calibration coefficient, $\alpha_{V(t)}$ was calculated from conductance volume measurements without the ECG signal by dividing the conductance-derived volume measurement by the absolute volume at either end-diastole or end-systole:

$$\alpha_{V(t)} = \frac{Q_{ED}}{V_{ED}} \text{ or } \frac{Q_{ES}}{V_{ES}} \quad (2)$$

We have previously demonstrated that the calibration coefficient, $\alpha_{V(t)}$ varies as a non-linear function of the absolute ventricular volume [10]. We used this non-linear $\alpha_{V(t)}$-volume relation to calibrate conductance volume measurements:

$$V_g(t) = \frac{1}{\alpha_{VV}} \cdot Q(t) \quad (3)$$

where α_{VV} is the $\alpha_{V(t)}$-volume relation [10].

Instantaneous pressure within the model ventricle was measured using a high-fidelity solid-state micromanometer laterally positioned between electrodes 5 and 6 within the conductance catheter. This pressure signal was amplified using a combined amplifier-interface unit (PCU-2000; Millar Instruments, Houston, TX, USA) and statically calibrated using a separate fluid-filled catheter-manometer system.

The cavitary electrogram, *ECGc* was measured as part of the conductance catheter technique. The conductance signal between electrodes 5 and 6 was measured, amplified and filtered using a second-order filter with a high cut-off frequency (–3dB; 125 Hz) in order to derive the *ECGc* [personal communication; CD Leycom, Zoetermeer, The Netherlands].

4. ECG INTEREFERENCE

The conductance signal was measured either with, $G_i^+(t)$, or without the ECG signal, $G_i^-(t)$. ECG interference, $G_{ECG}(t)$ was calculated as the difference between these two conductance signals:

$$G_{ECG}(t) = \sum_{i=1}^{7} G_i^+(t) - \sum_{i=1}^{7} G_i^-(t) \quad (4)$$

The difference between calibrated volume measurements made with, $V_g^+(t)$ and without the ECG signal, $V_g^-(t)$ was expressed as a percentage of the $V_g^-(t)$ signal:

$$\Delta V_{ECG} = \frac{[V_g^+(t) - V_g^-(t)]}{V_g^-(t)} \quad (5)$$

5. EXPERIMENTAL DATA

Analogue signals representing 7 segmental conductance signals, the pressure within the model ventricle, and both ECG_{input} and *ECGc* were all digitised at 12-bit accuracy and a sample frequency of 250 Hz. End-diastole and end-systole were retrospectively identified. End diastole was defined as the R wave on the ECG and end-systole was defined as the point immediately prior to IABP circuit deflation.

The effect of intracavitary volume (*V*), resistivity of the saline solution (ρ) and inter-electrode distance (*D*) on the intracavitary electrogram and conductance volume measurements were examined in turn. This involved a series of experiments, in which one variable was altered incrementally while the other two variables remained unchanged. This process was repeated until the data from the entire range was obtained (see above).

Each experiment was conducted under steady-state conditions and data from 5 consecutive cycles were analysed. The average within-experiment standard deviation was 0.71 ml and, at its worst, this represented <0.5% of the total conductance volume. All subsequent analyses were therefore based on the average data from each experiment.

6. DATA ANALYSIS

Data were analysed using SPSS for Windows (v12, SPSS Inc., Chicago, Il, USA). Data are expressed as mean ± SD and comparative analyses have been made using the t-test. The relationship between the ECG_{input}, *ECGc* and the ECG interference pattern was evaluated by least squares linear regression based on fractional polynomials of the data. The intracavitary volume, resistivity of the saline solution and the inter-electrode distance were included as covariables in the regression analyses. The coefficients of the linear regression analyses, in both the overall and covariance analyses, are expressed as mean ± standard error and a probability, $P<0.05$, was taken to represent statistical significance. The "goodness of fit" of the prediction equation was assessed as the square of the correlation between dependent and significant independent variables.

7. RESULTS

7.1. Comparison between ECG Signal and Cavitary Electrogram

The ECG signal, ECG_{input}, consisted of P, Q, R, S and T deflections that resembled the normal lead II electrocardiogram (**Figure 2A**). The P wave was 84 ms in duration with a peak of 6.2 mV (40 ms). The P wave represented approximately half the PR interval (156 ms). The positive QRS complex had an overall duration of 84 ms, with a maximum at 192 ms (150.4 mV) and two minima at 164 ms (–26.4 mV) and 224 ms (–34.1 mV). The ST segment was isoelectric (–8.2 mV) and 36 ms long. The duration of the T wave was 236 ms, with a peak of 41.4 mV (408 ms). Finally, the QT interval and TP segment were both isoelectric and 120 ms and 468 ms long, respectively.

The *ECGc* signal resembled the ECG_{input} signal turned upside down, with an inverted P wave, an rSR′ wave and an inverted T wave (**Figure 2B**). The overall relationship between the two signals was best approximated by a mathematical model in which the *ECGc* signal was inversely proportional to the ECG_{input} signal (r2=0.74, $P<0.001$):

$$ECGc(t) = \beta_0 - \beta_1 \cdot ECG_{input}(t) \quad (6)$$

Although the ECG_{input} and *ECGc* signals were similar, they were not identical There were differences in timing and amplitude of the two signals. The *ECGc* P and S wave minima and the R wave peak occurred either synchronously or within one data point (i.e. 4 ms) of the corresponding points on the ECG_{input} signal. By contrast, the *ECGc* R′ wave peak and the T wave misnimum were 12 ms and 32 ms earlier than the corresponding points on the ECG_{input} signal. The *ECGc* R′ wave peak was also disproportionately pronounced compared to the corresponding S wave of the ECG_{input} signal. The *ECGc* R′-S wave ratio was –0.51 ± 0.03 whereas the ECG_{input} S-R wave ratio was –0.17 ± 0.02 ($P<0.001$). In addition, the *ECGc* did not accurately reproduce the isoelectric phases of the ECG_{input} signal. During the PR, R′T and TP segments, the *ECGc* signal was initially elevated and decreased progressively towards the baseline signal.

In the covariance analyses, the intercept value ($\beta 0$) varied as a linear function of inter- electrode distance ($P<0.05$), but was not affected by variation in the other two factors. By contrast, the linear regression coefficient ($\beta 1$) varied as the inverse function of intracavitary volume and as a direct function of inter-electrode distance and resistivity of the solution (all $P<0.05$).

When these effects are combined, the relationship between *ECGc* and ECG_{input} was influenced by the inter-electrode distance (intercept value) and by the total resistance of the volume conductor (**Eq. 7**; $r^2=0.65$):

$$ECGc(t) = \beta_2 + (\beta_3 \cdot D) + \left[\beta_4 \cdot \left(\frac{D \cdot \rho}{V} \right) \cdot ECG_{input}(t) \right] \quad (7)$$

where $\beta_2 = 3.66 \pm 0.01$; $\beta_3 = 7.29 \cdot 10^{-3} \pm 1.53 \cdot 10^{-3}$ and $\beta_4 = -7.00 \cdot 10^{-3} \pm 0.06 \cdot 10^{-3}$ ($P<0.05$ for each coefficient).

7.2. Comparison between ECG Interference and ECG Input Signals

The ECG interference signal, $G_{ECG}(t)$ was characterised by a low amplitude biphasic P wave; a high amplitude equiphasic qRSr′ complex; and a low-amplitude biphasic T wave (**Figure 2C**). Each phase of the $G_{ECG}(t)$ signal was synchronous with the P wave, QRS complex and T wave of the ECG signal, respectively.

The amplitude of the $G_{ECG}(t)$ has been described as a percentage of the maximum $G_{ECG}(t)$ signal from the isoelectric line. The $G_{ECG}(t)$ P wave had a sine wave-like appearance with an initial upward deflection immediately followed by a downward deflection of comparable duration and amplitude. The maximum and minimum $G_{ECG}(t)$ P wave signals were 4.0 ± 2.6% (28 ms) above and 4.4 ± 2.6% (64 ms) below the isoelectric line. The spiked wave $G_{ECG}(t)$ qRSr′ complex had two maxima at 180 ms (R wave; 100 ± 3%) and 232 ms (r′ wave; 22 ± 3%) and two minima at 160 ms (q wave; –24 ± 3%) and 204 ms (S wave; –107 ± 3%). The $G_{ECG}(t)$ q and R waves occurred 32 ms and 12 ms before the ECG_{input} R wave whereas the $G_{ECG}(t)$ S and r′ waves occurred 12 ms and 40 ms after the ECG_{input} R wave. The $G_{ECG}(t)$ T wave had a similar overall appearance to the $G_{ECG}(t)$ P wave with an initial upward deflection immediately followed by an equivalent downward deflection. The maximum and minimum $G_{ECG}(t)$ T wave

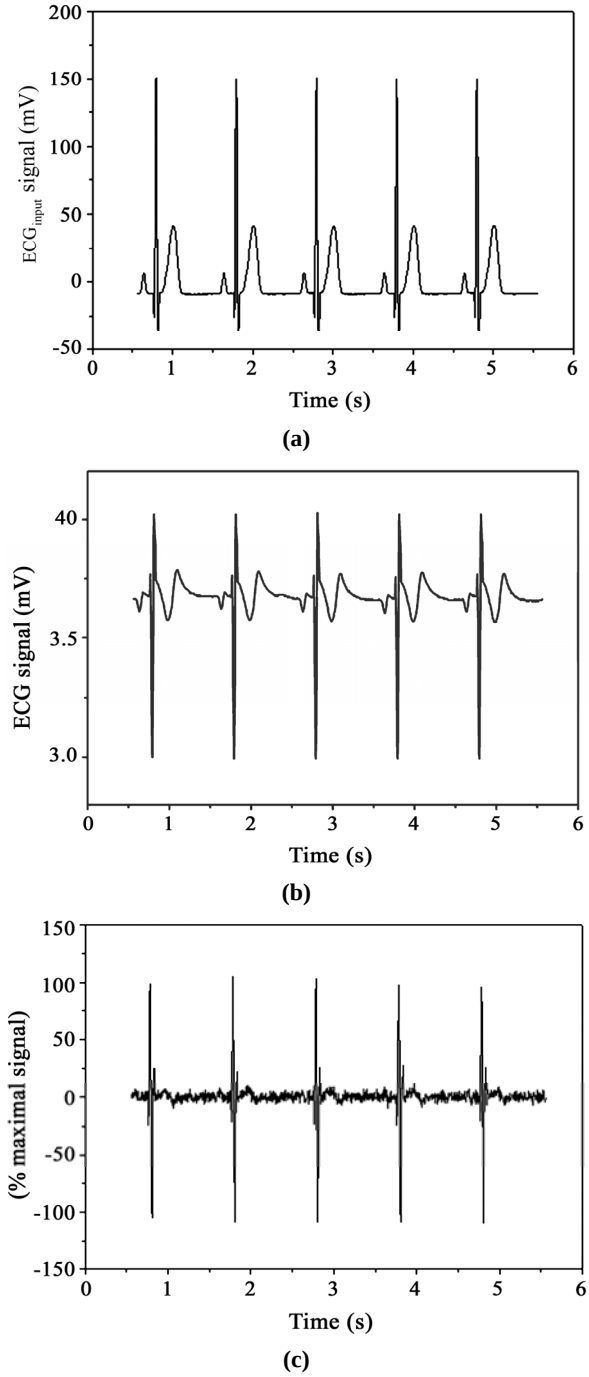


Figure 2. The ECG signal (ECG_{input}; A), cavitary electrogram ($ECGc$; B) and ECG interference ($G_{ECG}(t)$; C) signals versus time.

signals were 4.4 ± 2.6% (352 ms) and 4.8 ± 2.6% (452 ms) below the isoelectric line and occurred 80 ms and 180 ms after the start of the ECG_{input} T wave, respectively.

The relationship between $G_{ECG}(t)$ and ECG_{input} was well approximated (r^2=0.92, P<0.001) by a regression equation in which the interference signal varied proportionally to the first-derivative of ECG_{input} with respect to time:

$$G_{ECG}(t) = \beta_0 + \beta_1 \cdot dECG_{input}(t) \quad (8)$$

In the covariance analyses, the intercept (β_0) values varied as a linear function of the interelectrode distance (P<0.05), but was not affected by variation in the other two factors. By contrast, the linear regression coefficients (β_1), varied as a function of the intracavitary volume; and as the inverse function of both the inter-electrode distance and the resistivity of the solution (all P<0.05).

Overall, the relationship between $G_{ECG}(t)$ and the first derivative of ECG_{input} varied as a function of the inter-electrode distance and the conductivity of the volume conductor (**Eq. 9**; r^2=0.88):

$$G_{ECG}(t) = (\beta_0 \cdot D) + \left[\beta_1 \cdot \left(\frac{V}{D \cdot \rho} \right) dECG_{input}(t) \right] \quad (9)$$

where β_0 = $1.49 \cdot 10^{-3}$ ± $0.39 \cdot 10^{-3}$ and β_1 = $6.07 \cdot 10^{-3}$ ± $0.38 \cdot 10^{-3}$ (P<0.05 for both coefficients).

7.3. ECG Interference and Calibrated Conductance Volume Measurements

For the purposes of this simulation, the inter-electrode distance was assumed to change in accordance with the instantaneous volume within the latex balloon. The inter-electrode distance was estimated as the maximal short-axis diameter of the spheroid, which varied from7.0 cm (V_{ES} = 335 ml) to 8.6 cm (V_{ED} = 500 ml). Calibrated conductance volume measurements with the ECG signal, $V_g^+(t)$ were compared against synchronous calibrated conductance volume measurements without the ECG signal, $V_g^-(t)$.

A representative example of calibrated conductance volume measurements with and without the ECG signal is illustrated in **Figure 3**. The $V_g^-(t)$ signal had a smooth, sinusoidal pattern that varied throughout the model heart cycle. The $V_g^+(t)$ signal was broadly similar, but had an additional spiked-wave pattern that coincided with the simulated ventricular depolarisation. The difference between the two ventricular volumes measurements, ΔV_{ECG} during this phase of the cardiac cycle varied between –12% (i.e. an underestimation) and +9%. By contrast, the difference during the remainder of the cardiac cycle varied only slightly from –0.3 to +0.9%.

The pressure-volume loop obtained with the $V_g^-(t)$ signal had a quadrilateral shape with four distinct phases (**Figure 4**). End-diastole and end-systole were each identifiable as the single pressure-volume point at the lower right-hand and upper left-hand corners, respectively. The ECG interference pattern altered the shape of the pressure-volume loop, primarily affecting the late filling and early isovolumic contraction phases. As a result, end-diastole

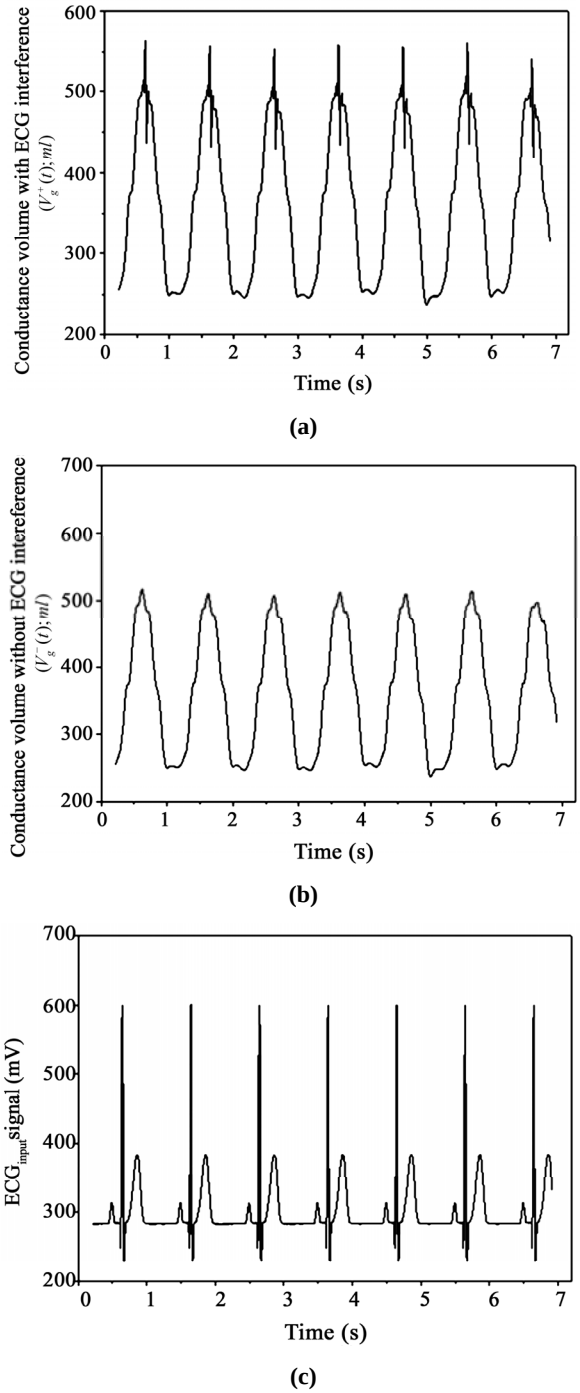


Figure 3. Calibrated conductance volume measurements versus time. Conductance volume measurements were made with ($V_g^+(t)$; A) and without ECG interference ($V_g^-(t)$; B) versus time. The ECG signal (ECGinput) has been plotted (C) for comparison.

could not be reliably identified (**Figure 4**).

Comparable results were obtained under all experimental conditions. Increasing the end-diastolic volume from 385 ml to 500 ml did not significantly change the discrepancy between the two ventricular volume measurements. While increasing the resistivity from 37 Ω·cm to 330 Ω·cm increased the median measurement discrepancy slightly, from +0.2% to +1.8%, it did not significantly alter the maximal range of the discrepancy.

8. DISCUSSION

The conductance catheter technique is an established method that enables continuous volume measurements based on the electrical conductance of the intraventricular blood pool. This study has demonstrated that other electrical signals within the ventricular cavity alter the measured conductance. This produces a conductance signal "artefact", which varies as a function of the first-derivative of the additional electrical signal. This artefact represents a novel and additional source of error that potentially affects the accuracy of ventricular volume measurements made using the conductance catheter technique.

Ventricular depolarisation and repolarisation cause a measurable electrical signal within the ventricular cavity [12]. In the present study, a simulated ventricular electrogram produced a biphasic signal with a highamplitude spiked wave pattern that coincided with the QRS complex and a comparatively low-amplitude sine wave- like pattern during the T wave. This signal was associated with a conductance volume measurement error that ranged between a 12% volume underestimation to a 9% volume overestimation. The entire range of this measurement discrepancy occurred within a 24 ms period during simulated ventricular depolarisation. By contrast, simulated repolarisation was associated with a small, clinically unimportant measurement error. The precise pattern, will vary with the morphology of ventricular electrogram [12,13,14]. The precise ECG interference pattern, $G_{ECG}(t)$ will vary with the morphology of ventricular electrogram [12,13,14]. Nevertheless, these in

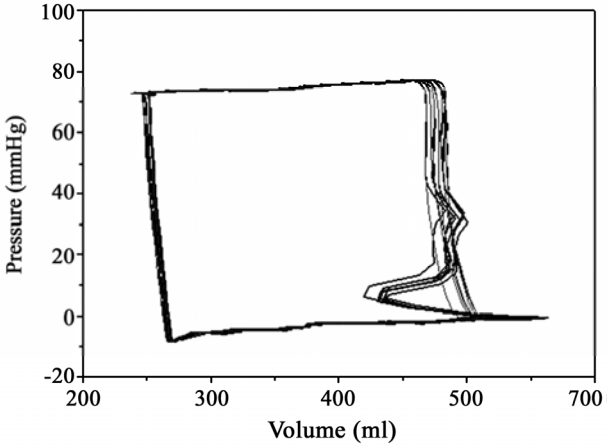


Figure 4. Pressure-conductance volume loop from the model ventricle. Conductance volumemeasurements were made with ($V_g^+(t)$; black line) or without ECG interference ($V_g^-(t)$; grey dashed line).

vitro findings are consistent with our previously unreported clinical findings.

The haemodynamic events during the cardiac cycle are best displayed by plotting the instantaneous left ventricular pressure versus volume [15]. Under steady-state conditions, this pressure-volume loop has a quadrilateral shape where each side represents one of four functional distinct phases: filling, isovolumic contraction, ejection and isovolumic relaxation. End-diastole and end-systole are identifiable as the single pressure-volume points in the lower right-hand and upper left-hand corners, respectively. However, ventricular depolarisation overlaps the rapid rise in intraventricular pressure that marks the onset of ventricular systole. The conductance signal artefact identified in this study meant that end-diastole could no longer be reliably identified on the pressure-conductance volume loop alone.

End-diastole may alternatively be defined using the surface electrocardiogram as the onset of the QRS complex [16]; the R wave peak [17]; or up to 40 ms after the R wave peak [18]. End-diastole may also be defined as the R wave peak on the ventricular electrogram. In our experience, this time-point occurs synchronously with the onset of ventricular systole [19]. However, conductance volume measurements at all of these time-points will be variably affected by the conductance signal artefact such that end-diastolic volume cannot be accurately measured using the conductance catheter technique. This in turn means that indices of ventricular function that are based on EDV, such as cardiac output, ejection fraction together with the quantitative assessment of ventricular compliance, will be adversely affected as a consequence of the conductance signal artefact.

9. STUDY LIMITATIONS

The limitations of the physical model have been described previously [10]. Electrical activity within the ventricle was represented using a fixed dipole within the ventricular cavity. This comparatively simple model enabled characterisation and quantification of a new conductance measurement error. However, the model did not include any representation of the ventricular wall and the effect of parallel conductance was not examined. A moving dipole or multiple dipoles within an artificial ventricular wall would also have provided a more physiological model.

10. CONCLUSIONS

This study has demonstrated that the accuracy of these conductance volume measurements is adversely affected by other electrical signals, such as the ventricular electrogram. The ventricular electrogram produced a clinically important volume measurement that meant end-diastole could neither be precisely identified nor accurately measured. These original findings have important implications for the quantitative assessment of ventricular function and, in particular the assessment of chamber compliance.

11. ACKNOWLEDGEMENTS

Simon McGuirk was supported by a British Heart Foundation Junior Research Fellowship (FS/03/102).

REFERENCES

[1] Baan, J., Aouw Jong, T. T., Kerkhof, P. L., Moene, R. J., van Dijk, A. D., van der Velde, E. T., and Koops, J. (1981) Continuous stroke volume and cardiac output from intraventricular dimensions obtained with an impedance catheter, Cardiovascular Research, **15**, 328–334.

[2] Mur, G. and Baan, J., (1984) Computation of the input impedances of a catheter for cardiac volumetry, IEEE Transactions on Biomedical Engineering, **31**, 448–453.

[3] Kass, D. A., (1992) Clinical evaluation of left heart function by conductance catheter technique, European Heart Journal, **13(Suppl E)**, 57–64.

[4] Wu, C. C., Skalak, T. C., Schwenk, T. R., Mahler, C. M., Anne, A., Finnerty, P. W., Haber, H. L., Weikle II, R. M. and Feldman, M. D., (1997) Accuracy of the conductance catheter for measurement of ventricular volumes seen clinically: Effects of electric field homogeneity and parallel conductance, IEEE Transactions on Biomedical Engineering, **44**, 266–277.

[5] Salo, R. W., (1989) The theoretical basis of a compartmental model for the determination of volume by impedance, Automedica, **11**, 299–310.

[6] Wei, C. L., Valvano, J. W., Feldman, M. D., and Pearce, J. A., (2005) Nonlinear conductance-volume relationship for murine conductance catheter measurement system, IEEE Transactions on Biomedical Engineering, **52**, 654–661.

[7] Salo, R. W., Wallner, T. G., and Pederson, B. D., (1986) Measurement of ventricular volume by intracardiac impedance: Theoretical and empirical approaches, IEEE Transactions on Biomedical Engineering, **33**, 189−195.

[8] Baan, J., van der Velde, E. T., de Bruin, H. G., Smeenk, G. J., Koops, J., van Dijk, A. D., Temmerman, D., Senden, J., and Buis, B., (1984) Continuous measurement of left ventricular volume in animals and humans by conductance catheter, Circulation, **70**, 812–823.

[9] Al-Khalidi, A. H., Townend, J. N., Bonser, R. S., and Coote, J. H., (1998) Validation of the conductance catheter method for measurement of ventricular volumes under varying conditions relevant to cardiac surgery, American Journal of Cardiology, **82**, 1248–1252.

[10] McGuirk, S. P., Barron, D. J., Ewert, D., and Coote, J. H., (2009) Calibrating volume measurements made using the dual-field conductance catheter, Journal of Biomedical Science and Engineering, In Press.

[11] Durrer, D., Van Dam, R. T., Freud, G. E., Janse, M. J., Meijler, F. L., and Arzbaecher, R. C., (1970) Total excitation of the isolated human heart, Circulation, **41**, 899–912.

[12] Taccardi, B., Arisi, G., Macchi, E., Baruffi, S., and Spaggiari, S., (1987) A new intracavitary probe for detecting the site of origin of ectopic ventricular beats during one

cardiac cycle, Circulation, **75**, 272–281.

[13] Khoury, D. S. and Rudy, Y. (1992) A model study of volume conductor effects on endocardial and intracavitary potentials, Circulation Research, **71**, 511–525.

[14] Derfus, D. L., Pilkington, T. C., Simpson, E. W., and Ideker, R. E., (1992) A comparison of measured and calculated intracavitary potentials for electrical stimuli in the exposed dog heart, IEEE Transactions on Biomedical Engineering, **39**, 1192–1206.

[15] Burkhoff, D., Mirsky, I., and Suga, H., (2005) Assessment of systolic and diastolic properties via pressure-volume analysis: A guide for clinical, translational, and basic researchers, American Journal of Physiology-Heart and Circulatory Physiology, **289**, H501–H512.

[16] Roelandt, J. and Gibson, D. G., (1980) Recommendations for standardization of measurements from M-mode echocardiograms, European Heart Journal, **1**, 375–378.

[17] Sugeng, L., Mor-Avi, V., Weinert, L., Niel, J., Ebner, C., Steringer-Mascherbauer, R., Schmidt, F., Galuschky, C., Schummers, G., Lang, R. M., and Nesser, H-J., (2006) Quantitative assessment of left ventricular size and function: Side-by-side comparison of real-time three-dimensional echocardiography and computed tomography with magnetic resonance reference, Circulation, **114**, 654–661.

[18] Møgelvang, J., Thomsen, C., Mehlsen, J. Bräckle, G., Stubgaard, M., and Henriksen, O., (1986) Evaluation of left ventricular volumes measured by magnetic resonance imaging, European Heart Journal, **7**, 1016–1021.

[19] Al-Khalidi, A. H., Lewis, M. E., Townend, J. N., Bonser, R. S., and Coote, J. H., (2001) A novel and simple technique to allow detection of the position of the R-waves from intraventricular pressure waveforms: application to the conductance catheter method, IEEE Transactions on Biomedical Engineering, **48**, 606–610.

J. Biomedical Science and Engineering, 2009, 2, 499-505 **JBiSE**
doi: 10.4236/jbise.2009.27072 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Effect of LDL-apheresis on plasma lipids, chitotriosidase and anti-oxLDL antibodies in heterozygous familial hypercholesterolemia

Maria Musumeci[1], Francesco Pappalardo[2,3], GianCarlo Tonolo[4], Fernando Torrisi[5], Francesca Gullo[3], Salvatore Musumeci[6]

[1]Department of Hematology, Oncology and Molecular Medicine, Italian National Institute of Health, Rome, Italy; [2]Institute for Computing Applications "M. Picone", National Research Council (CNR), Rome, Italy; [3]University of Catania, Catania, Italy; [4]Diabetology Unit, Azienda Sanitaria Locale 2 Olbia, Italy affiliated to Department of Clinical Sciences, Medical Genetics Unit, Lund University, Malmö, Sweden; [5]Institute of Medical and Environmental Research (IRMA), Acireale (Catania), Italy; [6]Department of Neurosciences and Mother to Child Sciences, University of Sassari, and Institute of Biomolecular Chemistry, National Research Council (CNR), Li Punti, Sassari, Italy.
Email: smusumeci@tiscalinet.it

Received 16 June 2009; revised 10 July 2009; accepted 17 July 2009.

ABSTRACT

Forty four consecutive subjects aged 29-58 years (21 males and 23 females) with a clinical diagnosis of heterozygous familial hypercholesterolemia periodically treated every 30 days with LDL-apheresis for statin resistance, were enrolled in this study. A lipid profile was obtained immediately before starting LDL-apheresis, a second profile was obtained within four hours after LDL-apheresis. Chit activity and anti-oxLDL levels were determined with appropriate methods in all patients before and after LDL-apheresis. Total cholesterol, LDL-cholesterol, HDL-cholesterol and triglycerides decreased significantly after LDL-apheresis, while the variations of Chit activity and anti-oxLDL were not significant after LDL-apheresis. The correlation between Chit and total cholesterol was negative (r= –0.44 and –0.50 respectively) before and after LDL-apheresis as between Chit and LDL-cholesterol (r= –0.45 and –0.55 respectively). Anti-oxLDL concentration before and after LDL-apheresis positively correlated with Chit activity (r= 0.52 and r = 0.63 respectively), negatively with total cholesterol (r= –0.33 and r = –0.35 respectively) and with LDL (r = –0.32 and r = –0.21 respectively). We think that removing LDL with LDL-apheresis the anti-oxLDL/oxLDL ratio could increase and the excess of anti-oxLDL could induce macrophage activation through the surface Fc receptors. Alternatively with high levels of LDL-cholesterol, the deposition of foam cells represent the characteristic evolution of atherosclerosis process. Macrophage activation in the heterozygous familial hypercholesterolemia could represent an attempt for re-modeling the vessel wall, reducing the growth of lipid plaques.

Keywords: LDL-Apheresis; Heterozygous Familial Hyper-Cholesterolemia; Lipids; Chitotriosidase; Anti-oxLDL Antibodies; Sardinia

1. INTRODUCTION

Familial hypercholesterolemia (FH) is a genetic alteration of lipoprotein metabolism caused by defects in the low density lipoprotein receptor (LDLR) [1]. High LDL levels, secondary to the LDLR homozygous defect, are associated to significant increase of oxidized LDL (oxLDL), which removed from circulation lead to massive lipid accumulation, foam cell formation in endothelial wall, often tendom xanthomas (TX) and corneal arcus [2]. Also the heterozygous FH (heFH), shows precocious coronary heart disease before 65 years old, if they are not treated [3]. Then FH represent a paradigmatic example of atherosclerosis produced by oxLDL accumulation and a model to study the role of macrophage activation in atherosclerosis process [1]. Generated oxLDL induces an immune response with production of anti-oxLDL antibodies and macrophage cells could remove from circulation immune complexes anti-oxLDL/oxLDL through the Fc receptor for antibodies [4]. Shoji *et al* 2000 [5] found a inverse correlation among anti-oxLDL and oxLDL in healthy individuals supporting the hypothesis that this mechanism is also operating in condition where oxLDL are

stably low. This mechanism could be not sufficient to protect from foam cells generation in presence of a defect in LDLR gene, when the levels of oxLDL are very high [6].

Chitotriosidase (Chit) is one of the most quantitatively represented marker of macrophage activation, such as occurs in Gaucher disease, sarcoidosis, nonalcoholic liver disease and atherosclerosis [7,8,9]. Plasma Chit activity has been associated with both the extension and prognosis of atherosclerotic vascular lesions in humans [10,11,12] and its phagocyte-specific expression supports a relevant role in innate immunity [13].

Considering the importance of anti-oxLDL in the pathogenesis of atherosclerosis lesions [14] and the involvement of Chit activity in the evolution of atherosclerotic vascular lesions [10,11], we hypothesize a relation among these two factors. The objective of this study is to establish the relationship between lipid parameters, Chit activity and anti-oxLDL levels in a group of subjects with clinically and genetically defined heFH (total cholesterol consistently more than 400 mg/dl) before and after LDL-apheresis treatment. LDL-apheresis represents an effective therapy in heFH patients, who had no response to highest doses of statin drugs, and could restore the physiological mechanism of anti-oxLDL/oxLDL immunocomplexes clearance altered by LDLR genetic defect [15].

2. MATERIAL AND METHODS

2.1. Study Subjects

Consecutive 44 subjects from Sardinia (Italy) aged 29–58 years (21 males and 23 females) with a clinical diagnosis of heFH were treated periodically with LDL-apheresis. The diagnosis of heFH was determined genetically in all patients [2]. They started LDL-apheresis treatment because their previous lipid lowering therapy (statin and benzafibrate) did not reduce the total and LDL cholesterol.

A combination of discontinuous blood centrifugation (MCS 3p Haemonetics Corp., Braintree, MA, USA) and 2 steps membrane differential filtration were performed at interval of 30 days. Clinical data, history of prior cardiovascular disease (CVD) at early onset, demographic and anthropometric measurements, and an accurate physical examination in search of tendon xanthomas (TX) were obtained from each subject.

Informed consent was obtained from all subjects and the ethical committee from each institution approved this study.

2.2. Lipid Concentrations

To obtain a baseline lipid profile, overnight fasting blood was drawn immediately before starting LDL-apheresis. A second profile was obtained within four hours after LDL-apheresis. Total cholesterol and triglyceride levels were measured with standardized enzymatic methods. HDL cholesterol was measured by precipitation methods and LDL cholesterol was estimated with the Friedewald's formula, since no patient had triglycerides over 300 mg/dl. Lipoprotein (a) was determined in immunonephelometry with specific antibodies (New Scientific Company S.r.L., Cormano (MI), Italy).

2.3. Chitotriosidase Enzyme Assay

Chitotriosidase enzyme assay was based on the method described by Hollak *et al.* 1994 [7], with minor modifications. Briefly, chitotriosidase activity was determined by incubating 5 µL of plasma with 100 µL of 22 mmol/L 4-methylumbelliferyl-ß-d-N,N′,N″triacetyl-chitotrioside fluorogenic substrate (Sigma-Aldrich S.r.L. Milano, Italy, catalogue M 5639) in McIlvain buffer (100 mmol/L citric acid and 200 mmol/L sodium phosphate, pH 5.2) for 15 minutes at 37°C. The reaction was stopped by using 2 ml of 0.5 mol/L Na_2CO_3-$NaHCO_3$ buffer, pH 10.7. The substrate hydrolysis by Chit produces the fluorescent molecule 4-methylumbelliferone, which was quantified with a Hitachi 2500 fluorometer, excitation at 366 nm and emission at 446 nm, and compared with a standard 4-methylumbelliferone calibration curve. Chit activity was expressed as nanomoles of substrate hydrolyzed per hour per milliliter of reaction mixture. Plasma Chit activity was measured by duplication and three QC samples from healthy adults were added in every set of determinations. The coefficient of variation was less than 5% in all cases.

2.4. Oxidized Low-Density Lipoprotein Antibodies

An enzyme-linked immunosorbent assay (ELISA) for the detection and quantification of IgG antibodies to oxidized low-density lipoprotein (oxLDL) in human plasma was used (ImmuLisa™, IMMCO Diagnostics, Buffalo, NY, USA). The intensity of the color changes, proportionally to the antibodies concentration, was read as absorbance at 405 nm. Three QC samples from healthy adults were also added in each plate containing a calibration curve. The absorbance values on native LDL were subtracted from the absorbance obtained on oxLDL for control, calibrators and specimens. The concentration of anti-oxLDL was determined from the calibration curve and the results are expressed in Enzyme Units per milliliter (EU/ml).

2.5. Apolipoprotein E Genotyping

DNA from patients isolated by peripheral blood cells was used in a polymerase chain reaction (PCR) and Apolipoprotein E genotypes were determined by HhaI digestion as the methods described by Hixon and Vernier [16], modified by Tsukamoto *et al* [17].

2.6. LDLR Genetic Analyses

For LDLR gene analysis the entire gene was sequenced. Mutations in the LDLR gene causing FH were classified when possible, as receptor-defective or receptor-negative on the basis of the residual LDLR activity. Mutation leading to a frame-shift and/or a truncated receptor were considered as receptor-negative [2,6].

2.7. Chitotriosidase Polymorphism

DNA from patients were used as template in subsequent PCR reactions. The duplication mutation analysis was performed using specific primers [Chs9 (AGCTATCT GAAGCAGAAG) and Chas8 (GGAGAAGCCGGCA-AAGTC)] and two fragments of 75 and 99 base pairs were amplified from the wild and mutant CHIT gene, respectively. Electrophoresis in Metaphore gel (4%), allowed the detection of both fragments.

3. STATISTICAL METHODS

The data are expressed as mean values with standard deviation (SD) for variables with normal distribution and as medians and range for variables with a skewed distribution. Statistical differences were computed by using the Student's t test or the Mann-Whitney U-test, respectively. Correlation and statistical analyses were performed with SSPS software (version 13.0), with significance set at $P<0.05$.

4. RESULTS

In **Table 1**, we show the clinical and biochemical characteristics of the 44 heFH subjects, including also the LDLR and CHIT genotype polymorphisms. No signi- ficant difference, among men and women, regarding age, body mass index, blood pressure was found and neither presence of hypertension nor diabetes. Differences were found in smoking habit, premature CVD, and age of first CVD event, being more smoker among the men group, who had more premature CVD, with minor age at the first event. Mean values of total cholesterol, LDL- cholesterol, triglycerides in the last few months were elevated in all subjects with no significant difference between men and women (see **Table 1**). Only men had significantly lower HDL cholesterol levels compared to women. No Chit deficient patients was found at the genotype analysis, 27 (61.4 %) were wild/wild and 17 (38.6%) wild/mut. Chit activity was found elevated before LDL-apheresis, without difference between males (11.94±8.68 nmol/ml/h) and females (12.26±7.70 nmol/ml/h), considering that 7.3±1.9 nmol/ml/h activity was found in healthy control of the same age in our laboratory. Also the anti-oxLDL were found elevated before LDL-apheresis in males (38.20±17.64 UI/ml) and in females (55.80±38.97), considering that < 20 UI/ml was found in healthy control of the same age in our laboratory. The difference between males and females was not significant (P = 0.0646). In **Table 2** are reported the levels of total cholesterol, LDL-cholesterol, HDL-cholesterol, triglycerides, Chit and anti-oxLDL before and after LDL-apheresis, considering separately males and females. All lipid parameters significantly decreased after LDL-apheresis ($P<0.0001$). The reduction of Chit activity after LDL-apheresis was not significant as well as the

Table 1. Clinical and laboratory characteristics of 44 heterozygous FH patients (21 men and 23 women).

	Men N=21	Women N=23	p
Age, years	35.2 (29-54)	36.9 (32-58)	N.S.
Body mass index, kg/m2	28.0±3.68	28.5±4.84	N.S.
Smoking status: Never, n (%) Current, n (%) Former, n (%)	 7 (33.3) 3 (14.3) 11 (52.4)	 15 (65.2) 5 (21.7) 3 (13.0)	0.0001
Smoking, packs x years	26.1±10.0	26.5±7.42	N.S.
Systolic blood pressure, mmHg	130.3±17.1	129.4±19.1	N.S.
Diastolic blood pressure, mmHg	77.4±10.3	76.8±11.3	N.S.
Hypertension, n (%)	4 (19.0)	5 (21.7)	N.S.
Diabetes, n (%)	1 (4.8)	1 (4.3)	N.S.
Premature CVD, n (%)	8 (38.1)	6 (26.1)	0.016
Age of first CVD event, years	45.4±6.16	49.5±7.53	0.0578
Family history premature CVD, n. (%)	10 (47.6)	11 (47.8)	N.S.
Tendon xanthomas, n. (%)	8 (38.1)	11 (47.8)	N.S.
Total cholesterol, mmol/L	8.99±4.03	8.29±1.53	N.S.
Triglycerides, mmol/L	1.46±1.24	1.04±0.53	N.S.
LDL cholesterol, mmol/L	7.25±1.20	6.82±1.33	N.S.
HDL cholesterol, mmol/L	1.33±0.28	1.48±0.35	0.0228
Chitotriosidase (nmol/ml/hr)	11.94±8.68	12.26±7.70	N.S.
Anti-oxLDL (UI/ml)	38.20±17.64	55.80±38.97	0.0646
ApoE genotype E3/E3, n (%) E3/E4, n (%) E3/E2, n (%) E4/E4, n (%) E2/E2, n (%)	 16 (76.2) 5 (23.8-0) 0 (0) 0 (0) 0 (0)	 16 (69.6) 5 (21.7) 2 (8.7) 0 (0) 0 (0)	N.S.
LDLR gene mutation Defective, n (%) Negative, n (%) Undefined, n (%)	 9 (45.0) 7 (35.0) 4 (20.0)	 8 (33.7) 9 (39.1) 6 (26.1)	N.S.
Chit gene mutation Mut/Mut, n (%) Wild/Mut, n (%) Wild/Wild,n (%)	 0 (0) 13 (61.9) 8 (38.1)	 0 (0) 14 (60.8) 9 (39.2)	N.S.

Table 2. Lipid parameters before and after plasmapheresis in 44 heterozygous FH patients (21 men and 23 women).

		Before Plasmapheresis	After Plasmapheresis	Student T test
Total cholesterol, mmol/L	(men)	8.99±4.03	3.23±0.85	0.0001
" " " "	(women)	8.29±1.53	2.98±0.70	0.0001
Triglycerides, mmol/L	(men)	1.46±1.24	0.62±0.26	0.0001
" " " "	(women)	1.04±0.53	0.44±0.21	0.0001
LDL cholesterol, mmol/L	(men)	7.25±1.20	2.10±0.74	0.0001
" " " "	(women)	6.82±1.33	1.97±0.63	0.0001
HDL cholesterol, mmol/L	(men)	1.33±0.28	0.93±0.21	0.0001
" " " "	(women)	1.48±0.35	1.03±0.22	0.0001
Chitotriosidase (nmol/ml/hr)	(men)	11.94±8.68	9.04±4.36	N.S.
" " " "	(women)	12.26±7.70	9.16±4.92	N.S.
Anti ox-LDL (UI/ml)	(men)	30.20±17.64	26.45±15.48	N.S.
" " " "	(women)	55.80±38.97	48.00±30.01	N.S.

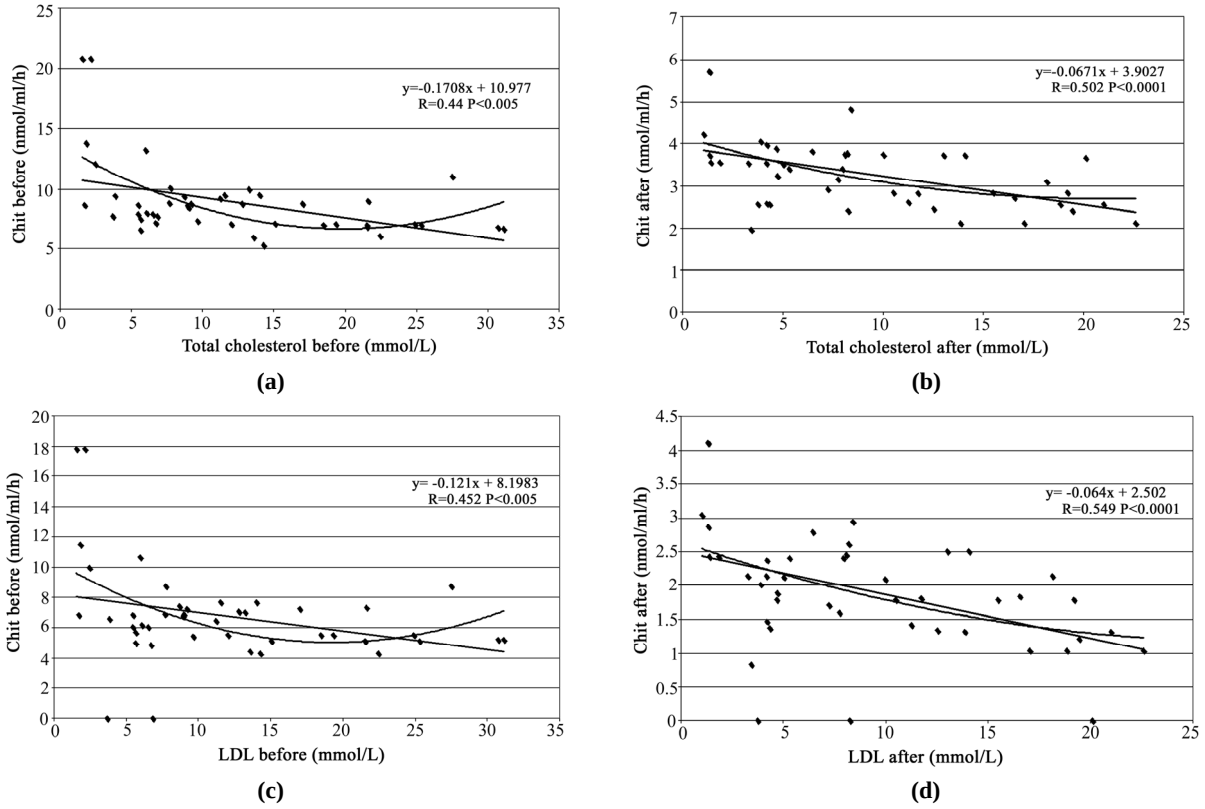


Figure 1.

reduction of anti-oxLDL concentration. Before and after LDL-apheresis the correlation between Chit and total cholesterol was negative (r=–0.44 and –0.50 respectively) and the same between Chit and LDL-cholesterol (r = –0.45 and –0.55 respectively) **Figures 1a, b, c, d**). The correlation between HDL-cholesterol and Chit before and after LDL-apheresis was not significant. The anti-oxLDL concentration positively correlated with Chit activity (r = 0.52) and negatively with LDL-cholesterol (r = –0.32) and total cholesterol (r = –0.33) before LDL-apheresis (**Figures 2a, c, e**). After LDL-apheresis the correlation with Chit was maintained high (r = 0.63) and the correlation with LDL-cholesterol and total cholesterol remained significant (r = –0.21 and r = –0.35 respectively) (see **Figures 2b, d, f**).

However negative correlations were not maintained for high levels of total cholesterol and LDL-cholesterol, before LDL-apheresis. This was supported by the order-2 polynomial trend behavior showed in **Figures 1a** and **c**, **Figures 2c** and **e**. On the contrary, these trends were maintained, when the level of total cholesterol and LDL-cholesterol decreased after LDL-apheresis.

5. DISCUSSION

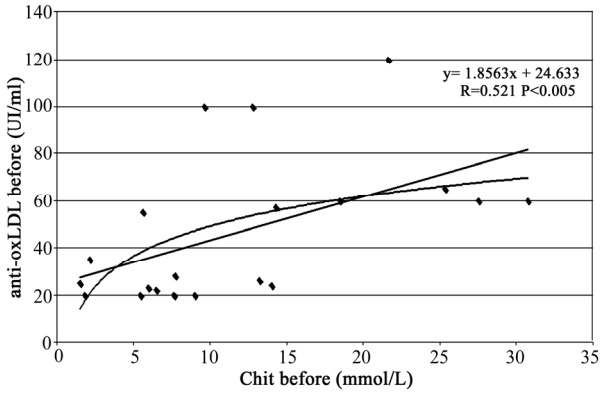
Plasma filtration represent an effective therapy in heFH

patients, who had no response to highest doses of statin and could restore through a macrophage activation the physiological mechanism of remotion of anti-oxLDL/ oxLDL immunocomplexes altered by LDLR genetic defect [15].

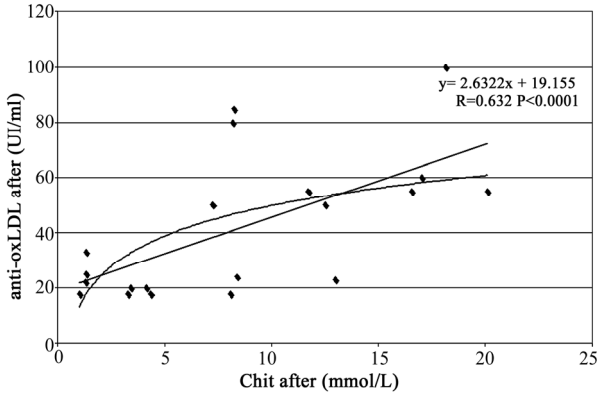
In fact in this study we demonstrated a negative correlation among Chit activity, total and LDL-cholesterol before and after LDL-apheresis, which seems to be dependent by active macrophage removal of anti-oxLDL/oxLDL immunocomplexes.

By removing LDL with LDL-apheresis, the anti-oxLDL/oxLDL ratio could increase and the excess of anti-oxLDL could induce macrophage activation through the surface Fc receptors. Alternatively with high levels of LDL-cholesterol, the deposition of foam cells represent the natural evolution of atherosclerosis process (as highlighted by the trend inversion above mentioned).

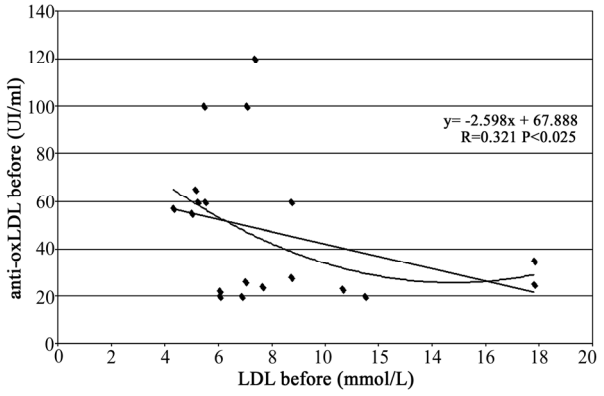
Hulthe *et al* 1998 [18] found, in a post hoc analysis, lower antibody titers in patients with a history of myocardial infarction, suggesting that antibodies against



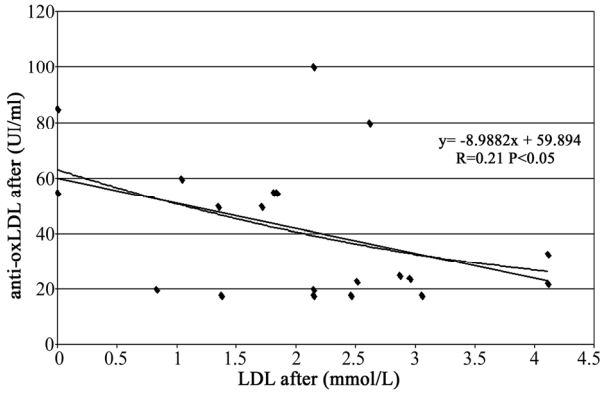
(a)



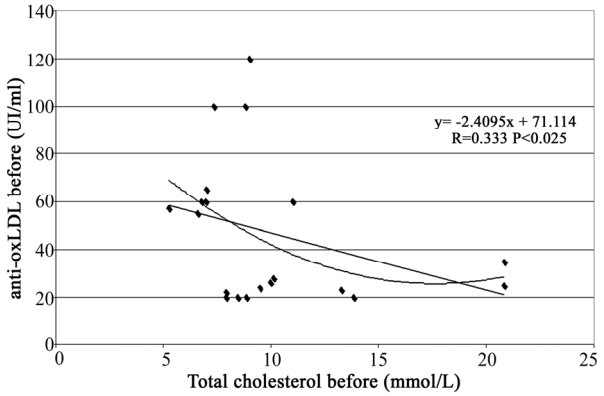
(b)



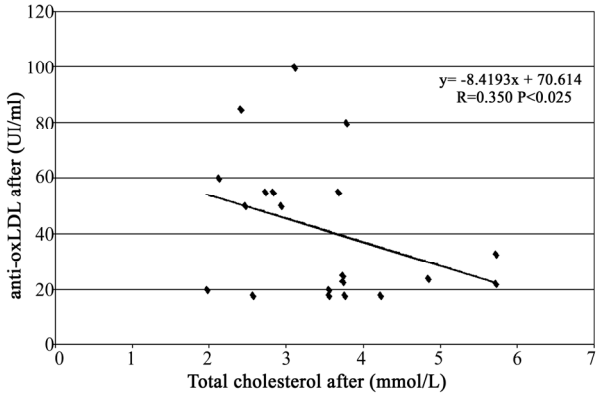
(c)



(d)



(e)



(f)

Figure 2.

oxLDL could have a protective effect. This observation is in agreement with Tinahones *et al* 2002 [19], who reported in literature a negative correlation among anti-oxLDL and total cholesterol in a population of 400 subjects from Malaga (Spain), suggesting again the active role of macrophage cells in the remotion of oxLDL. The same authors [20] in a larger population (1354 subjects) found a very significant increase ($P < 0.0001$) of anti-oxLDL: younger persons (16-35 years) had higher levels of anti-oxLDL (MDA-LDL) antibodies than persons older than 35 years ($P = 0.05$) and immune complexes were significantly higher ($P = 0.05$) in persons aged 5–15 years than in persons older than 40 years. On the contrary Artieda *et al* 2003 [10] reported an increased plasma Chit activity, independent from the distinct allelic or genotype distribution, in Spanish patients with atherothrombotic stroke and correlated plasma Chit activity to the severity of the atherosclerotic lesion [11]. Moreover, in Spanish hypercholesterolemic patients Canudas *et al* [21] found that no correlation exist between plasma Chit activity and the variation of plasma lipid levels before and after treatment with atorvastatin or bezafibrate. In our study plasma Chit activity, total and LDL-cholesterol levels correlated negatively before and after LDL-apheresis. These contrasting results could be due to different mechanism depending from the high level of oxLDL in plasma, which could inhibit the re-modeling of vessel wall in defect of anti-oxLDL.

In fact the activation of macrophages within the atherosclerotic lesion could represent an attempt for re-modeling the vessel wall, which may control the growth of lipid plaques. In condition where the total cholesterol and LDL-cholesterol levels are maintained elevated, the negative trend which seems to demonstrate an active mechanism of immunocomplexes removal disappears, supporting the foam cells formation and subsequent deposition. Some years ago we demonstrated "in vitro" that the activation of macrophages was induced by formation of immunocomplexes anti-oxLDL/ oxLDL through the Fc fragment receptor on the macrophage cell surface [22]. Recently this hypothesis was resumed by Fostergard *et al* 2007 [23] who demonstrated the importance of antibodies response, comparing two populations with different cardiovascular disease prevalence from New Guinea and from Sweeden. The discriminant factor among this two groups was a significant level of anti-phosphorylcholine antibodies of IgM subclass, which are active in the remotion of oxLDL. Also in atherosclerosis mice model, the induced increase of anti-oxLDL appears to be protective, in fact several investigators have demonstrated with oxLDL immunization [24,25] or with S. Pneumoniae immunization [26], that in such model the atherosclerosis development attenuates. This was also confirmed by an "in vivo" study in Watanabe heritable hyperlipidemic rabbits, lacking the LDL receptor and mimicking the human familial hypercho-lesterolemia. In this mice model the continuous auto-immunization with malondialdehyde-modified LDL, miming oxLDL, results in a very high concentration of antibodies against oxLDL, leading to significantly reduced progression of atherosclerosis [27]. This experiment opens the way to a model for the oxLDL vaccination, with the aim to activate the macrophage system to clean the atheromatosis lesions. Moreover it is evident that the mechanism of active cleaning of atheromatosis lesions is working when the level of total cholesterol and LDL-cholesterol is kept sufficiently low [5].

This evidence was also supported by our computational model that was able to reproduce experimental data and is actually used for making predictions and evaluate biological hypothesis [28].

In fact, LDL-apheresis is, up to date, the more active system for lowering LDL level in FH, and we found a negative correlation between anti-oxLDL, total cholesterol and LDL-cholesterol before and after LDL-apheresis, even plasma Chit activity and total and LDL-cholesterol correlated negatively before and after treatment. This stable response is explained by active remotion of immunocomplexes in these patients in the interval between LDL-apheresis. The increase of Chit activity was not found after statin treatment, probably because the modification of total cholesterol and LDL was not rapid enough to induce an activation of macrophage cells [21]. On the contrary Orem *et al* 2002 [29] reported after lipid lowering therapy with atorvastatin (10 mg/day) a significant decrement of anti-oxLDL and an increase of anti oxidant capacity of plasma LDL, which suggested again an active removal of immune complexes. They suggest also that the measurement of antibodies against oxLDL during lipid-lowering therapy may be used as an important marker for monitoring in-vivo LDL oxidation and atherosclerosis processes.

In individuals where the LDL were reduced by LDL-apheresis or by statine treatment, an active mechanism may be operative clearing the surface of vessel as was reported in our beta-thalassemia patients [30] and in diabetes patients in statine treatment [22] protecting from oxLDL vascular alteration.

6. ACKNOWLEDGEMENTS

Authors are grateful to Mister Rapicavoli Giuseppe, laboratory technician, for his precious assistance in the preparation of samples and in the determination of chitotriosidase activity.

7. AUTHORS CONTRIBUTION

Authors have equally contributed to the manuscript.

REFERENCES

[1] Goldstein, J. L., Hobbs, H. H., and Brown, M. S., (2001) Familial hypercholesterolemia, The Metabolic and Molecular Basis of Inherited Disease, Scriver CR, Beaudet, A. L., Sly, W. S., Valle D, eds., McGraw-Hill, New York, 2863−2913.

[2] Ross, R., (1999) Atherosclerosis: An inflammatory

disease, N. Engl. J. Med., **340**, 115−126.
[3] Civeira, F., (2004) Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia, Atherosclerosis, **173**, 55−68.
[4] Kern, F., (1990) Cholesterol metabolism, LDL, and the LDL receptor, Edited by Myant NB, Academic press, NewYork, 465.
[5] Shoji, T., Nishizawa, Y., Fukumoto, M., Shimamura, K., Kimura J., Kanda, H., *et al*., (2000) Inverse relationship between circulating oxidized low density lipoprotein (oxLDL) and anti-oxLDL antibody levels in healthy subjects. Atherosclerosis; **148**, 171−7.
[6] Bertolini, S., Cantafora, A., Averna, M., Cortese, C., Motti, C., Martini, S., *et al*., (2000) Clinical expression of familial hypercholesterolemia in clusters of mutations of the LDL receptor gene that cause a receptor-defective or receptor-negative phenotype, Arterioscler Thromb Vasc Biol, **20**, 41−52.
[7] Hollak, C. E., van Weely, S., van Oers, M. H., and Aerts, J. M., (1994) Marked elevation of plasma chitotriosidase activity, A novel hallmark of Gaucher disease, J. Clin. Invest., **93**, 1288−1292.
[8] Boot, R. G., van Achterberg, T. A., van Aken, B. E., Renkema, G. H., Jacobs, M. J., Aerts, J. M., *et al*., (1999) Strong induction of members of the chitinase family of proteins in atherosclerosis: Chitotriosidase and human cartilage gp-39 expressed in lesion macrophages, Arterioscler Thromb Vasc Biol, **19**, 687−694.
[9] Malaguarnera, L., Di Rosa, M., Zambito, A. M., dell'Ombra, N., Nicoletti, F., and Malaguarnera, M., (2006) Chitotriosidase gene expression in Kupffer cells from patients with non- alcoholic fatty liver disease, Gut, **55**, 1313−1320.
[10] Artieda, M., Cenarro, A., Gañán, A., Jericó, I., Gonzalvo, C., Casado, J. M., *et al*., (2003) Serum chitotriosidase activity is increased in subjects with atherosclerosis disease, Arterioscler Thromb Vasc Biol., **23**, 1645−1652.
[11] Artieda, M., Cenarro, A., Gañán, A., Lukic, A., Moreno, E., Puzo, J., *et al*., (2007) Serum chitotriosidase activity, a marker of activated macrophages, predicts new cardiovascular events independently of C-Reactive Protein, Cardiology, **108**, 297−306.
[12] Hansson, G. K., (2005) Inflammation, atherosclerosis, and coronary artery disease, N. Engl. J. Med., **352**, 1685−1695.
[13] van Eijk, M., van Roomen, C. P., Renkema, G. H., Bussink, A. P., Andrews, L., Blommaart, E. F., *et al*., (2005) Characterization of human phagocyte-derived chitotriosidase, a component of innate immunity, Int. Immunol., **17**, 1505−12.
[14] Shaw, P. X., Hörkkö, S., Tsimikas, S., Chang, M. K., Palinski, W., Silverman, G. J., *et al*., (2001) Human-derived anti-oxidized LDL autoantibody blocks uptake of oxidized LDL by macrophages and localizes to atherosclerotic lesions in vivo, Arterioscler Thromb Vasc Biol, **21**, 1333−9.
[15] Bláha, M., Cermanová, M., Bláha, V., Blazek, M., Malý, J., Siroký, O., *et al*., (2007) Safety and tolerability of long lasting LDL-apheresis in familial hyperlipoproteinemia. Ther Apher Dial, **11**, 9−15.
[16] Hixson, J. E. and Vernier, D. T., (1990) Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with HhaI, J. Lipid. Res., **31**, 545–8.
[17] Tsukamoto, K., Watanabe, T., Matsushima, T., Kinoshita, M., Kato, H., Hashimoto, Y., Kurokawa, K., and Teramoto, T., (1993) Determination by PCR-RFLP of apo E genotype in a Japanese population, J. Lab. Clin. Med., **121**, 598−602.
[18] Hulthe, J., Wikstrand, J., Lidell, A., Wendelhag, I., Hansson, G. K., and Wiklund, O., (1998) Antibody titers against oxidized LDL are not elevated in patients with familial hypercholesterolemia, Arterioscler Thromb Vasc Biol, **18**, 1203−1211.
[19] Tinahones, F. J., Gomez-Zumaquero, J. M., Rojo-Martinez, G., Cardona, F., Esteva de Antonio, I. E., Ruiz de Adana, M. S., *et al*., (2002) Increased levels of anti-oxidized low-density lipoprotein antibodies are associated with reduced levels of cholesterol in the general population, Metabolism, **51**, 429−31.
[20] Tinahones, F. J., Gomez-Zumaquero, J. M., Garrido-Sanchez, L., Garcia-Fuentes, E., Rojo-Martinez, G., Esteva, I., *et al*., (2005) Influence of age and sex on levels of anti-oxidized LDL antibodies and anti-LDL immune complexes in the general population, J. Lipid. Res., **46**, 452−7.
[21] Canudas, J., Cenarro, A., Civeira, F., García-Otín, A. L., Arístegui, R., Díaz, C., *et al*., (2001) Chitotriosidase genotype and serum activity in subjects with combined hyperlipidemia: Effect of the lipid-lowering agents, atorvastatin and bezafibrate, Metabolism, **50**, 447−450.
[22] Brizzi, P., Tonolo, G., Bertrand, G., Carusillo, F., Severino, C., Maioli, M., *et al*., (2004) Autoantibodies against oxidized low-density lipoprotein (oxLDL) and LDL oxidation status, Clin. Chem. Lab. Med., **42**, 164−70.
[23] Frostegård, J., Tao, W., Georgiades, A., Råstam, L., Lindblad, U., and Lindeberg, S., (2007) Atheroprotective natural anti-phosphorylcholine antibodies of IgM subclass are decreased in Swedish controls as compared to non-westernized individuals from New Guinea, Nutr Metab (Lond), **20**, 7.
[24] Binder, C. J., Chang, M. K., Shaw, P. X., Miller, Y. I., Hartvigsen, K., Dewan, A., *et al*., (2002) Innate and acquired immunity in atherogenesis, Nat. Med., **8**, 1218–26.
[25] Palinski, W., Miller, E., Witztum, J. L., (1995) Immunization of low density lipoprotein (LDL) receptor-deficient rabbits with homologous malondialdehyde-modified LDL reduces atherogenesis, Proc. Natl. Acad. Sci., USA, **92**, 821–5.
[26] Binder, C. J., Hörkkö, S., Dewan, A., Chang, M. K., Kieu, E. P., Goodyear, C. S., *et al*., (2003) Pneumococcal vaccination decreases atherosclerotic lesion formation: molecular mimicry between Streptococcus pneumoniae and oxidized LDL, Nat. Med., **9**, 736–43.
[27] Ameli, S., Hultgardh-Nilsson, A., Regnstrom, J., Calara, F., Yano, J., Cercek, B., *et al*., (1996) Effect of immunization with homologous LDL and oxidized LDL on early atherosclerosis in hypercholesterolemic rabbits, Arteri-oscler Thromb Vasc Biol, **16**, 1074–9.
[28] Pappalardo, F., Musumeci, S., and Motta, S., (2008) Modeling immune system control of atherogenesis, Bioinformatics.
[29] Orem, C., Orem, A., Uydu, H. A., Celik, S., Erdol, C., and Kural, B. V., (2002) The effects of lipid-lowering therapy on low-density lipoprotein auto-antibodies: Relationship with low-density lipoprotein oxidation and plasma total antioxidant status, Coron Artery Dis, **13**, 65–71.
[30] Brizzi, P., Isaja, T., D'Agata, A., Malaguarnera, L., Malaguarnera, M., and Musumeci, S., (2002) Oxidized LDL antibodies (OLAB) in patients with beta thalassemia major, J. Atheroscler Thromb, **9**, 139−44.

J. Biomedical Science and Engineering, 2009, 2, 506-515 **JBiSE**
doi: 10.4236/jbise.2009.27073 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Correlation of selected molecular markers in chemosensitivity prediction

David King, Thomas Keane, Wei Hu

Department of Computer Science, Houghton College, Houghton, NY 14744, USA.
Email: Wei.Hu@houghton.edu

Received 3 July 2009; revised 19 August 2009; accepted 20 August 2009.

ABSTRACT

Finding effective cancer treatment is a challenge, because the sensitivity of the cancer stems from both intrinsic cellular properties and acquired resistances from prior treatment. Previous research has revealed individual protein markers that are significant to chemosensitivity prediction. Our goal is to find correlated protein markers which are collectively significant to chemosensitivity prediction to complement the individual markers already reported. In order to do this, we used the D' correlation measurement to study the feature selection correlations for chemosensitivity prediction of 118 anti-cancer agents with putatively known mechanisms of action. Three datasets on the NCI-60 were utilized in this study: two protein datasets, one previously studied for chemosensitivity prediction and another novel to this topic, and one DNA copy number dataset. To validate our approach, we identified the protein markers that were strongly correlated by our analysis with the individual protein markers found in previous studies. Our feature analysis discovered highly correlated protein marker pairs, based on which we found individual protein markers with medical significance. While some of the markers uncovered were consistent with those previously reported, others were original to this work. Using these marker pairs we were able to further correlate the cellular functions associated with them. As an exploratory analysis, we discovered feature selection correlation patterns between and within different drug mechanisms of action for each of our datasets. In conclusion, the highly correlated protein marker pairs as well as their functions found by our feature analysis are validated by previous studies, and are shown to be medically significant, demonstrating D' as an effective measurement of correlation in the context of feature selection for the first time.

Keywords: Cancer; Chemosensitivity; Correlation; D'; Feature Selection; Genetic Algorithm; Markov Blanket; Memetic Algorithm; NCI-60

1. INTRODUCTION

The success of cancer treatment as well as the severity of the side effects of said treatment is heavily dependent on the sensitivity of the cancerous tissue to chemical treatment. Clinics face a great challenge in predicting treatment success, because chemosensitivity is determined by both intrinsic genomic and proteomic characteristics of the cancer as well as resistances induced through prior treatment. When trying to choose a therapy that will work best for a patient, it is important to evaluate their physical responses to different drugs. Because of this, many studies have been done to improve drug response prediction accuracy.

Data profiling of cancer cells at genomic, proteomic, chromosomal and functional levels has long been used in the analysis of pharmacological sensitivity of the cancer cells [1,2,3,4]. A primary source of cancer data in this field is a set of 60 human cancer cell lines provided by the National Cancer Institute (NCI-60) [5]. These cell lines have been in use since 1990 and over 100,000 chemical compounds have been tested on them [6]. The NCI-60 includes melanomas, leukemias and samples of ovarian, prostate, renal, breast, colon, lung and central nervous system cancers.

1.1. Related Works on the NCI-60

One study [7] used protein expression profiles to predict responses to a set of 118 anti-cancer agents with known or experimentally supported mechanisms of action [8]. Well known machine learning algorithms such as Random Forest, Nearest Neighbor and Relief were used to make chemosensitivity predictions. One Random Forest based classifier was built for each of the 118 drugs. To measure the significance of their predictions, this study compared the computed predictions against random pre-

dictions, which can be measured by a standard P-value. The P-value was the percentage of 1000 random predictions with higher accuracy than the calculated predictions. The study found chemosensitivity prediction accuracies ranging from 50 to 90%, with the vast majority being between 50 and 70%. Every prediction had *P*-values less than 0.019, and 97 of the predictions had *P*-values equal to 0.00.

A subsequent study by the same research group used a combination of the previously used proteomic data and new transcriptional data [9]. This integrative approach demonstrated its advantage, achieving higher accuracy and statistical significance, with *P*-values for all 118 drugs less than 0.001, calculated in the same manner as in [7].

A separate study [10] analyzed the correlation between DNA copy number variations, gene expression levels, and chemosensitivies to the same 118 drugs as in [7,9]. The analysis indicated that the correlations of gene expression and DNA copy number are particularly evident among leukemias and ovarian cancers.

An additional study [6] used four gene expression datasets, two of which were original to the paper, and one proteomic dataset. These data sets were used to observe the effectiveness of transcript profiling for the prediction of different protein expression levels. In addition, a consensus set selected from the four gene expression datasets was constructed. This consensus set was found to have a correlation to the protein dataset of 65%; a notable percentage that was higher than most reports done with mammalian cells. Further, this consensus dataset was used to predict tissue origin with a higher accuracy than any of its parent datasets.

1.2. Feature Selection and Motivation

New technologies in biomedical studies, such as microarrays, have made the analysis of large volumes of complex data a necessity [11]. Frequently, a majority of these data contain noise, i.e., features not relevant to a particular task at hand, such as classification of cancer types with gene expression data.

Both studies conducted by [7,9] used Random Forest as a feature selection technique to improve the accuracy of chemosensitivity predictions and to single out protein markers that were particularly important to this task.

In studying the effects of feature selection on chemosensitivity prediction, we observed disparity between expected and observed results. We ranked and ordered all features in the smaller protein dataset used in this study according to the Relief algorithm provided by Weka. We used Random Forest to make predictions based on incrementing feature subsets, using the top two ranked features, then three, four, etc. up to 40 features, as in **Figure 1**. We observed that contrary to our expectations, some higher ranked features decreased prediction accuracy, while some lower ranked features increased accuracy.

This led us to hypothesize that features contribute to the prediction accuracy collectively, rather than independently. To test this hypothesis, we developed a new technique using the D' measure [12] in order to study the correlations between feature pairs. As a demonstration of the utility of this technique, we apply it to those protein markers found to be significant in [9].

2. MATERIALS AND METHODS

2.1. Datasets

Three datasets derived from the NCI-60 were used in our study: two sets of protein data, and one of DNA data.

Protein expression data. The first protein expression dataset had 162 protein markers, hereafter referred to as Protein162, and was created by Shankavaram *et al* [6] and can be found at http://discover.nci.nih.Gov/datasets.jsp. The second dataset, which contains 52 protein markers (Protein52), available at http://discover.nci.nih.gov/host/2003_profilingtable7.xls., was generated by a study of the proteomic profiles of the NCI-60 [13], and was also used by two studies on chemosensitivity prediction [7,9].

DNA copy number. The DNA copy number variation dataset was presented in a study of the correlation between mRNA and DNA copy number [10]. It is available at http://discover.nci.nih.gov/datasets.jsp.

Drug activity data. Our drug resistance information contained activity data from 118 anti-cancer agent activity profiles. They were screened by Scherf *et al* [8] and recorded using the NCI-60 cancer cell lines. The file containing this data can be found at http://discover.nci.nih.gov/nature2000/data/selected_data/dataviewer.jsp?baseFileName=a_matrix118&&nsc=2&dataStart=3.

Defining drug sensitivity and resistance. As in [7,9] we used a threshold to define sensitivity to a drug into

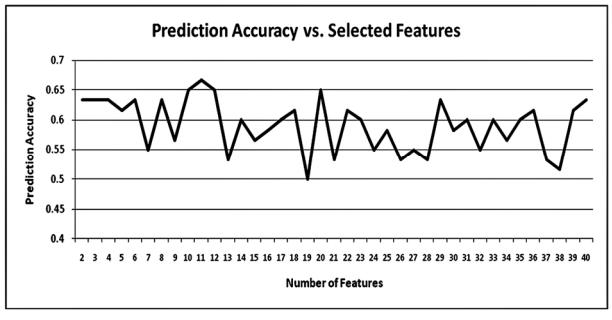


Figure 1. Random forest prediction accuracy. This plot shows the prediction accuracies of Random Forest using the same protein dataset used in [7,9]. The drug on which the prediction was performed was Bisentrene (NSC # 337766).

three categories. A $\log_{10}$ (GI_{50}) was taken for each cell line to determine sensitivity. Cell lines with sensitivities at least 0.5 standard deviations above the average were given the label 'resistant.' Those with sensitivities at least 0.5 standard deviations below the average were 'sensitive.' The remaining cell lines were defined as 'intermediate' [7,9].

2.2. D' Formula

A standard measurement for the correlation between pairs of events *i* and *j* in a set of sequences is *D'*, which can be defined by the following formulae:

$$D_{ij} = x_{ij} - p_i q_j \qquad D'_{ij} = \frac{D_{ij}}{D_{\max}}$$

where x_{ij} is the frequency at which both event *i* and event *j* occur in a single sequence, p_i is the frequency of event *i* and q_j is the frequency of event *j*. If $D_{ij} < 0$,

$$D_{max} = min\left[p_i q_j, (1-p_i)\left(1-q_j\right)\right], \text{ and if } D_{ij} > 0,$$

$$D_{max} = min\left[p_i \left(1-q_j\right), (1-p_i) q_j \right].$$

The *D'* formula was introduced by Richard Lewontin as a measurement of linkage disequilibrium of alleles at two or more loci on the same chromosome [12]. The *D'* formula has been shown to be a more reliable measurement than other measurements of correlation between pairs of events [14], but this study is the first to use it to correlate pairs of selected features.

2.3. Markov Blanked-Embedded Genetic Algorithm (MBEGA)

Genetic Algorithms have been used as a strategy for feature selection [15] due to their ability to generate better feature subsets than other feature selection algorithms. In some cases, these genetic algorithms are combined with memetic operations in order to fine tune results beyond what would be produced by classical genetic algorithms alone.

One particular implementation of these memetic algorithms is the Markov blanket-embedded genetic algorithm (MBEGA), which uses an approximation of a Markov blanket to reduce redundancy in selected features. Pseudocode for the MBEGA can be found in **Figure 2**.

In each generation of the algorithm, the MBEGA uses add and delete operations to add and delete features from some of the elite feature subsets in the population; the elite feature subsets are improved by adding important features and removing those that are less important. After the memetic operations, standard genetic algorithm techniques such as linear ranking, crossover and mutation methods occur to generate the next population [16].

The MBEGA was selected in our study for two reasons: First, the MBEGA generates a population of feature subsets in each generation, rather than generating a single final subset as in classical feature selection algothms. The feature subsets from each generation are represented as binary strings, with a 1 representing a present feature and 0 representing an absent feature, to calculate the D' values of our correlation analysis. Second, the MBEGA does not require a predefined number of features to be selected. Rather, the MBEGA gradually optimizes both the size of the feature subset as well as the accuracy of the classifier.

Markov Blanket Embedded Genetic Algorithm (MBEGA)

BEGIN
(1) **Initialize:** Randomly generate an initial population of feature subsets encoded as binary strings
(2) **For** the number of iterations to run
(3) Evaluate all feature subsets in the population based on prediction accuracy
(4) Select a number of elite feature subsets from the population to undergo the Markov blanket memetic operations
(5) **For** each feature subset create a set of all present features X and all absent features Y
Add operation **BEGIN**
1) Rank the features in Y according to their correlation to the class label.
2) Select a feature *Yi* in Y so that the larger the correlation of a feature in Y the more likely it will be picked.
3) Add *Yi* to X.
END
Delete operation **BEGIN**:
1) Order the features in X according to their correlation to the class label.
2) Select a feature *Xi* in X so that the larger the correlation of a feature in X the more likely it will be picked.
3) Eliminate all features in X which are less correlated than *Xi*. If no feature is eliminated, remove *Xi*.
END
(6) Replace the original elite feature subset with the improved feature subset.
(7) **End For**
(8) Perform crossover and mutation to create the next generation of feature subsets.
(9) **End For**
END

Figure 2. Pseudocode for MBEGA.

2.4. Correlation Analysis Using D' Formula

We used the D' formula to calculate correlation between pairs of features selected in each generation of the MBEGA. Because the MBEGA begins with a randomized feature subset and becomes more selective as the algorithm progresses, we decided to use only the last 20% of the feature subsets generated. We calculated the D' values for every pair of selected features, using the presence of one feature within the encoded binary sequence as event *i* and the presence of the other as event *j*.

2.5. Feature Selection Using Weka's Relief Algorithm

There are three primary types of feature selection algothms: filter, wrapper and embedded algorithms. Filter algorithms have advantages in their speed and scalability,

however they ignore feature dependencies. They also do not interact with classifiers, which is both an advantage, because they can select features independently, and a disadvantage, because they are unable to take the classifier into account when determining the feature subset. Wrapper algorithms, on the other hand, do interact with the classifier, and are therefore able to produce more informative feature subsets. They are also less prone to local optima. They are, however, computationally intense, and have a higher risk of over fitting. Embedded algorithms are built directly into a classifier. As such, they are able to interact with the classifier in the same manner as wrapper algorithms, but are far less computationally intense.

Relief, a filter feature selection algorithm implemented in WEKA, was used to assess the features pairs found by our correlation analysis. Relief ranks features by assigning them weights according to their ability to

Table 1. Highly correlated protein marker pairs in protein52 based on significant chemosensitivity protein markers. The protein markers in column one and associated drugs, expressed as their NSC drug numbers, in column two were found in [9]. The remaining columns are the ten protein markers with the highest correlation to the protein marker in the first column. The markers notated with a * are those also selected by Weka's Relief algorithm.

Protein Marker	NSC Drug #	1	2	3	4	5	6	7	8	9	10
ISGF3g	56410	MAPK1	EP300*	MSN	NME1*	GSK3B	FADD	STAT1*	STAT3	STAT5A	MSH6
ISGF3g	354646	MGMT	VIL1	RIPK1*	EP300	EP300	MSN	GSK3B	FADD	STAT5A	MSH2
STAT3	56410	EP300*	EP300	MSN	GSK3B	FADD	ISGF3G*	STAT1*	STAT5A	STAT6*	MSH2
NME1	353451	FN1	MVP	RELA	MSN	CDH1*	MGMT	GSK3B	FADD	ISGF3G*	STAT5A
NME1	344007	KRT18	EP300	MAPK1	CDH1*	MGMT	GSK3B	FADD	ISGF3G	STAT3*	STAT5A
NME1	102816	TP53	EP300	MGMT*	CCNE	MAP2K1	CDH1*	MGMT*	GSK3B	FADD	STAT3*
NME1	107392	KRT8	MSN*	CDH1	MGMT*	GSK3B	FADD	ISGF3G	STAT3*	STAT5A	MSH2
MGMT	95466	KRT18*	CDH2	EP300	MSN	EP300	MSN*	CDH1	NME1*	GSK3B	ISGF3G*
CCNE	95441	KRT8*	CCNA2	CCNB1	VIL1*	CDH1	RELA	RIPK1	JAK1	MAP2K2	STAT5A
EP300	119875	KRT18	EP300	CDH2	KRT20	FN1	MSN	CCNB1*	JAK1	MAP2K1	MAP2K2
EP300	606497	EP300	CDH2	KRT20	FN1	KRT8	CCNB1	CCNE	RIPK1	STAT3*	MAP2K1
FN1	135758	KRT18	CDH2	KRT20	KRT8	CCNA2	CCNB1	CCNE	VIL1	MAP2K2	ISGF3G
MSN	301739	KRT20	MAPK1	MCP	MCM7	CDK6	G22P1	MVP	PGR	MAP2K2	FADD
MSN	755	KRT18	CDH2	MCM7	CDK6*	G22P1	MVP	PGR	MAP2K2	FADD	STAT1
MSN	376128	PCNA	MCP	CDK6	G22P1	MVP*	PGR	CCNE	VIL1	CDH1*	EP300
PGR	354646	MSN	MVP	CCNA2*	CCNB1*	CCNE	CDH1	CASP2	RIPK1*	EP300	FADD
STAT1	354646	KRT20	MAPK1	G22P1	MVP	MAP2K2*	MSN	NME1*	FADD	STAT3	STAT5A
STAT6	354646	EP300	PCNA	MAPK1	FADD	ISGF3G	STAT3	STAT5A	MSH2	MSH6	EP300
CASP2	264880	CCNA2*	CCNE	VIL1	CDH1	RELA*	RIPK1	JAK1*	STAT3	EP300	STAT1
CDH1	71261	MAPK1	CCNE	VIL1*	RELA*	CASP2	EP300	EP300	CDH1	NME1	MSH2*
MCP	740	TP53	EP300	EP300	KRT20	ACVR2*	MCM7*	CDK6	CCNB1	VIL1	EP300
KRT18	19893	TP53	EP300	CDH2	EP300	FN1*	KRT8	PGR	JAK1	MAP2K2	EP300
KRT18	757	TP53	EP300	RELA	STAT3	EP300	CDH1	GSK3B	STAT5A	MSH6	EP300
KRT18	33410	TP53	EP300	CDH2	EP300	KRT20*	RIPK1	MAP2K2	MSN	MSH6	EP300
KRT18	125973	TP53	EP300	CDH2	EP300*	KRT20	CDK6	CCNA2	CCNE	RELA	EP300
KRT18	658831	TP53	EP300	KRT20	FN1*	MAPK1	MSN	G22P1	JAK1	CDH1	EP300
KRT18	673188	TP53	EP300	CDH2	FN1	GSK3B*	FADD*	STAT5A	STAT6	MSH6	EP300
KRT18	671867	TP53	EP300	CDH2	EP300	CCNB1	CASP2	EP300	MAP2K1	MSH6	EP300
KRT18	664402	TP53	EP300	CDH2	EP300	MVP	PGR	JAK1	EP300	STAT6	EP300
KRT18	661746	TP53	EP300	CDH2	EP300	ACVR2	MCP	EP300	MSN	CDH1*	EP300
KRT18	673187	TP53	EP300	CDH2*	ACVR2	CDK6	VIL1	STAT6	MSH2*	MSH6*	EP300
KRT18	664404	TP53	EP300	CDH2	KRT20	RB1	MAPK1	EP300	EP300	STAT5A*	EP300
KRT18	671870	TP53	EP300	CDH2	KRT20*	FN1	PCNA	CDH1	CASP2	STAT6	EP300*
KRT18	666608	TP53	EP300	CDH2	EP300	KRT20	MAPK1	CCNB1	RIPK1	STAT3	MGMT
KRT18	600222	TP53	EP300*	CDH2*	RB1	G22P1	PGR	CCNA2	MSN	STAT3	STAT5A
KRT18	656178	EP300	EP300	EP300	MGMT	MAPK1	MSN	G22P1	MAP2K1	STAT5A*	STAT6
TP53	19893	KRT18	CDH2	RB1	MAPK1	EP300	CDK6	MSN	STAT6	MSH6	EP300
TP53	125973	KRT18*	KRT20	FN1*	MGMT	MAPK1	ERBB2	MCM7*	STAT6	MSH6	EP300
RELA	153353	CDH2*	MSN	G22P1	VIL1	CDH1	CASP2	RIPK1*	JAK1	MAP2K1	MAP2K2
G22P1	224131	EP300	MAPK1	MSN	MVP*	CCNA2	CCNB1	RIPK1	EP300*	MAP2K2	EP300

discriminate between neighboring patterns. In each iteration of the algorithm, an instance x containing features $(x_1,x_2,\ldots,x_n)$ is selected randomly, and one nearest neighbor from the same class (called NH) along with one nearest neighbor from a different class (called NM) are found. The weights of the features in x are updated such that they will be greater if x is similar to the NH and dissimilar to the NM, and less if the opposite is true.

3. RESULTS AND DISCUSSION

To generate sequences for D' analysis, we ran a ten-fold cross validation of the MBEGA on all three datasets. Each fold of the MBEGA ran for 100 generations, with a population of 51. Each fold generated 5100 sequences, of which we used the last 20% generated, or 1020 from each fold. The final number of sequences used for the D' analysis was 10200 for each dataset.

3.1. Correlation Analysis of Individual Protein Markers from Previous Study

A previous study [9] discovered 18 individual protein markers from Protein52 along with their functions, including transcriptional factoring, tumor suppressing, DNA repair, cell adhesion, and apoptosis, among others, that are significant to the prediction of chemosensitivity to 33 of the 118 anti-cancer agents. These drugs represent 12 out of the 15 total mechanisms of action present in the 118 anti-cancer agents, with a large number of them being tubulin active antimitotic agents. In order to investigate the protein markers highly correlated with those found in [9], for each protein marker/drug combination identified there, we found the ten protein markers with the highest D' value. We also sought to validate these pairs using a ten-fold cross validation of the Relief feature selection algorithm provided by Weka, which measures feature significance individually. As seen in **Table 1**, the highly correlated protein marker pairs our analysis discovered are validated not only by the protein markers reported in [9], but also by Relief.

In order to discover any patterns in the correlation of the protein markers selected in **Table 1**, we took a frequency count of them, as illustrated in **Figure 3**. While most protein markers had frequencies within the same range, mostly between 4 and 10, there were some which clearly stood out. In particular, the protein marker CDH2, a cell adhesion protein, is highly correlated with 19 out of the 40 protein markers in **Table 1**. CDH2 was not selected in the previous study, but is very similar in both function and family to CDH1, which was selected. Another protein with a high frequency, 16 out of 40, was TP53 whose function is tumor suppression and apoptosis. We found that in most occurrences TP53 was paired with protein marker KRT18. Both of these protein markers are involved in protein death, and both were found to be strong chemosensitivity predictors for the drug Taxol in the previous study [9]. Lastly, we noticed STAT5A is both from the same family as and is highly correlated with the protein markers STAT1 and STAT6, both of which were highlighted in previous study [9].

We were also interested in observing how the functions of the individual protein markers from [9] correlated with the functions of the highly correlated protein markers found in **Table 1**. We grouped the previously reported protein markers according to their function, and then selected the protein markers that were most frequently correlated with them. We only included those protein functions which had three or more protein markers associated with them, as in **Table 2**.

Because we used two protein datasets in this study, we wanted to conduct the same analysis on the Protein162 dataset in order to explore the possibility of discovering new protein markers highly correlated with those found in [9]. All but 2 of the previously reported protein markers from Protein52, G22P1 and CCNE, were also present in Protein162, so we used the same protein marker/drug combinations as in **Table 1** when generating **Table 3** for Protein162.

We also created a selection frequency histogram for

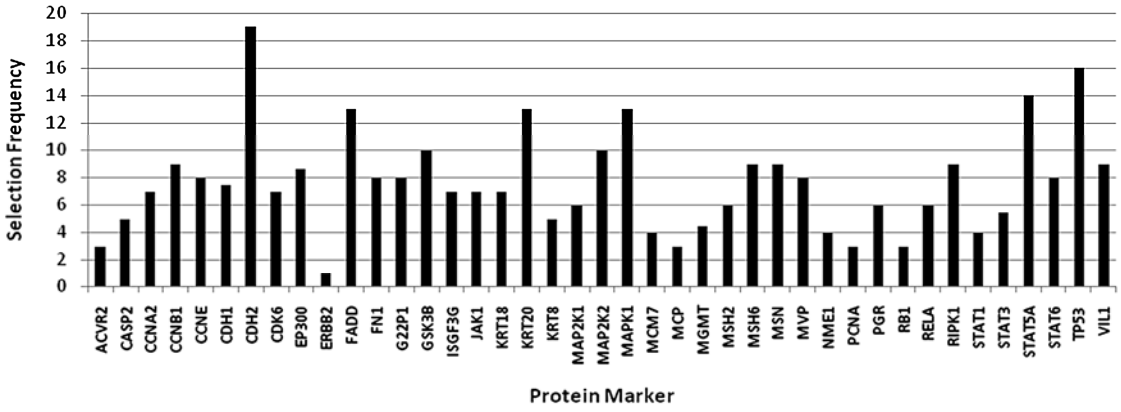


Figure 3. Frequency of protein52 protein markers present in **Table 1**.

Table 2. Correlation of the functions of protein markers in **Table 1**.

Protein Function	Reported Protein Markers	Correlated Protein Markers	Functions of Correlated Protein Markers
Transcriptional Factor	ISGF3G	STAT5A	Transcriptional Factor
	STAT3	FADD	Apoptosis
	EP300		
	STAT1		
	STAT6		
	RELA		
Integrin Signaling	NME1	STAT6	Transcriptional Factor, Inteferon Signaling
	EP300	MSH6	DNA Repair
	TP53	KRT18	Structural Protein; Biomarker of cell death
		FADD	Apoptosis
Tumor Suppressors	CCNE	FADD	Apoptosis
	TP53	FN1	Cell Adhesion; Integrin Signaling
	RELA	GSK3B	Hormonal Control
		KRT18	Structural Protein; Biomarker of cell death
Cell Apoptosis	CASP2	CDH2	Cell Adhesion
	KRT18	KRT20	Structural Protein
	TP53	MSH6	DNA Repair
	RELA	TP53	Tumor Suppressor; Cell Cycle and Apoptosis

Table 3. Highly correlated protein marker pairs in protein162 based on significant chemosensitivity protein markers. The protein markers in column one and associated drugs, expressed as their NSC drug numbers, in column two were found in [9]. The remaining columns are the ten protein markers with the highest correlation to the protein marker in the first column. The protein markers G22P1 and CCNE present in **Table 1** are excluded here because neither is present in Protein162. The markers notated with a * are those also selected by Weka's Relief algorithm.

Protein Marker	NSC Drug #	1	2	3	4	5	6	7	8	9	10
ISGF3g	56410	ANXA1	CASP7	CRK	EP300	EP300	EP300	JAK1	RELA	RIPK1	TP53
ISGF3g	354646	AKAP5	AP2M1	CDC2*	MSN	PRKCI	RIPK1*	SMARCB1	FASLG	TP53	VIL2
STAT3	56410	CCNA2	CDH1	CDKN2A	HRAS	KRT18	MAPK1*	MVP	STAT1*	STAT3	VASP
NME1	353451	ANXA4	EP300	EP300	FN1	GTF2B*	IRS1	MGMT	NCAM1	PCNA	TP53
NME1	344007	CASP7	CASP7	CDH2	ENAH	EP300	HSPA4	JAK1	MCC	PRKCI	TP53
NME1	102816	CASP2	CASP7	EP300	FADD	ISGF3G	MAP2K1	MGMT	MSN	NCAM1*	TUBB2A
NME1	107392	ANXA1	CASP7	EP300	EP300	MAP2K2	NCAM1	PCNA	PRKCA	RB1	RELA
MGMT	95466	ACVR2A	BCAR1	EP300	FADD	MGMT*	RB1	STAT1	TP53	TP53	YWHAG
EP300	119875	PARP1	CTNNB1*	EP300	GRB2*	JAK1	MCC	MSN	RIPK1	STAT6*	TP53
EP300	606497	ADNP	ATXN2*	EP300	EP300	GRB2	JAK1	MCP*	MSH6	STAT3*	TP53
FN1	135758	AKAP8	CDK4	CTTN	EP300	EP300	EP300	GSTP1	PCNA	STAT6	FASLG
MSN	301739	ADNP	CDH1	EP300	JAK1	MAP2K2	MSN*	MVP	RIPK1	SMARCB1	STAT3
MSN	755	CDK5	CTNNB1	EP300	ISGF3G	JAK1	KRT18	MAPK1	MCC	RB1	RIPK1
MSN	376128	PARP1	CDC2	CDH2	EP300	EP300	MGMT	MVP*	PTPN6	RB1*	TP53
PGR	354646	CDH2	ENAH	EP300	EP300	EP300	EP300	MSN	RELA	EXOC4	TP53
STAT1	354646	CASP2	EP300	EP300	EP300	FADD	MCM7	MSH6*	PRKCB1	RELA	TYR
STAT6	354646	PARP1	CDK4*	CDK7	CRK	ENAH	EP300	MSH6*	RB1	RELA	STAT3
CASP2	264880	CASP7	CTNNB1*	DSG1	EP300	EP300	EP300	FADD*	ISGF3G	KRT7*	PCNA
CDH1	71261	FN1	KRT19	MGMT	PTPN11	RELA*	RELA	RELA	STAT6	TP53	TRADD
MCP	740	CASP7	DSG1	EP300	ESR1	FADD	KRT19	MAP2K2	MCM7	RB1*	SMARCB1
KRT18	19893	PARP1	PARP1	CDKN2A	EP300	MCM7*	MSN	PRKCA	RB1	RELA	STAT5A
KRT18	757	CDK4	ENAH	EP300	EP300	EP300	KRT19*	MCP	SMARCB1	STAT5A	STAT6
KRT18	33410	AKAP8	EP300	EP300	EP300	EP300	JAK1	KLK3	KRT19	STAT1	VASP
KRT18	125973	CDC2	CDK5	GSTP1	IRS1	KRT19*	TP53	TP53	TP53	TP53	VIL2
KRT18	658831	ATXN2	CCNA2	IRS1	MCM7*	MSH2*	EXOC4	SMARCB1*	TP53	TP53	TRADD
KRT18	673188	AKAP5	CDK5	EP300	JAK1	KLK3	KRT19*	MAP2K2	RELA	TGFB1I1	VASP
KRT18	671867	ADNP	BCAR1	CCNA2	CDC2	EP300	KLK3	MAP2K1	MAPK1	MVP	STAT3
KRT18	664402	ADNP	CCNA2	EP300	EP300	KRT19*	KRT7	PTPN11	RELA	RELA	STAT6
KRT18	661746	AP2M1	CCNB1	CDH2	DSG1	EP300	KRT19	PRSS8	RELA	STAT5A	VASP
KRT18	673187	CCNA2	EP300	EP300	ERBB2	FADD	XRCC6	MGMT	MVP*	STAT3	TP53
KRT18	664404	CDK6	EP300	ISGF3G	KLK3	MCC	MSH6	PRKCB1	STAT1	STAT6	FASLG
KRT18	671870	CASP2	CCNB1	CDH2	EP300	MAP2K1	MCP	MLH1	PCNA	RELA	EXOC4
KRT18	666608	CDH1*	CDK7	ENAH	EP300	FADD	GSTP1	KLK3	KRT20	PRSS8*	VIL1
KRT18	600222	ADNP	AKAP5	AKAP8	CDH2	EP300*	GSTP1*	KLK3	MAP2K1	PTPN11	TYR
KRT18	656178	CDH2	EP300	GTF2B	KRT19	KRT8*	EXOC4	TP53	TP53	TP53	TRADD
TP53	19893	CASP7	EP300	EP300	ESR1	GTF2B	KRT8	MAP2K2	STAT1	STAT1	STAT6
TP53	125973	CASP7	CCNA2	CDH1	EP300	JAK1	KLK3	KRT19*	PRKCA	STAT1	TP53
RELA	153353	ADNP	CASP7	CDK7	EP300	EP300	EP300	PRKCI	PRSS8	PTPN11	STAT3

Table 3, illustrated in **Figure 4**. We observed that the frequencies were lower by roughly a factor of 2 for this dataset when compared to Protein52. We believe this is because the number of unique protein markers in Protein162 was roughly twice that of the unique protein markers in Protein52; however we chose the top ten most correlated pairs in both instances.

Many of the most selected protein markers from **Table 1**, including CDH2, TP53, and STAT5A, had only an average or even low frequency in **Table 3**. We selected 8 protein markers from **Table 3** whose average frequencies were above 4. These were KRT18, KLK3, CCNA2, ADNP, MVP, RIPK1, SMARCB1, and ENAH. Because the Protein162 dataset contains protein markers not present in Protein52, we found 5 protein markers which were not reported in the previous study [9]. These proteins, as well as their cellular functions and associated drugs can be seen in **Table 3**. The most frequently selected protein marker from **Table 3** was KRT19, a structural protein from the same family as KRT18, a protein marker found to be significant in [9], and KRT20, a protein marker frequently selected in **Table 1**. KLK3 had the second highest frequency of selection. Serum levels of the KLK3 protein, called PSA, are used to diagnose and monitor prostatic carcinoma. Members of the KLK family are also thought to be biomarkers for cancers and diseases. CCNA2 has a functional relationship with CDC2, another protein marker with an above-average selection frequency in **Table 3**. ADNP affects both normal cell growth and cancer proliferation. In addition, ADNP is a transcription factor, a trait held in common with six of the eighteen significant protein markers in [9]. MVP is a protein which is over-expressed in multi-drug resistant cancer cells, and is potentially useful as a signal for drug resistance. MVP also bears a functional relation with STAT1, one of the important protein markers reported in [9]. RIPK1 is an apoptosis protein related to cell death, much like the KRT18, TP53 and CASP2 found in both the previous study [9] and in **Table 1**. SMARCB1 functions as a tumor suppressor, but mutations within the protein are associated with rhabdoid tumors. ENAH is a cell adhesion protein which is present in some breast cancers, and may be used as a marker for such.

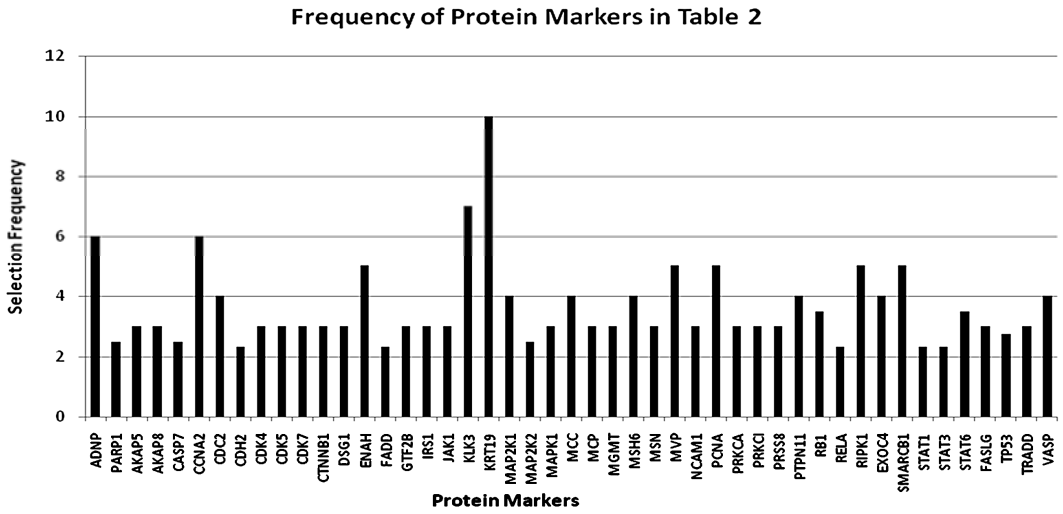


Figure 4. Frequency of protein162 protein markers present in **Table 3**. This plot shows the frequency with which protein markers from the Protein162 dataset were selected in **Table 3**. Only those protein markers with an average frequency above 2 are shown due to limited space.

Table 4. Highly correlated protein markers for the evaluated anticancer drugs**.** Protein markers denoted with a * were unique to the Protein162 dataset, and as such not reported in previous study [9].

Protein Marker	Function	NSC Drug Numbers
KRT19*	Structural Protein	71261, 740, 757, 33410, 125973, 673188, 664402, 661746, 656178
KLK3*	Biomarker of Prostatic Carcinoma	33410, 673188, 671867, 664404, 666608, 600222, 125973
ADNP*	Cell Growth, Cancer Proliferation, Transcription Factor	125973, 301739, 671867, 664402, 600222, 153353
CCNA2	Binding & Activating Agent	56410, 658831, 671867, 664402, 673187, 125973
MVP	Mediating Drug Resistance, Over-expressed in multi-drug resistance cancer cells	56410, 301739, 376128, 671867, 673187
RIPK1	Apoptosis Protein	56410, 354646, 119875, 301739, 755
SMARCB1*	Tumor Suppressor	354646, 301739, 740, 757, 658831
ENAH*	Cell Adhesion Protein	344007, 354646, 354646, 757, 666608

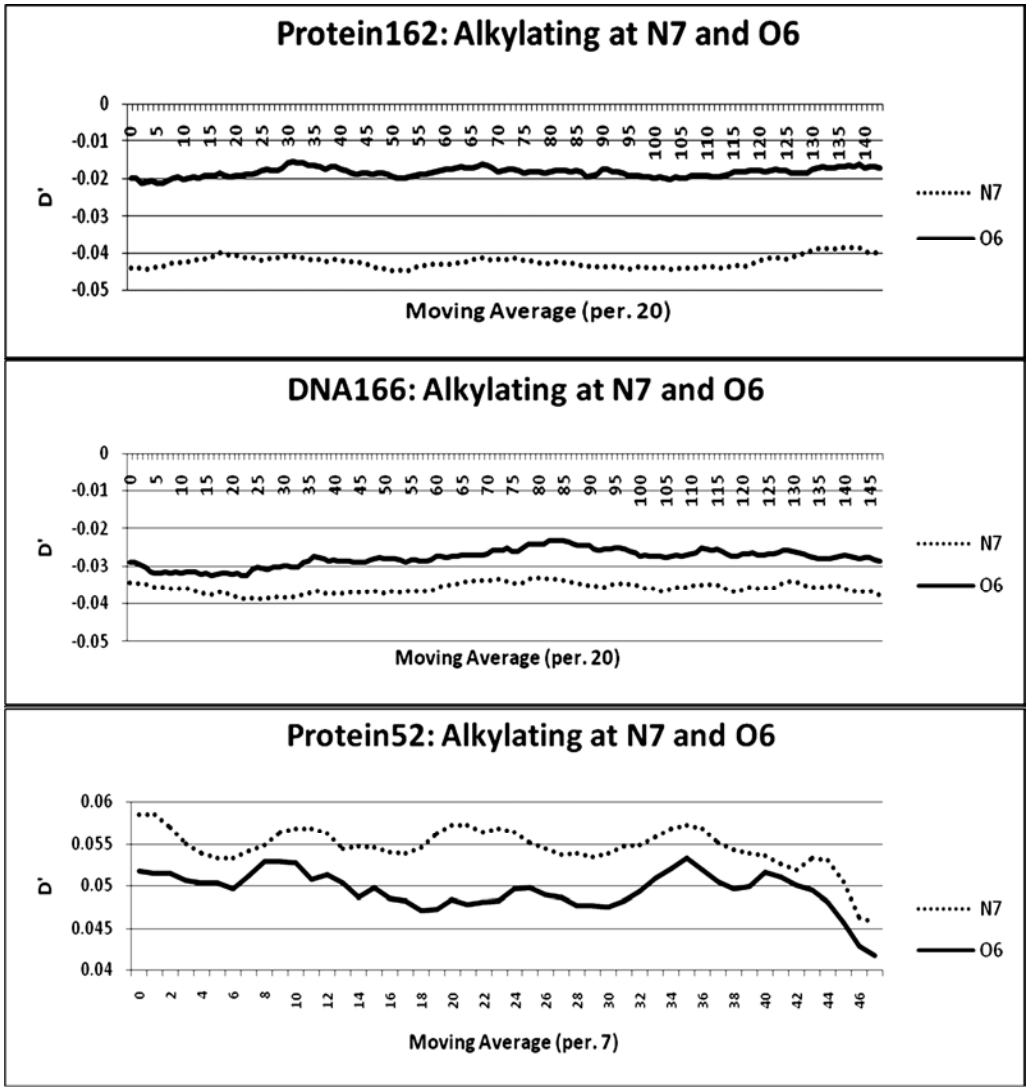


Figure 5. D' patterns categorized according to mechanism of action. The drugs which alkylate at N7 and O6 produce similar patterns of D' values in each of the three datasets. For readability, the curves displayed are moving average curves with a period of 20 for Protein162 and DNA166 and a period of 7 for Protein52.

3.2. Feature Correlation Analysis of All 118 Drugs

In addition to simply calculating the D' values of the feature pairs, we also calculated the average D' values for each feature based on the D' values of all pairs associated with that feature. We performed this analysis for both protein datasets as well as the DNA166 dataset, using the protein markers as features for the protein datasets, and the DNA copy number variations as features for the DNA166 dataset. As an exploratory analysis, we attempted to use the average D' values to find the trends of feature correlation within and between the mechanisms of action of the 118 anti-cancer agents.

Each dataset generated unique patterns of feature correlation in each of the 118 drugs. We did observe, however, similar patterns of feature correlation in drugs with related mechanisms. In particular, we noticed that topoisomerase I inhibitors and topoisomerase II inhibitors have very similar trends of feature correlation to one another in all three datasets. Drugs which alkylate at positions N7 (24 drugs) and O6 (7 drugs) of guanine also have very similar trends of feature correlation, as shown in **Figure 5**. This implies that related drug mechanisms tend to produce similar patterns of correlation between feature pairs. Our analysis indicated that this is not necessarily true of drugs with similar chemical structure.

We also grouped drugs with similar mechanisms into three larger categories: drugs which alkylate at specific positions of guanine (Alkylating), drugs which inhibit topoisomerase (TIM Inhibitors), and all other drugs (Other). The D' values of these larger categories were generated by averaging the D' values of the individual drugs within that larger category. We found the correlation

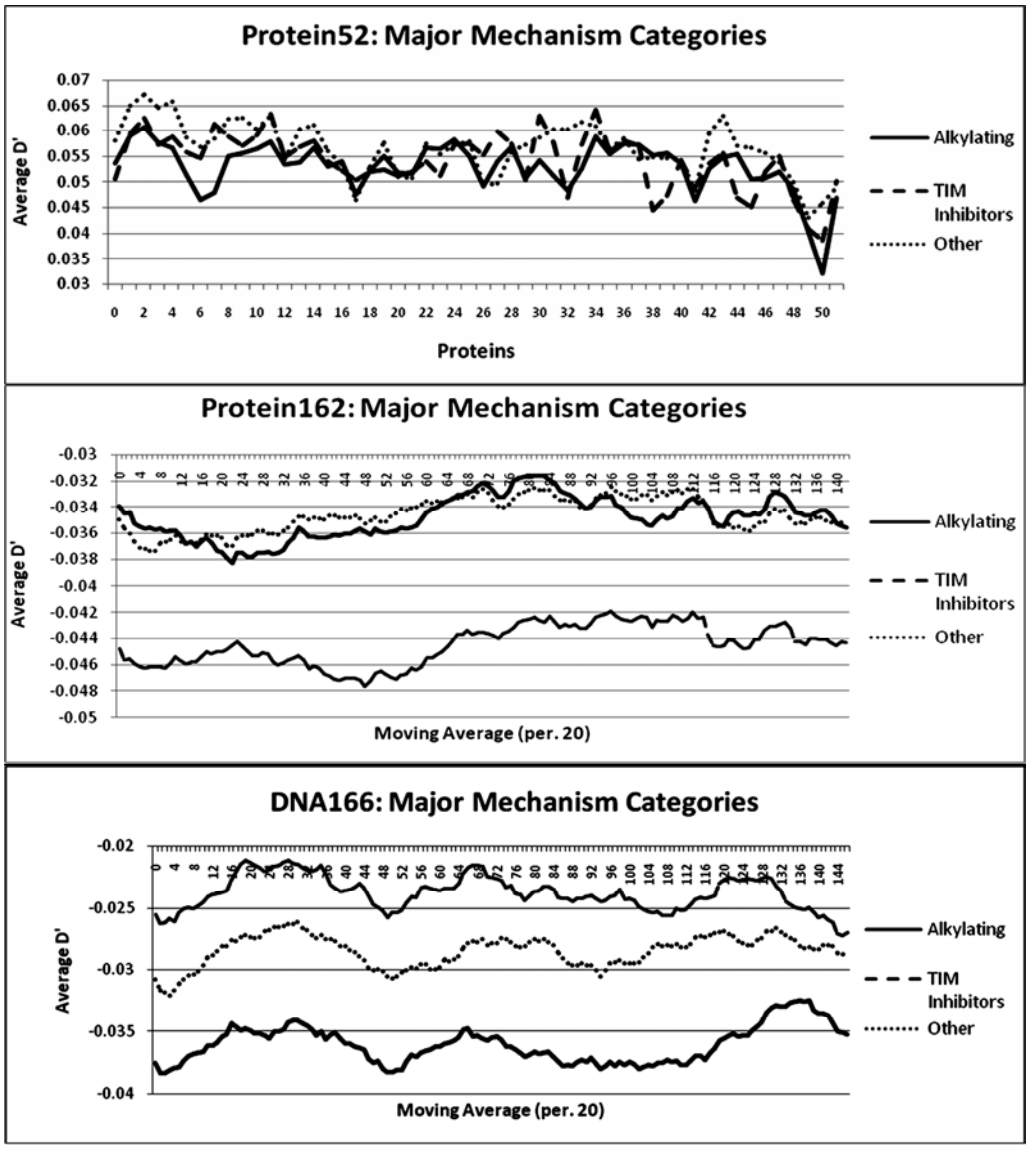


Figure 6. D' patterns according to major mechanism of action categories. Each dataset yields different levels of D' values between the three categories. The values produced by Protein52 are very similar in all three mechanism categories, whereas TIM Inhibitor values for Protein162 are distinct from the others. All three categories produce distinct values in DNA166. For readability, the plots of Protein162 and DNA166 are moving averages with a period of 20, whereas the plot of Protein52 shows the curve without a moving average.

trends of these three categories to be different for all three datasets.

In Protein52, we observed that while each of the larger categories carries unique drugs and drug mechanisms, the averaged D' values of all three of these categories were very similar, with the averages being 0.053 for the Alkylating category, 0.054 for the TIM Inhibitor category, and 0.06 for the Other category. All averaged D' values were positive.

The same analysis of the Protein162 dataset revealed that while the averaged D' values of the Alkylating and the Other categories where very similar, with average D' values –0.0348 and –0.0345 respectively, the TIM Inhibitors category was quite distinct with an average D' value of –0.044. All averaged D' values were negative.

For DNA166, all three curves have distinct averaged D' values, with the Alkylating category having an average D' of –0.0382, the Other category an average of –0.0286 and the TIM Inhibitors category an average of –0.0240. Again, all D' values were negative.

These plots, available in **Figure 6**, illustrate the benefit of using multiple datasets in this type of study. While all of our datasets are based on the NCI-60, they each provide unique insight into physical responses of the cell lines to the 118 anti-cancer drugs. If we were only using the DNA data, we might be tempted to claim

that these three mechanism categories produce distinctly different D' values, whereas if we were only using the data from Protein52, we might claim the opposite. It is only when a wide range of data are used in study that a holistic understanding of the effects of these 118 drugs becomes possible.

4. CONCLUSIONS

We found that each of our datasets provides a unique insight into the analysis of feature correlations in the study of the chemosensitivity of the NCI-60 cancer lines. Two of the three (Protein52, DNA166) have been used in previous studies for the prediction of chemosensitivity of cancerous cells, and though Protein162 was novel to this topic, we have shown that both Protein162 and Protein52 contain a number of protein markers that are both medically significant and highly correlated to individual protein markers highlighted in previous study [9].

In addition, we have shown D' to be an accurate measure of correlation in the context of feature selection for the first time.

5. ACKNOWLEDGEMENTS

We would like to thank Houghton College for its financial support.

REFERENCES

[1] Staunton, J. E., Slonim, D. K., Coller, H. A., Tamayo, P., Angelo, M. J., Park, J., Scherf, U., Lee, J. K., Reinhold, W. O., Weinstein, J. N., Mesirov, J. P., Lander, E. S., and Golub, T. R., (2001) Chemosensitivity prediction by transcriptional profiling, PNAS, **98**, 10787−10792.

[2] Potti, A., Dressman, H. K., Bild, A., Riedel, R. F., Chan, G., Sayer, R., Cragun, J., Cottrill, H., Kelley, M. J., Petersen, R., Harpole, D., Marks, J., Berchuck, A., Ginsburg, G. S., Febbo, P., Lancaster, J., and Nevins, J. R., (2006) Genomic signatures to guide the use of chemotherapeutics. Nature Medicine, **12**, 1294−1300.

[3] Paweletz, C. P., Charboneau, L., Bichsel, V. E., Simone, N. L., Chen, T., Gillespie, J. W., Emmert-Buck, M. R., Roth, M. J., Petricoin, E. F., and Liott, L. A., (2001) Reverse phase protein microarrays which capture disease progression show activation of prosurvival pathways at the cancer invasion front, Oncogene, **20**, 1981−1989.

[4] Lee, J. K., Bussey, K. J., Gwadry, F. G., Reinhold, W., Riddick, S. L. Pelletier, S. Nishizuka, G. Szakacs, J. Annereau, G., Shankavaram, U., Lababidi, S., Smith, L. H., Gottesman, M. M., and Weinstein, J. N., (2003) Comparing cDNA and oligonucleotide array data: Concordance of gene expression across platforms for the NCI-60 cancer cells, Genome, Biol., **4**, R82.

[5] Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C. F., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., and Brown, P. O., (2000) Systematic variation in gene expression patterns in human cancer cell lines. Nat. Genet., **24**, 227−235.

[6] Shankavaram, U. T., Reinhold, W. C., Nishizuka, S., Major, S., Morita, D., Chary, K. K., Reimers, M. A., Scherf, U., Kahn, A., Dolginow, D., Cossman, J., Kaldjian, E. P., Scudiero, D. A., Petricoin, E., Liotta, L., Lee, J. K., and Weinstein, J. N., (2007) Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integromic microarray study, Mol. Cancer Ther., **6**, 820−832.

[7] Ma Y., Ding Z., Qian Y., Shi X., Castranova V., Harner, E. J., and Guo L., (2006) Predicting cancer drug response by proteomic profiling, Clin. Cancer Res., **12**, 4583−4589.

[8] Scherf U., Ross D. T., Waltham M., Smith L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., Scudiero, D. A., Eisen, M. B., Sausville, E. A., Pommier, Y., Botstein, D., Brown, P. O., and Weinstein, J. N., (2000) A gene expression database for the molecular pharmacology of cancer, Nat. Genet., **24**, 236−244.

[9] Ma, Y., Ding, Z., Qian, Y., Wan, Y., Tosun, K., Shi, X., Castranova, V., Harner, E. J., and Guo, N. L., (2009) An integrative genomic and proteomic approach to chemosensitivity prediction, International Journal of Oncology, **34**, 107−115.

[10] Bussey, K. J., Chin, K., Lababidi, S., Reimers, M., Reinhold, W. C., Kuo, W., Gwadry, F., Ajay, Kouros-Mehr, H., Fridlyand, J., Jain, A., Collins, C., Nishizuka, S., Tonon, G., Roschke, A., Gehlhaus, K., Kirsch, I., Scudi ero, D. A., Gray, J. W., and Weinstein, J. N., (2006) Integrating data on DNA copy number with gene expression levels and drug sensitivities in the NCI-60 cell line panel, Mol. Cancer Ther., **5**, 853−867.

[11] Saeys, Y., Inza, I., and Larranaga, P., (2007) A review of feature selection techniques in bioinformatics. Bioinformatics, 23: 2507−2517.

[12] Lewontin, R. C., (1964) The interaction of selection and linkage, I. General considerations, Heterotic Models, Genetics, **49**, 49−67.

[13] Nishizuka, S., Charboneau, L., Young, L., Major, S., Reinhold, W. C., Waltham, M., Kouros-Mehr, H., Bussey, K. J., Lee, J. K., Espina, V., Munson, P. J., Petricoin, E., Liotta, L. A., and Weinstein, J. N., (2003) Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays, Proc. Natl. Acad. Sci., USA, **100**, 14229−14234.

[14] Hedrick, P. W., (1987) Gametic disequilibrium measures: proceed with caution, Genetics, **117**, 331−341.

[15] Leardic R., Boggia, R., and Terrile, M., (2005) Genetic algorithms as a strategy for feature selection, Journal of Chemometrics, **6**, 267−281.

[16] Zhu, Z., Ong, Y., and Dash, M., (2007) Markov blanket-embedded genetic algorithm for gene selection, Pattern Recognition, **40**, 3236−3248.

J. Biomedical Science and Engineering, 2009, 2, 516-520 **JBiSE**
doi: 10.4236/jbise.2009.27074 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Efficiency of diagnostic model to predict recurrent suicidal incidents in diverse world communities

Vatsalya Vatsalya[1], Kan Chandras[2], Shweta Srivastava[3], Robert Karch[1]

[1]Health Promotion Management American University, Washington D.C., USA; [2]Behavioral Sciences, Fort Valley State University, Fort valley GA, USA; [3]Department of Biology, Georgetown University, Washington D.C., USA.
Email: vvatsalya@gmail.com

Received 16 July 2009; revised 26 August 2009; accepted 27 August 2009.

ABSTRACT

Suicidal attempts have a very significant effect on the society, and they also reflect on the efforts of the supporting health care and counseling facilities; and the mental health professionals involved. The impact of suicide is further magnified by the needs of persons who attempt suicide multiple times, requiring emergency health care and rehabilitation. Preventing such activities becomes a major task for the support providing agencies as soon as patient with such tendencies are identified. There are repetitive traits that can be observed during the entire therapeutic program among the high-risk group individuals, who are susceptible to this kind of activity and such traits indicate for specific profiling. The aim of the instrument is to prevent the occurrence of the repetitive suicidal attempts of the patients in various world regions, which may have significantly higher and concerning suicide rates. This profile has been constructed on the various parameters recognized in the statistical analysis of the patient population, which have been identified or can be under treatment for their suicidal behavior. This instrument is developed to predict the probability of population segments who may attempt suicide and repetitively, by matching the parameters of the profile with that of the patient pool. Building a profile for the purpose of predicting behavior of this kind can strengthen the intervention strategies more comprehensively and reduce such incidents and health care requirements and expenses.

Keywords: Instrument; Parameters; Predictor; Risk Category; Suicidse

1. INTRODUCTION

Several world regions face the emerging concern of suicidal incidents and such growth patterns have suggested typical geographical factors playing major contributory roles [1]. Many regions so far have been identified with consistent year wise rising suicide rate and from such locations, Washington State has been included in the study to evaluate this phenomenon and derive a model to investigate and predict suicidal incidents in the general populations. Nationally, Washington State has the 16th highest suicide rate and suicide is the eleventh leading cause of death with 814 death reported in year 2005 with a rate of 13.1 incidents per 100,000 (age adjusted) of population, compared to the national US rate of approximately 11.0 [2]. Suicidal attempts can be five time or more than the suicide death rate [3] (**Figure 1**).

There have been consistencies in reports that even after getting initial medical and other support services, patients attempt suicide again in Washington State with similar rate as without the interventions [4]. During 2003–2005, males in Washington accounted for 79% of completed suicides, with the highest rate of suicide, observed in the group among the age group of 75 years and more [5]. In 2004; in the 15–24 age groups, suicide was the second most leading cause of death and altogether 17% deaths [6]. Nearly two-thirds of teens with clinical depression go unnoticed and may not get treated; among these, males 15 to 19 years old are five times more likely

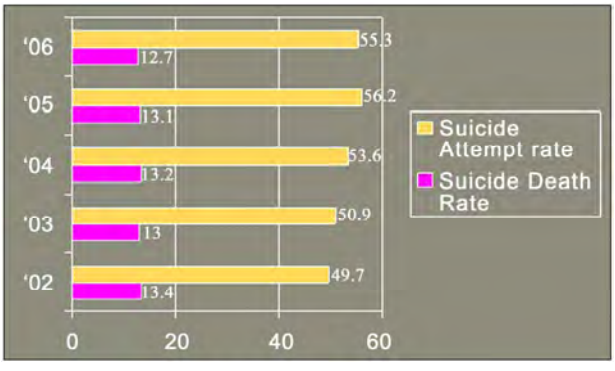


Figure 1. Yearly suicide rate representation per 100,000 populations (Suicide attempt and death rate versus years as x:y coordinates).

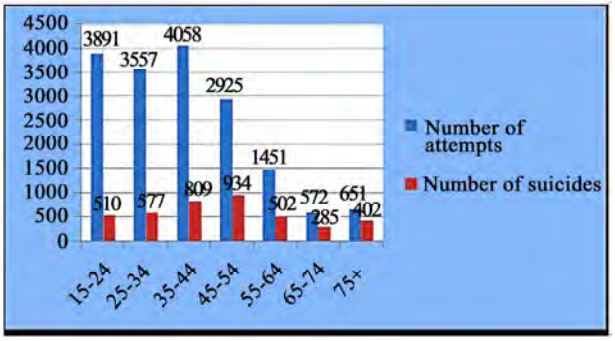


Figure 2. Age-group based comparisons of number of suicidal incidents (Age-groups versus number of incidents as x:y coordinates).

than females to complete suicides [7] (**Figure 2**). The frequency of suicide attempts can increase, if the causative factors are not resolved. Extent for emotional and physical pain cannot be assessed for individuals and their relatives, though approximately three billion US dollars of loss associated with the suicidal incidents has been recorded so far during the year 2002–2006 [8]. Preventive procedures have been particularly promoted to reduce such incidences, which have priority concern to precisely identify target population. Development of a predicting instrument, which can provide appropriate behavioral evaluation, can be one of the choices to this requirement.

This instrument is based on set of primary repetitive characters, which can provide a predictable analysis of such behavior among the high-risk group individuals, who are susceptible to commit or attempt suicide. The concept of predictor is to identify the patients with such tendency or ideations and inform the healthcare professionals and people concerned, to prepare and implement specific therapeutic plan to minimize such incidents further. A predictor profile for the purpose of interpreting specific behavior would strengthen the intervention strategies, reduce the frequency of such incidents, and regularize involvement of health care and financial resources.

2. METHODOLOGY

Data source for the profile analysis has been collected from yearly published hospitalization reports of suicide incidents from Washington State public health statistical data bank, Washington State Injury and Violence Prevention Program, 2008; and GIS from 2002 to 2006 year. Data from 15,826 cases with suicides and single or multiple identified attempts of history during the interval 2002–2006, between the age group of 15 and 75 + years, have been included in the study (**Figure 3**). Significant cohort parameters have been studied from the target population for the development of the instrument, namely gender, age groups, health condition, biological and environmental factors (namely economic, academic, social, ethnic and likewise). Statistical analysis has been performed to identify repetitive characteristics of the parameters with annual rate and frequency comparisons. An analog model is constructed utilizing the parameters and their characteristics under observations. The characteristics have been broadly grouped as parameters of interests and highly repetitive prevalent characteristics observed in the evaluation have been consigned under high risk category (Colored in red, **Figure 6**); average repeating characteristics as intermediate and less frequently occurring as low risk category. These parameters were constructed in the form a pathway with the promi-

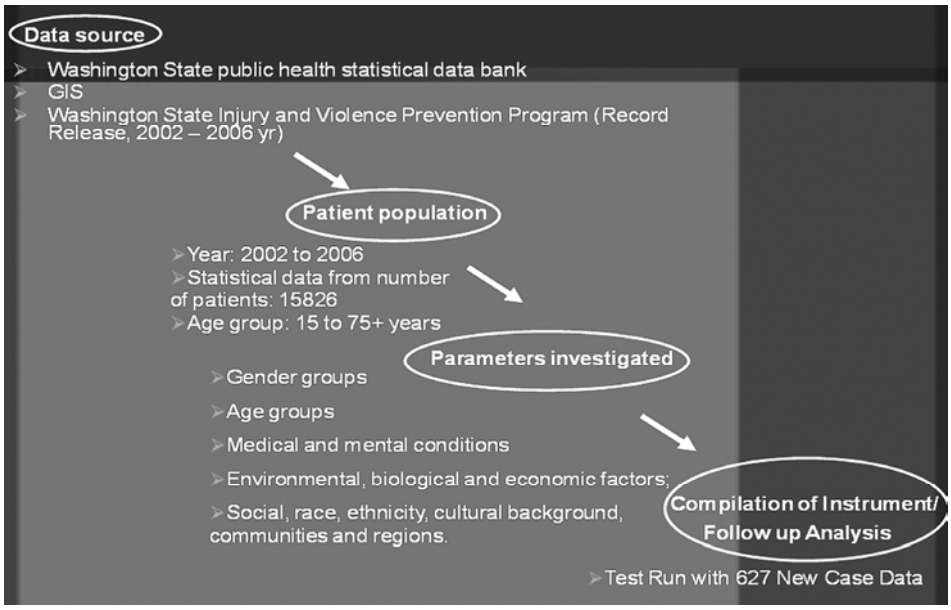


Figure 3. Methodology illustration for investigating model parameters.

nent parameters evaluated initially within which all the characteristics of each parameter undergo comparison for in track corresponding profile similarities.

Characteristics which have not demonstrated significant prevalence in the analysis or are not significantly identifiable in the data search and evaluations have not been included in the development of the analog at this stage of the study. A follow up evaluation assessment has been conducted for comparing this profile with the data from new reports to evaluate the predictability the evaluations and to develop the validity of the model.

3. RESULTS

The parameters are assessed in relation to the numbers of attempts and actual suicide incidents reported between year 2002 and 2006. The age groups 15–24 and 35–44 yrs have the maximum incidents of attempted suicides (**Figure 2**).

The leading conversion of the attempts into actual suicide is demonstrated by the age group 45–54 yrs and from 2002 year onwards, there is a steady increase in the rate of suicide and suicide attempts in this age group (**Figure 4**). Males in Washington accounted for 79% of total suicides, though majority of females have attempted suicides with a rate of 62%. There has been a consistent annual rise in the number of suicidal attempts as evident from annual health reports (**Figure 5**). Suicide comparison ratio among the adult male and female patient population rate has been measured as 7:2, which is evident from the finding that during 2003–2005 period. Among adolescent and young age groups, approximately 25% of attempted suicide by youth male results in death; whereas three–four percent deaths result in females. 15 to 19 year old males are five times more likely than females to commit suicides. Caucasian and American Indians demonstrate largest proportion of actual number of suicidal attempts and incidents as 14 every 100,000 resident population each.

Data reflects that in Caucasian and Native Indian groups, suicide is occurring with 14 per 100,000 rates each though African-American, Asian and Hispanic groups demonstrated non-significant rate of eight, eight and six consecutively. High competition, economic and social hardship can be observable factors among Caucasian though relationship cannot be established directly; partially available particulars are only confirmed presently for family predisposition, abuse, addiction and health concerns among with Native Americans too [9]. Suicide rate as observed in relation with academic background can be described with the ratio of 28:14:9 with 28 with patients having 12th grade or less education, 14 with some college level programs and 9 with graduation or post graduation. Linear regression of suicide rate has been observed with progression of academics in the statistical evaluation. Poisoning is identified as the most common means for attempted suicide cases among 15–24 and 35–54 age group whereas firearms are as the most common means among the young and adult groups for suicide deaths. Based on the results of statistical analysis and evaluations, several parameters and their characteristics have been identified and incorporated for the development of the model for the instrument with the order as gender, age group, ethnicity, health, education, environmental condition, and; exposure and access to perform self-fatality for the Washington region (**Figure 6**).

The construction of the analog utilizes major and recurring parameters first and tracks subsequent parameters and their characteristics in the descending order of evaluation. These Follow-up evaluations with parameters evaluated from various new case studies (a total of 627 case data) randomly, indicated results in compliance with an accuracy of more than 90% of the total cases, when the parameters are run through the analog model.

4. DISCUSSION

Age groups 15–24 and 35–54 year have demonstrated the majority of incidents of suicide and suicidal attempts. Parallel emergence of economic hardship can be projected as one of the reasons for the growth of suicide rate among the 45–54 year age group [10].

Higher suicide rate among is recorded among males, approximately 25% of attempted suicide by youth male results in death, whereas 3–4 % death results in females, though higher attempt rate among females is observed [11]. Health conditions (depression, substance abuse,

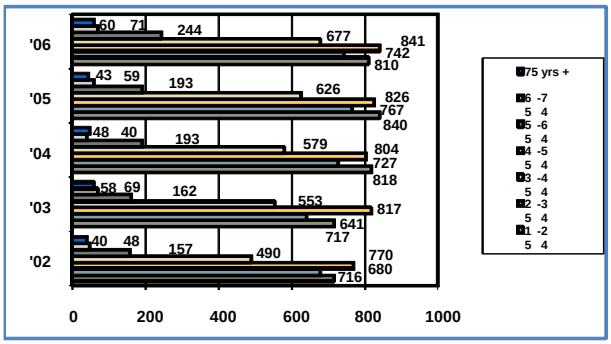


Figure 4. Yearly distribution of suicidal incidents in various age-groups (Number of incidents versus years as x:y coordinates).

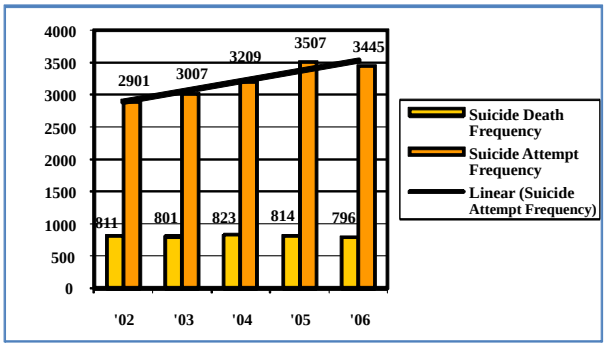


Figure 5. Yearly distribution of total number of suicidal incidents (Years versus number of incidents as x:y coordinates).

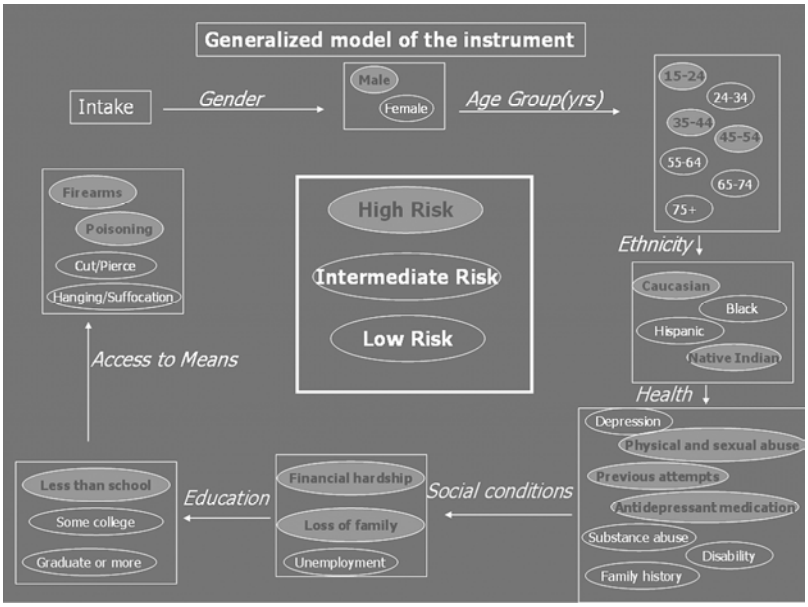


Figure 6. Analog construction of diagnostic model.

ideation and conditions like family history, sexual orientation and abuse, disability) show higher probability of suicidal incidents (**Figure 7**). Individuals with physical and sexual abuse have been observed to have four times higher suicide attempt rate than non-abused. Among African American communities, some of the major factors are community segregation, economic hardship and access to weapon; though cultural encapsulation and stress is prevalent among other minority groups. From 2002 year onwards, there is a steady increase yearly, in the number and rate of suicide incidents in the 35–54 age groups [12]; and the proportion of people living in poverty increased, though data is not sufficient enough to identify relation to economic hardship [13]. Post-Traumatic Stress Disorder, economic hardship, social constructs are being studied as an extension of the ongoing study.

5. CONCLUSIONS

The complexity of suicidal behavior and ideation requires multiple and immediate preventive approach and application of a predictor model can supplement targeting population, requiring adequate health care services to prevent suicide. Since much published material and clinical experience demonstrate a number of causative factors associated with suicide, the early identification and appropriate treatment of this condition is an important strategy for prevention. Improvements may be required for broadening data collection and investigatory techniques for possible suicidal attempts, social constructs and economic causes; biological and environmental factors and minority groups from both fatal and

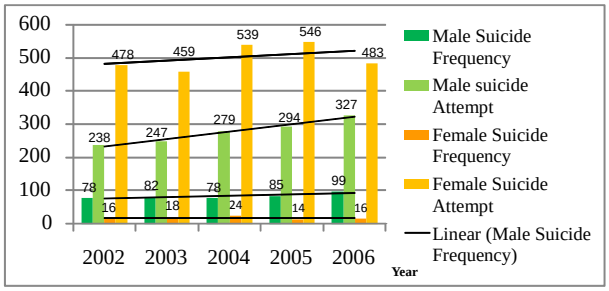


Figure 7. Yearly distribution of depression-originated youth suicidal incidents (Years versus number of incidents as x:y coordinates).

non-fatal suicide incidents. Utilization of the instrument at the key regions and community level; significantly social and community based agencies, academic institutions and care centers can be highly valuable and recommended. Design, planning and Implementation of intervention strategies for newly growing suicidal incidents in the 45–54 year age group should be focused appropriately. A preliminary investigation in various world regions with above-average suicide rate have identified points of interests namely Kutznetsk Basin, former USSR republics (between 1980 and 1995 citing economic instability); among indigenous peoples in both Australia and Canada in the last 20 years [1]. Depending upon the various factors and parameters in different world locations, with specific standardization, an appropriate construction of the predictor model and evaluation may be utilized for similar diagnostic procedures and further research may provide essentials and scope for development for area specific revisions.

REFERENCES

[1] World Report on Violence and Health: Self-Directed Violence (1999). Geneva, Wold Health Organization (document WHO/ HSC/PVI/99.11).

[2] Injury and Violence Prevention Program. Washington State Department of Health [online], 2007, http://www.doh.wa.gov/hsqa/emstrauma/injury/data-tables.

[3] Washington State Department of Health, (2004). Suicide chapter, The Health of Washington State 2004 Supplement. Olympia, WA, http://www.doh.wa.gov/HWS/HWS2004supp.htm.

[4] American Journal of Psychiatry [online], 2006, http://ajp.psyh.iatryonline.org/cgi.

[5] Health of Washington State Report. Washington State Department of Health [online], 2007, http://www.doh.wa.gov/HWS/default.htm.

[6] Simon, G. E., Savarino, J., Operskalski, B., and Wang, P. S., (2006) Suicide risk during antidepressant treatment, Am. J. Psychiatry, 163, 41–47

[7] Center of Health Statistics (Death data). Washington State Department of Health [online], (2007), http://www.doh.wa.gov/ehsphl/chs-data/death/dea_VD.htm.

[8] Center of Health Statistics (Hospital data). Washington State Department of Health [online], (2007), http://www.doh.wa.gov/EHSPHL/hospdata/default.htm.

[9] Mayer-Gross, W., Slater, E., and Roth, M., (1960) Clinical Psychiatry, London, Cassell.

[10] U.S. Centers for Disease Control and Prevention, National Centers for Injury Prevention and Control. (2007). Web-based Injury Statistics Query and Reporting System (WISQARS), Retrieved January 3, 2007 from http://www.cdc.gov/ncipc/wisqars.

[11] Rotheram-Borus, M. J., Piacentini, J., Cantwell, C., Beline, T. R., and Sone, J. (2000). The 18-month impact of an emergency room intervention for adolescent female suicide attempters. Journal of Counseling and Clinical Psychology, 68(6), 1081–1093.

[12] Centers for Disease Control and Prevention (CDC). Web-based Injury Statistics Query and Reporting System (WISQARS) [Online], (2005) National Center for Injury Prevention and Control, CDC (producer), Available from http://www.cdc.gov/ncipc/wisqars/default.htm.

[13] Geographic Information Systems. Division of Information Resource Management, Washington State Department of Health [online], 2007, http://www.doh.wa.gov/gis/gisdata.htm.

J. Biomedical Science and Engineering, 2009, 2, 521-525 **JBiSE**
doi: 10.4236/jbise.2009.27075 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

An oversampling system for ECG acquisition

Yu Zhou

School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai, China.
Email: Zyusst@yahoo.cn

Received 31 July 2009; revised 31 August 2009; accepted 3 September 2009.

ABSTRACT

Traditional ECG acquisition system lacks for flexibility. To improve the flexibility of ECG acquisition system and the signal-to-noise ratio of ECG, a new ECG acquisition system was designed based on DAQ card and Labview and oversampling was implemented in Labview. And analog signal conditioning circuit was improved on. The result indicated that the system could detect ECG signal accurately with high signal-to-noise ratio and the signal processing methods could be adjusted easily. So the new system can satisfy many kinds of ECG acquisition. It is a flexible experiment platform for exploring new ECG acquisition methods.

Keywords: ECG Acquisition; Oversampling; DAQ; Labview

1. INTRODUCTION

ECG acquisition has been a mature technology since Doctor Willam Enthoven in Holland detected human ECG firstly. ECG can reflect the physiology status and pathology status of heart. Heart diseases can be detected in time by ECG. So ECG is used widely in clinic [1]. ECG can be divided into regular ECG, body surface potential mapping, high frequency ECG, ventricular late potentials electrogram, surface Hisbundle electrogram and etc. Their detection objects are all with low amplitude but distributing in different low frequency ranges [2]. Traditional ECG acquisition systems are designed according to different detection objects. Their analog signal conditioning circuit including multistage amplifiers and filters focus on special object. The disadvantage is that the special ECG acquisition system only can detect special object. So it lacks for flexibility and wastes resources.

The DAQ card and Labview produced by NI Company provide a flexible solution for data acquisition and processing [3,4]. DAQ card can detect signal accurately with high speed. And digital signal processing can be implemented easily in Labview. There are many works using DAQ card and Labview in ECG. But they almost focus on the ECG's analysis and processing [5,6].

To resolve the problem about flexibility, analog signal conditioning circuit is improved on and a new ECG acquisition system is developed with DAQ card and Labview.

Oversampling technology can exert the advantage of high speed A/D and improve the signal-to-noise ratio of signal detected. There has been detailed discussion about the use of oversampling in biopotential detection [1,7,8]. This technology is implemented in Labview and used in the new ECG acquisition system to improve the signal-to-noise ratio of ECG.

2. SYSTEM DESCRIPTION

The oversampling system for ECG acquisition shown in **Figure 1** is consisted of analog signal conditioning circuit, DAQ card and Labview.

Research supported by The Special Funds for Scientific Research for Selecting and Training Outstanding Young Teachers in Shanghai Universities (slg08029).

The analog signal conditioning circuit condition the ECG signal detected by electrodes to an analog signal with certain amplitude and frequency restriction. The analog signal is converted to digital signal by DAQ card with high speed sampling rate. And oversampling and other processing methods for the digital signal are implemented in Labview.

3. ANALOG SIGNAL CONDITIONING CIRCUIT DESIGN

Analog signal conditioning circuit included the circuit from the electrodes to A/D input of DAQ card. To fit for the need of the oversampling system for ECG acquisition, traditional circuit was improved on.

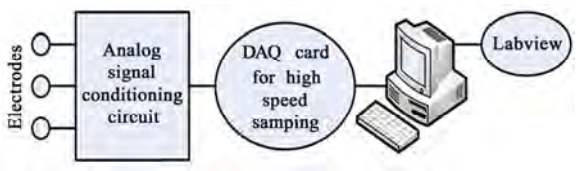


Figure 1. Lock diagram of the oversampling system for ECG acquisition.

3.1. Circuit Structure

The circuit was shown in **Figure 2**. As traditional circuit, it included protection circuit, amplifying circuit based on AD620, high-pass circuit for eliminating polarization voltage, main amplifier, low-pass circuit and driven-right-leg circuit. But the parameters set of analog filter was different. In traditional circuit the filter parameters were set according to the object detected and couldn't be adjusted. Thus it lacked for flexibility. In oversampling system for ECG acquisition, the analog filter parameters were set with 0.03Hz–4KHz. Although the filter parameters were also fixed, the frequency range was wide and special digital filter parameters could be set further in Labview. This scheme guaranteed the currency of the system.

3.2. Isolated Circuit and Power

DAQ card and Labview worked based on PC. Isolated circuit was needed to protect the object detected from electric shock. HCNR200 was an optical coupler with high linearity and fit for the oversampling system for ECG acquisition. The isolated circuit based on HCNR 200 was shown in **Figure 3**. Isolated power supply was from 9V battery. LM2576 and ICL7660 got energy from 9V battery and produced ±5V. Non-isolated power supply was from DC supply.

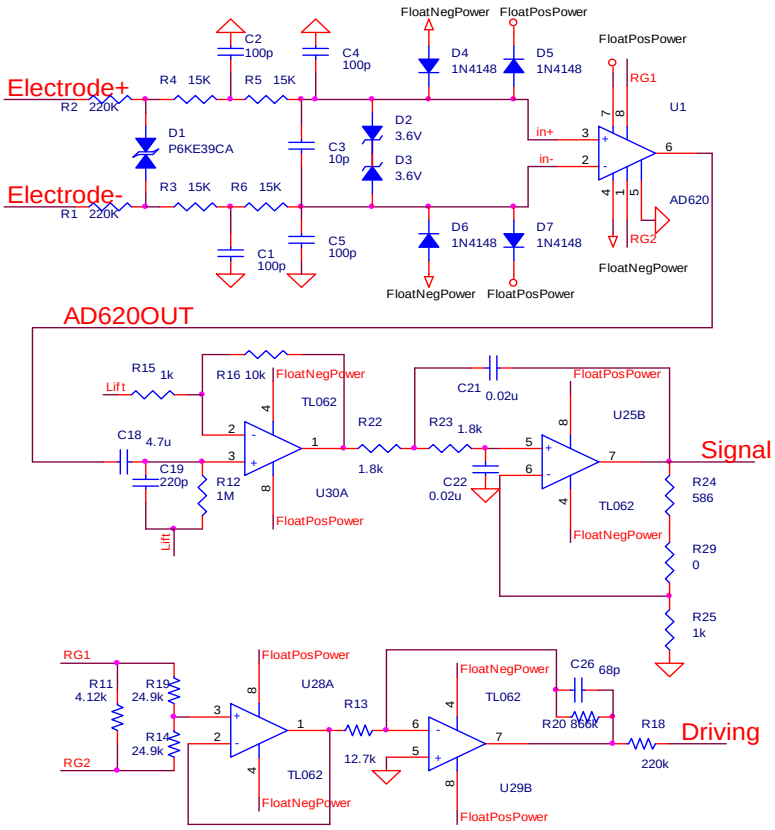


Figure 2. Analog signal conditioning circuit.

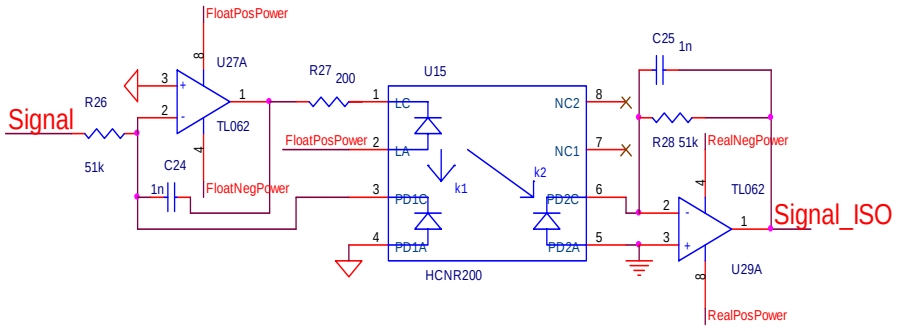


Figure 3. Isolated circuit.

4. LABVIEW PROGRAM DESIGN

Labview program was the main body of the oversampling system for ECG acquisition. It included the set of DAQ card, digital filter, oversampling and etc.

4.1. Setting DAQ Card

DAQ6221 was a data acquisition card with PCI interface. Its highest sample rate for single channel was 250KHz. In the oversampling system for ECG acquisition, the sample rate was set with 250KHz and continuous sampling mode was chosen. And the reference voltage for A/D in DAQ card could be chosen. ±5V was set because the analog signal conditioning circuit had amplified the signal enough.

4.2. Signal Filters

Filters including low-pass filter and 50Hz notch filter were needed in Labview. In the basis of analog filters with wide range, special digital filters were used for special object detected in Labview. It was the flexibility of the oversampling system for ECG acquisition. 50Hz-interference caused by power line was an important problem in biopotential detection [9,10]. Notch filter was the tool to eliminate the interference. In Labview digital 50Hz notch filter with more orders could get much better result than analog notch filter. At the same time the delay caused by digital filter was very small because of the high-speed cpu of PC.

In experiments, for low-pass filter, type was set with Butterworth, order was set with 50 and cut-off frequency was set with 90Hz. For notch filter, type was set with Bessel, order was set with 50.

4.3. Oversampling

The criteria of using oversampling were that the original sample rate was high enough and the sample rate after oversampling could satisfy Nyquist sampling theorem.

The sample rate in DAQ card was set with 50KHz. So the sample rate after oversampling was 50KHz/256≈ 195 Hz if oversampling coefficient was set with 256. It was enough for ECG to satisfy Nyquist sampling theorem [1,7]. But if the oversampling coefficient was increased, the Nyquist sampling theorem couldn't be satisfied. If sampling rate was improved with 4 multiples, the resolution of ADC could be improved with 1 bit [1,11]. So the resolution of DAQ card' ADC was improved with 4 bits. Continuous sampling mode was chosen for DAQ card. So continuous 256 samples were sampled and accumulated and output. When this cycle went on, oversampling result was get. Oversampling coefficient and filter parameters were both decided by the object detected and could be set in Labview according to need. In experiments, 250KHz was also used for sample rate in DAQ card sometimes. Then the oversampling coefficient should be changed to 1024 and the sample rate after oversampling was 244Hz. It was also right for ECG.

5. EXPERIMENTS DESIGNS AND RESULTS

Some experiments were used to validate the design.

5.1. Oversampling Experiment

A sine wave with 20Hz and 5mV Vpp similar to ECG produced by signal generator was input to DAQ card. The results were shown in **Figure 4**. Oversampling result was better than the direct sampling result apparently.

FFT was applied for the two waveforms in **Figure 4** and the results were shown in **Figure 5**. The main com-

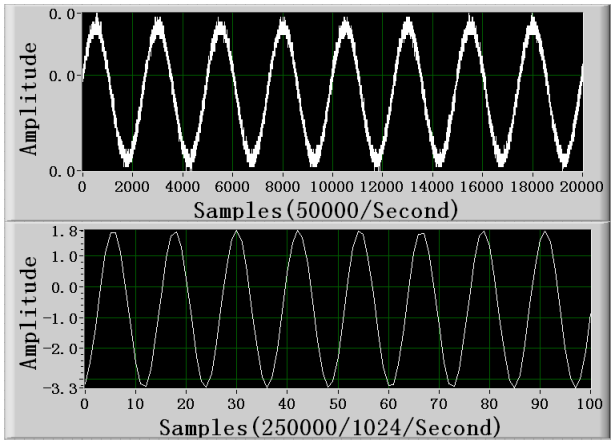


Figure 4. Direct sample result (upside, sampling rate: 50KHz) vs. Oversampling result (downside, sample rate: 250KHz, Oversampling coefficient: 1024).

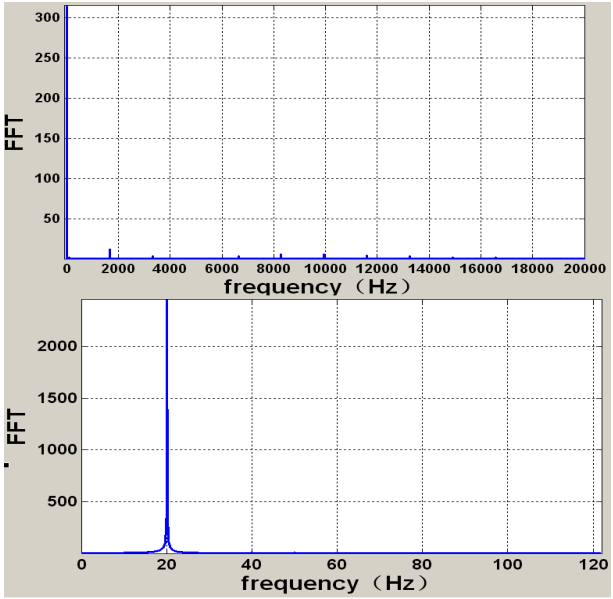


Figure 5. FFT for direct sample result (upside) vs. FFT for Oversampling result (downside).

ponents in both FFT were in 20Hz. But there were many components in other frequencies in the FFT for direct sample result apparently. The FFT for oversampling result was clearer. Through careful observation, the value in 20Hz in FFT for direct sample result was 316 and the value in 50Hz which was the frequency of power line interference in FFT for direct sample result was 1.46. The corresponding values in FFT for oversampling result were 2454 and 10.3 respectively. Only considering these components, the signal to noise ratio were 216 for direct sample result and 238 for oversampling result. So the FFT for oversampling result was clearer and its signal to noise ratio was higher. These were the reasons that oversampling result was better than direct sample.

5.2. ECG Acquisition Experiment with 3 Electrodes

In the oversampling system for ECG acquisition, acquisition mode with 3 electrodes could be chosen where driven-right-leg electrode was included. Two electrodes for detection were pasted in the left wrist and right wrist respectively. The detection results were shown in **Figure 6**. The 1st was direct sampling signal sampled by DAQ card, the 2nd was signal after oversampling, the 3rd was signal after low-pass filter and the 4th was signal after low-pass filter and notch filter from upside to downside. It could be observed that the effect of the digital filters was good and oversampling could improve the signal-to-noise ratio of ECG. The components of ECG were shown clearly.

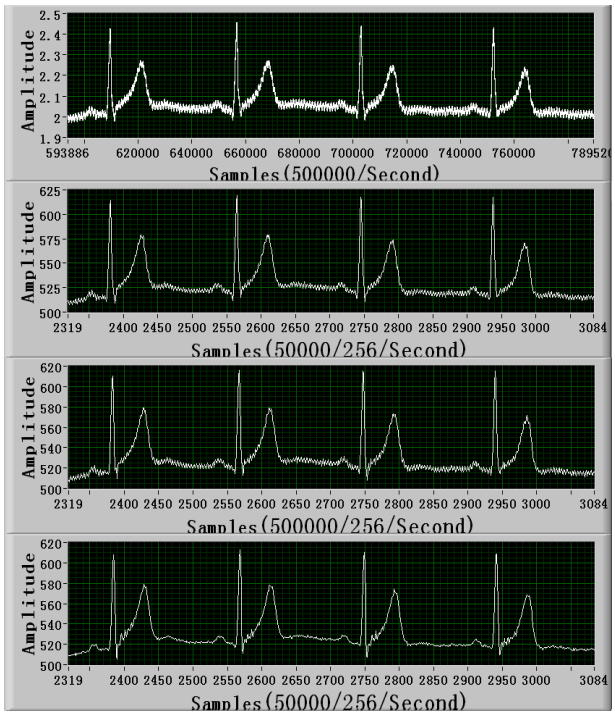


Figure 6. Sampling results with driven-right-leg circuit.

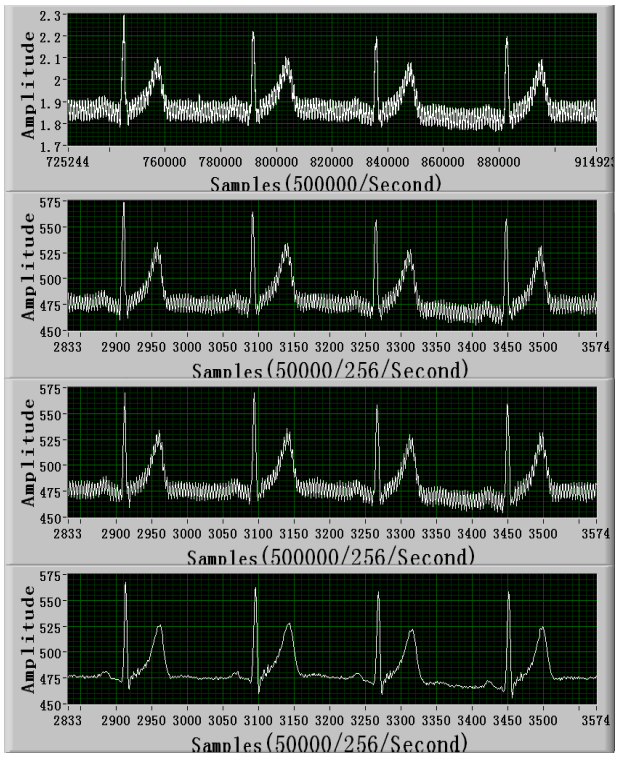


Figure 7. Sampling results without driven-right-leg circuit.

5.3. ECG Acquisition Experiment with 2 Electrodes

Also acquisition mode with 2 electrodes could be chosen for the oversampling system for ECG acquisition where driven-right-leg electrode was not included. Two electrodes were pasted in the left wrist and right wrist respectively. The detection results were shown in **Figure 7**. The 1st was direct sampling signal sampled by DAQ card, the 2nd was signal after oversampling, the 3rd was signal after low-pass filter and the 4th was signal after low-pass filter and notch filter from upside to downside. It could be observed that the power line interference was worse due to lack of driven-right-leg circuit before notch filter. But after notch filter ECG became very clear and the effect could be equivalent with the effect with driven-right-leg circuit. So profiting from digital filter with high orders, the detection result from 2 electrodes was satisfying. And oversampling could improve the signal-to-noise ratio of ECG further.

5.4. ECG Acquisition Experiment in Left Arm

In ECG monitoring field, it was a problem waiting to be resolved that how to acquire ECG with 2 electrodes expediently. The correlative experiment was attempted. An electrode was pasted in left wrist and the other was pasted in the left shoulder. The results were shown in **Figure 8**. It was observed that the ECG could be detected from one arm and oversampling made result a little clearer.

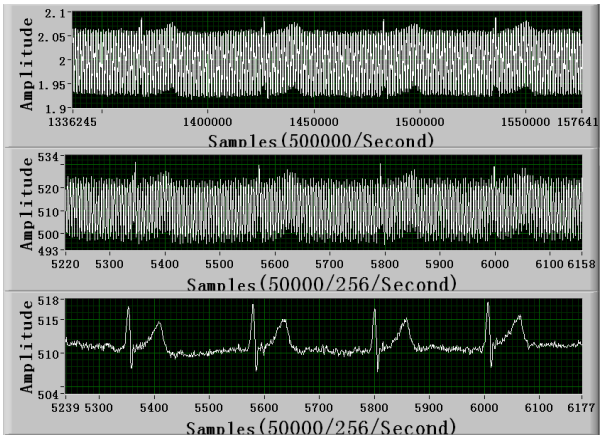


Figure 8. Sampling results from one arm (direct sample result (upside), Oversampling result (middle), result after oversampling and low-pass filter and notch filter (downside)).

6. CONCLUSIONS

The results of experiments indicate that the new oversampling system for ECG acquisition could detect ECG accurately, digital signal processing can be implemented easily in the system with DAQ card and Labview, ECG can be detected clearly with both 3 electrodes mode and 2 electrodes mode profiting from digital filters with high orders, oversampling improve the signal-to-noise ratio of ECG especially in one arm ECG detection. And the new system is flexible because the signal processing methods can be adjusted easily in Labview such as filter parameters and oversampling coefficient. So the oversampling system for ECG acquisition with flexible set and high signal-to-noise ratio can satisfy many kinds of ECG detection.

REFERENCES

[1] Li, G., Zhang, L. J., Lin, L., and He, F., (2008) Research of Bioelectricity Signal Detection Based on Oversampling, Acta Electronica Sinica, **36**, 1465–1467.

[2] Wang, B. H., (2003) Biomedical Measurement and Instrument, 1st Edition, Fudan Press, Shanghai.

[3] Koutroulis, E. and Kalaitzakis, K., (2003) Development of an integrated data-acquisition system for renewable energy sources systems monitoring. Renewable Energy, **28**, 139–152.

[4] Yeung, K. and Huang, J., (2003) Development of a remote-access laboratory: a dc motor control experiment. Computers in Industry, **52**, 305–311.

[5] Wei, Y. C., Lee, Y. H., and Young, M. S., (2008) A Portable ECG Signal Monitor and Analyzer. The 2nd International Conference on Bioinformatics and Biomedical Engineer- ing. Shanghai, China, pp.1336–1338.

[6] Yeon, N., Sung, P., Kyu, H., *et al.*, (2006) A Study of Significant data Classification between EDR extracted and frequency analysis of Heart Rate Variability from ECG using Conductive textile. World Congress on Medical Physics and Biomedical Engineering. Seoul, Korea, pp.4100–4103.

[7] Li, G., Zhang, L. J., Lin, L., and He, F. (2008) Weak Signal Detection Based on Over-Sampling and Saw-Tooth Shaped Function. ACTA ELECTRONICA SINICA, **36**, 756–759.

[8] Goes, J., Paulino, N., Pinto, H., Monteiro, R., Vaz, B., and Garcao, A. S., (2005) Low-power low-voltage CMOS A/D sigma-delta modulator for bio-potential signals driven by a single-phase scheme. Circuits and Systems I: Regular Papers, IEEE Transactions on, **52**, 2595– 2604.

[9] Huhta, J. C. and Webster, J. G. (1973) 60-Hz Interference in Electrocardiography. IEEE Transactions on Biomedical Engineering, **20**, 91–101.

[10] Spinelli, E. M. and Mayosky, M. A., (2005) Two-electrode biopotential measurements: power line interference analysis. Biomedical Engineering, IEEE Transactions on, **52**, 1436–1442.

[11] Oppenheim, A. V., Schafer, R. W., and Buck, J. R., (1999) Discrete-time Signal Processing, 2nd Edition, Prentice Hall, New Jersey.

J. Biomedical Science and Engineering, 2009, 2, 526-531 **JBiSE**
doi: 10.4236/jbise.2009.27076 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Enzymatic pretreatment and microwave extraction of asiaticoside from Centella asiatica

Cheng-Xiao Wang, Wei Han*, Liang Fan, Chun-Li Wang

Engineering Center for Traditional Chinese Medicine Modernization, East China University of Science and Technology, Shanghai, China.
E-mail: *whan@ecust.edu.cn

Received 3 July 2009; revised 13 July 2009; accepted 19 July 2009.

ABSTRACT

The extraction of asiaticoside from Centella asiatica by enzymatic pretreatment and microwave extraction (EPME) was studied in this article. The effects of several important factors such as temperature of enzymatic pretreatment, liquid to solid ratio and microwave radiation time were investigated by quadric regression orthogonal design experiment and were analyzed by response surface. An extraction model with well forecast performance was then established. The results indicate that the optimum extraction condition was as follows: liquid to solid ratio was 36mL/g, temperature of enzymatic pretreatment was 45℃, enzymatic time was 30min, and microwave radiation time was 110s. On such conditions the yield was 27.10%.

Keywords: Asiaticoside; Enzymolysis; Microwave Extraction; Centella Asiatica

1. INTRODUCTION

Centella asiatica (L.) Urban, a perennial herb belonging to the umbelliferae family, is well known as a traditional Chinese herbal medicine. Its conventional efficacy is clearing away heat and toxic, inducing diuresis and reducing edema. The major active constituents of *Centella asiatica* are asiaticoside and madecassoside. It has been used for the treatment of hot and humid jaundice, traumatic injuries, infectious hepatitis and dermatosis [1]. Currently, some conventional extraction methods are mostly adopted for the extraction of asiaticoside [2,3,4], such as aqueous extraction and ethanol extraction.

In recent years, new kinds of extraction techniques appeared, including enzymolysis and microwave extraction. The former has impressive effects with characteristics of high catalytic efficiency, high specificity, mild reactive conditions and preserving the original efficacy of active compounds to the maximum [5]. The later method has many advantages, such as shorter time, less solvent, higher extraction rate and better products with lower cost [6,7]. However, the application of the combination of these two methods on plant materials was rarely reported.

For the purpose of improving the efficient of asiaticoside extraction, reducing the cost, the EPME method is employed in this article, quadric regression orthogonal design is adopted to investigate the effects of three main extraction parameters including temperature of enzymatic pretreatment, liquid to solid ratio and microwave radiation time on the yield of asiaticoside, and optimum extraction process is worked out.

2. MATERIALS AND METHODS

2.1. Equipments and Reagents

An ER-692 microwave oven (as shown in **Figure 1**) with a power output of 650W, operating at 2450MHz, was mechanically modified to fit a reflux system that enables extraction performed under atmospheric settings and permits solvent containment. The extraction vessel was a 250-mL three-necked round-bottomed flask connected to a water condenser.

Thermostat with magnetic stirrer (Model DF-101S, Yuhua Experimental Apparatus Co., Shanghai) was used as enzymolysis device. UV spectrophotometer (Model 1900PC, Yayan Electronic Co., Shanghai) was used for determination.

Dry *Centella asiatica* (fitted for the Chinese Pharmacopoeia); Asiaticoside standard (Provided by National Institute for the Control of Pharmaceutical and Biological Products, Beijing); Cellulase (≥400U/mg); Anhydrous ethanol and concentrated sulfuric acid used in the experiment were all of analytical grade.

2.2. Analytic Method of Asiaticoside

The concentration of the asiaticoside in this process was determined by ultraviolet spectrophotometry [8] and the result was expressed as extraction yield, i.e. unit extrac-

tion quantity (g asiaticoside/g *Centella asiatica*).

The asiaticoside standard sample solution (was diluted by anhydrous ethanol) and extraction solution were scan ed at 200-400nm respectively, there were the same absorption peak at 277nm (**Figure 2**), which was close to the literature values [9]. So 277nm was chosen for the UV wavelengths.

The asiaticoside standard sample solution (concentration: 92μg/mL) was precisely measured at 0.0, 0.5, 1.0, 1.5, 2.0, 2.5mL, and put into a 10mL volumetric flask. Firstly, volatilized out the solvent ethanol in the boiling water bath, and then added the 2mL concentrated sulfuric acid (H_2SO_4) after the flask cooling to the ambient temperature, heated for 30min in 80℃ water bath. Finally, added anhydrous ethanol to the scale after the flask getting to the ambient temperature. According to the ultraviolet spectrophotometry, the prepared solution was measured at 277nm. Regression equation and correlation coefficient were *y=43.40x-2.49* and r=0.9992 (n=7) respectively. The linear range was 4.6~23.0μg/mL.

The 0.5mL test sample solution was accurately measured and placed into a 10mL volumetric flask, following the preparation method of regression equation. The absorbance of test sample was determined, and the yield of asiaticoside in the test sample was calculated in accordance with the following equation:

Yield of asiaticoside (%, w/w)=

$$\frac{(43.40A-2.49)Vn}{3\times10^6}\times100\% \quad (1)$$

where, A, the absorbance of asiaticoside in test sample; V, volume of solvent, mL; n, diluted times.

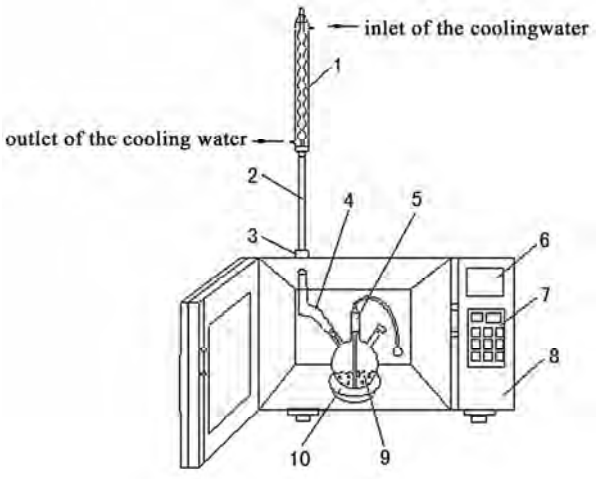


Figure 1. Microwave equipment diagram; 1-Water condenser; 2-Air condenser; 3-Copper tube; 4-Tailored tube; 5-Air agitator; 6-Status display; 7-Microwave oven timer; 8-Microwave oven; 9-Flask; 10-Base of flask.

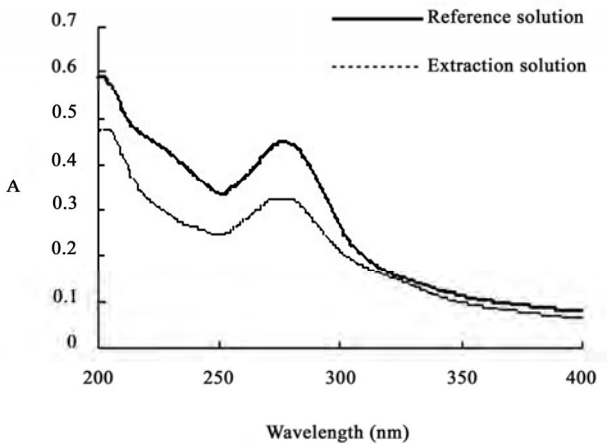


Figure 2. UV spectra of reference solution and extraction solution.

2.3. Enzymatic Pretreatment and Microwave Extraction

The dry *Centella asiatica* (sieved through 10 screen mesh) 3.0g was accurately weighed and placed into a three-neck flask with 3% cellulase solution (dissolved by deionized water), then the deionized water as solvent was added in according to a certain ratio (mL/g) of material to solvent volume, and the mass of system was weighed.

Three-neck flask was put into the thermostat with magnetic stirrer, setting enzymolysis time at 30min according to the pre-experiment, while the temperature of enzymatic reaction and stirring speed were set at certain values. Then the flask was taken out and placed into the microwave oven immediately. The radiation exposure was 30s for preventing the serious evaporation of solvent. At the end of each exposure, the system was brought back to ambient temperature during 2~3 min interval by cooling it with a water bath. An extraction cycle was defined as the combination of phases of radiation and phases without radiation in which the solvent cooled. The sum of radiation exposure of processing extraction cycles served as the overall intensity of microwave radiation. The extraction solution was agitated with an air pump to promote heat uniformity while exposing to microwave radiation. After the radiation, the flask was taken out to weigh the total mass again, and the lost weight was supplemented. The extract was filtered through 0.45μm millipore filter, and then abandoned the initial filtrate, added 0.5mL subsequent filtrate to a 10mL volumetric flask with anhydrous ethanol as the test sample.

3. RESULTS AND DISCUSSIONS

3.1. Effect of Temperature of Enzymatic Pretreatment on EPME

As shown in **Figure 3**, the results indicate that the yield of asiaticoside was increased with the increase of temperature of enzymatic pretreatment, reached a high point

at 45℃. Because the temperature of enzymatic pretreatment was a significant factor in the process of enzymolysis, it affected the enzyme activity as well as the rates of enzyme-catalyzed reactions. When the temperature was lower than 45℃, the increase of temperature can improve the cellulase activity, accelerate the degradation of cytoderm. While the temperature of enzymatic pretreatment was higher than 45℃, the cellulase activity was decreased, leading to the reduction of the yield. Therefore the temperature of enzymatic pretreatment for 45℃ was used.

3.2. Effect of Liquid to Solid Ratio on EPME

As it is known, the liquid to solid ratio is very important in the extraction. From the perspective of mass transfer, it mainly affects the concentration gradient between liquid phase and solid phase. **Figure 4** shows that the yield of asiaticoside was increased with the increase of liquid to solid ratio. After the peak, the yield of asiaticoside was decreased with the increase of liquid to solid ratio. The higher liquid to solid ratio, the longer time for the solution elevated to the same temperature was required. At the same microwave radiation, the temperature of the system with larger amount of solvent was relatively lower; the solute diffusion would be affected. So the ratio for 30mL/g was chosen.

3.3. Effect of Microwave Radiation Time on EPME

Figure 5 described that the extraction yield sharply increased before 60s and was asymptotic to a steady value during 60~120s, then falls down after 120s. At the preliminary stage of extraction, the velocity of molecular thermal motion quickened and the asiaticoside was quickly separated from the cell into solution. The extraction process had tended towards equilibrium since 60s. When

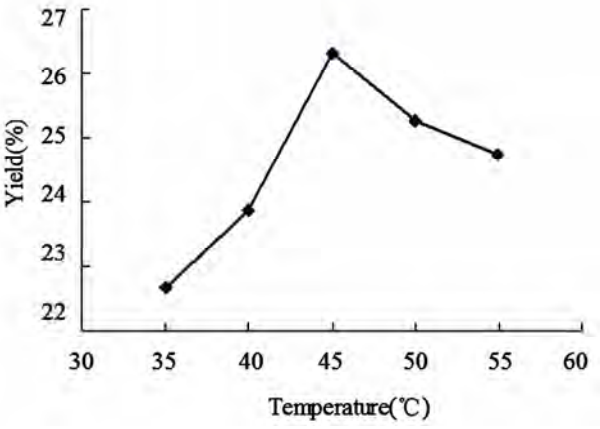


Figure 3. Effect of temperature of enzymatic pretreatment on EPME.

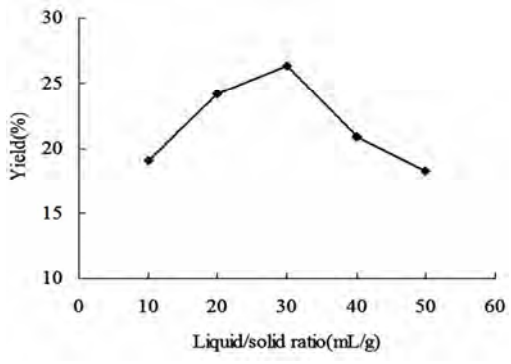


Figure 4. Effect of liquid to solid ratio on EPME.

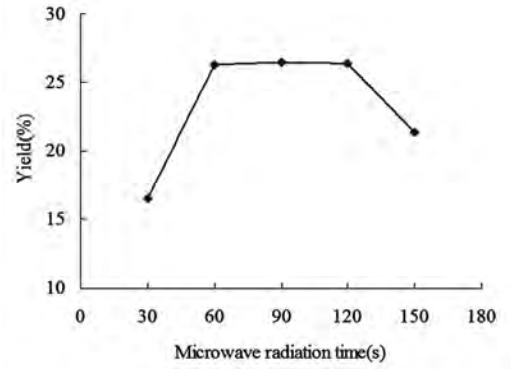


Figure 5. Effect of microwave radiation time on EPME.

the microwave radiation time exceeded 120s, the system was hyperthermal, the vaporization reflux of solution increased. All of above brought about the reduction of effective contact interval of solvent and plants, and caused the decrease of the thermal effects. In addition, the extension of microwave radiation time would increase the consumption of energy. Thus the microwave radiation time for 60s was used in the experiment.

3.4. Quadratic Regression Orthogonal Design Results

The quadratic regression orthogonal design was employed to evaluate the relevance of the three main extraction factors including enzymatic temperature, liquid to solid ratio and the microwave radiation time, while other factors including grinding degree of plant material, concentration of enzyme, enzymatic pretreatment time and pH of the solvent were constant according to the pre-experiments. The multivariate study allows the identification of interactions between variables and provides a complex exploration of the experimental domain to be studied with a number of experiments optimized.

The three key variables studied were pointed at five separate coded levels [11], –1, 0, +1, +γ (=1.682) and their values were selected on the basis of previous experiments. The natural values and coded levels used in this multivariate study are presented in **Table 1**.

Table 1. Factors of orthogonal experiment.

	X_1-Microwave	X_2-Liquid to	X_3-Enzymatic
-γ	9	13	28
-1	30	20	35
0	60	30	45
1	90	40	55
γ	111	47	62
Δ	30	10	10

The statistical analysis software was applied in this experiment to establish a regression **Eq.2**, in this equation, the terms of Y, X_1, X_2, and X_3 respectively represent the yield of asiaticoside, microwave radiation time, liquid to solid ratio and enzymatic temperature:

$$Y=25.8963+1.8150X_1-0.7646X_2+0.3464X_3 +0.9363X_1X_2+0.4688X_1X_3+0.1363X_2X_3 -0.7065X_1^2-0.6852X_2^2-1.7547X_3^2 \quad (2)$$

Analysis of variance was carried out in order to test the signification of the regression model. Thus, various statistical data such as sum of squares (SS), mean squares (MS), F-ratio were given in **Table 2**.

The F-ratio in **Table 2** was the ratio of the meansquared error to the pure error obtained from the replicates at the design center. The significance of the F-ratio depends on the number of degrees of freedom (d. f.) in the model. Thus, the effects lower than 0.05 in this column were significant.

As shown in **Table 2**, $F_{Regression}=10.866>F_{0.05}$ (9, 8)=3.388, the regression of (2) was significant, while $F_{Lack\ of\ fit}=3.478<F_{0.05}$ (5, 8) =3.687, the lack of fit was non-significant. Therefore, (2) had well predictivity under the experimental condition. In the test of regression coefficient, F-ratio for terms X_3, X_1X_3, X_2X_3 (1.253, 1.344, 0.114) were lower than $F_{0.05}$ (1, 8) =5.318, so these terms were not significant. On the contrary, the F-ratio for other terms (X_1, X_2, X_1X_2, X_1^2, X_2^2 and X_3^2) which were higher than $F_{0.05}$, indicated the significance of these terms. As the orthogonality of this experiment, the insignificant terms were cut out to simplify the (3):

$$Y=25.8963+1.8150X_1-0.7646X_2+0.9363X_1X_2 -0.7065X_1^2-0.6852X_2^2-1.7547X_3^2 \quad (3)$$

Table 2. Variance analysis of test results.

Source	d.f.	SS	MS	F	$F_{0.05}$
X_1	1	44.995	44.995	34.409	5.318
X_2	1	7.985	7.985	6.106	
X_3	1	1.639	1.639	1.253	
X_1X_2	1	7.012	7.012	5.363	
X_1X_3	1	1.758	1.758	1.344	
X_2X_3	1	0.148	0.148	0.114	
X_1^2	1	7.935	7.935	6.068	
X_2^2	1	7.463	7.463	5.707	
X_3^2	1	48.949	48.949	37.432	
Regression	9	127.885	14.209	10.866	3.388
Lack of fit	5	22.740	4.548	3.478	3.687
Pure error	8	10.461	1.308		
Total	22	161.087	7.322		

Eq.3 indicated that the microwave radiation time and liquid to solid ratio were the main factors that influence the yield because of the significance of the terms X_1 and X_2 .The significance of X_1X_2 suggested the obvious interaction between microwave radiation time and liquid to solid ratio. And the significance of all the quadratic terms demonstrated the nonlinear relationship between the three factors and the yield of extraction.

3.5. Analysis of Response Surface

The 3D surface curves were drawn to illustrate the main and interactive effects of the three factors on the yield. The response surfaces are shown in **Figures 6,7,8** with one variable kept at optimum level and the other two varied within the experimental range.

Figure 6 shows the effect of liquid to solid ratio(X2) and extraction time(X1) on the yield. A quadratic effect for both factors on the response can be observed. At a low level of X2 (-2), the system was readily hyperthermal and vaporization of the solvent could reduced the yield with the increasing of microwave time. And at a high level of X2 (2), the yield displayed an increasing curve in the experimental range of X1. It is due to the distinctly interaction between X1 and X2, which was implied in the **Eq.3**. The maximum yield was predicted when X1 was in the range of 1.5 to 1.7 and X2 varied from 0.4 to 0.6.

Figure 7 depicts the effect of enzymatic temperature(X3) and extraction time(X1)，as both them exerting a quadratic effect. As shown in the **Figure 7**, an increase in yield resulted when X3 was increased in the level range from –2 to 0, then the curve started to go down, which may indicate that a level of X1 at approximate 0 is required to achieve maximum yield. Likewise, an increase in yield resulted when X1 was increased in the code range from –2 to 1.5, and then the yield was slightly reduced. In the response surface, X1 exerted a more significant effect on yield than X3, and no obvious interaction between X1 and X3 was observed, which was well in agreement with **Eq.3**.

In **Figure 8**, yield showed quadratic curve depending upon the liquid to solid ratio(X2), whereas no significant effect was observed in the enzymatic temperature(X3). Because the ratio was a key factor which influences the impetus in mass transfer of both enzymatic pretreatment and microwave extraction processes, it exerted a more significant effect on yield than the factor of enzymatic temperature. According to the response surface, there was no obviously interaction between X2 and X3, and it was also supported by the **Eq.3**.

3.6. Optimization of EPME Condition

Yield of extraction was employed as the evaluation objective in the optimization of the parameter to ensure it reaches the peak under constraint conditions. According

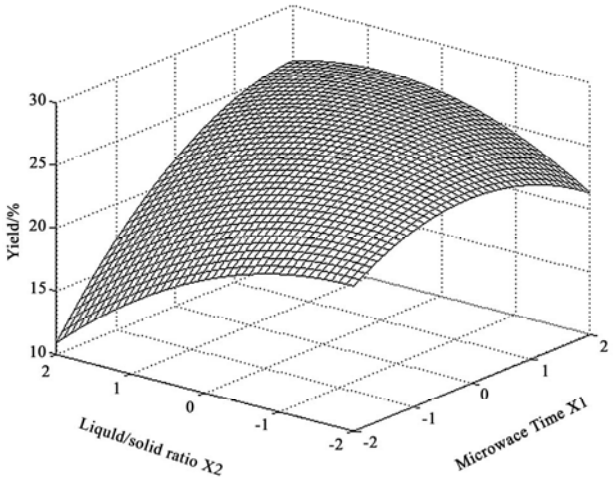


Figure 6. Response surface graph of microwave radiation time and liquid/solid ratio.

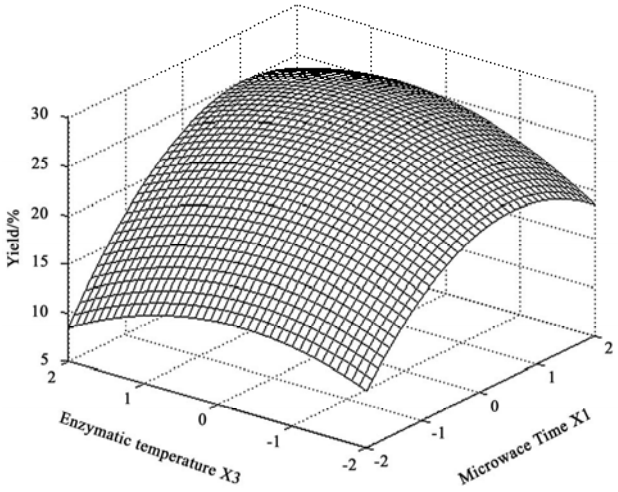


Figure 7. Response surface graph of microwave radiation time and enzymatic temperature.

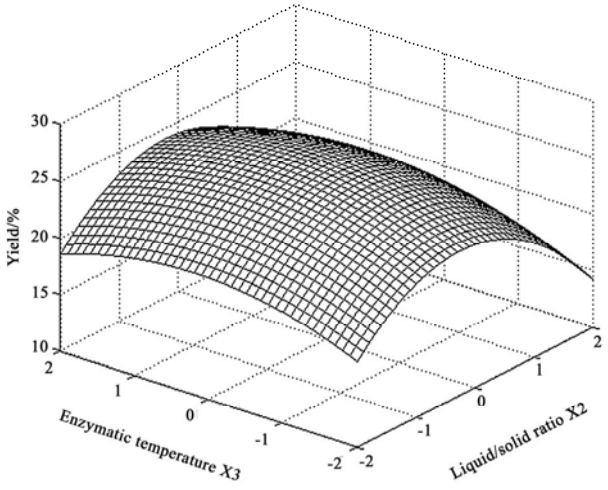


Figure 8. Response surface graph of liquid/solid ratio and enzymatic temperature.

Table 3. Optimum values and verification results

	Coded Value	Calculated Value	Yield/%	
			Esti-mated Value	Measured Value
X_1(s)	1.672	110		
X_2(mL/g)	0.584	36	27.19	27.10
X_3(℃)	0	45		

to (3), model was then built up as described hereinafter [10]:

Objective function: $Y(X_1,X_2,X_3)$;

Constraint conditions: $-1.682 \leq X_i \leq 1.682$; (i=1,2,3)

The optimum parameters and maximal yield were worked out by Newton's iteration method. The results were shown in **Table 3**. To compare the predicted result (82.10%) with the practical value, the rechecking experiment was performed using this deduced optimal condition. The mean value of 27.10% (n=3), obtained from real experiments, demonstrated the validity of the model, since there was no significant differences between 27.19% and 27.10%. The strong correlation between the real and the predicted results confirmed that the extraction model was adequate to reflect the expected optimization.

4. CONCLUSIONS

The quadratic regression orthogonal design was used in this research, an extraction model which can accurately predict the yield of asiaticoside extraction under the experimental condition was established.

Through the analysis of experiment data, it can be found that the microwave radiation time and liquid to solid ratio significantly influence the yield of the extraction. Especially, there was an obvious interaction between the microwave radiation time and liquid to solid ratio.

The optimum combination of the parameters for the extraction of asiaticoside was obtained by the mathematical methology; it was microwave radiation time for 110s; liquid to solid ratio for 36mL/g, and enzymatic pretreatment temperature for 45℃. On this condition, the maximum yield of extraction was 27.10%, closed to the estimated value 27.19%.

The EPME procedure had the advantages of less time, high efficiency of extraction, and environmentally friendly. It can be applied to other extraction of plant materials. But the expansive cost of the enzyme and difficulty in industrialization of microwave extraction would limit the further application of EPME.

5. ACKNOWLEDGMENTS

The work was supported by National Natural Science Foundation of China (No. 20006003).

REFERENCES

[1] Randriamampionona, D., Diallo, B., and Rakotoniriana, F., (2007) Comparative analysis of active constituents in Centella asiatica samples from Madagascar: Application for ex situ conservation and clonal propagation, Fitoterapia, **78**, 482–489.

[2] Verma, K., Bhartariya, G., and Gupta, M., (1999) Reversephase high performance liquid chromatography of asiaticoside in Centella asiatica, Phytochem. Anal., **10**, 191–193.

[3] Sarma, K., Khosa, L., and Chansauria, N., (1995) Antistress activity of Tinospora cordfolia and Centella asiatica extracts, Phytother. Res., **10**, 181−183.

[4] Flora, S. and Gupta, R., (2007) Beneficial effects of Centella asiatica aqueous extract against arsenic-induced oxidative stress and essential metal status in rats, Phytother. Res., **21**, 980−988.

[5] Li, B. B., Smith, B., and Hossain, M., (2006) Extraction of phenolics from citrus peels II, Enzyme-assisted extraction method, Sep. Purif. Technol., **48**, 189−196.

[6] Bayramoglu, B., Sahin, S., and Sumnu, G., (2008) Solventfree microwave extraction of essential oil from oregano, J. Food Eng., **88**, 535−540.

[7] Proestos, C. and Komaitis, M., (2008) Application of microwave-assisted extraction to the fast extraction of plant phenolic compounds, Food Sci. Technol-LEB, **41**, 652−659.

[8] Pan, X. J., Niu, G. G., and Liu, H. Z., (2003) Microwave-assisted extraction of tea polyphenols and tea caffeine from green tea leaves, Chem. Eng. Process, **42**, 129−133.

[9] Lu, X. Y., "Determination of asiaticosides by UV- spectrophotography. (2005) J Modern Food Drug, **16**, 15−16.

[10] Ma, S. F., Wang, L. Q., and Hu, Z. C., (2005) Enzymatic extraction of the submerged mycelium polysaccharide from Pleurotus nebrodensis, T CSAE, **22**, 198−201.

[11] Virot, M., Tomao, V., and Colnagui, G., (2007) New microwave-integrated Soxhlet extraction an advantageous tool for the extraction of lipids from food products, J. Chromatogr. A, **1174**, 138−144.

J. Biomedical Science and Engineering, 2009, 2, 532-537 **JBiSE**
doi: 10.4236/jbise.2009.27077 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Application of quartz crystal nanobalance in conjunction with a net analyte signal based method for simultaneous determination of leucine, isoleucine and valine

Maryam Shojaei[1], Abdolreza Mirmohseni[2], Maryam Farbodi[2]

[1]Department of Natural sciences, Faculty of Animal sciences, University of Tabriz, Tabriz, Iran; [2]Polymer Research Technology Laboratory, Department of Applied Chemistry, Faculty of Chemistry, University of Tabriz, Tabriz, Iran.
Email: mshojaei@tabrizu.ac.ir; mirmohseni@tabrizu.ac.ir; farbodi@tabrizu.ac.ir

Received 22 June 2009; revised 16 July 2009; accepted 19 July 2009.

ABSTRACT

The aim of the present investigation was to develop a biosensor for the detection of amino acids, Leucine, Isoleucine and Valine based on a quartz crystal nanobalance. leucine (Leu), isoleucine (Ile), and valine (Val) were selectively determined by quartz crystal nanobalance (QCN) sensor in conjunction with net analyte signal (NAS)-based method called HLA/GO. An orthogonal design was applied for the formation of calibration and prediction sets including Leu, Ile and Val compounds. The selection of the optimal time range involved the calculation of the net analyte signal regression plot in any considered time window for each test sample. The searching of a region with maximum linearity of NAS regression plot (minimum error indicator) and minimum of PRESS value was carried out by applying a moving window strategy. On the base of obtained results, the differences on the adsorption profiles in the time range between 1 and 300 s were used to determine mixtures of compounds by HLA/GO method. The results showed that the method was successfully applied for the determination of Leu, Ile and Val.

Keywords: Quartz Crystal Nanobalance; Net Analyte Signal; Leucine; Isoleucine; Valine; HLA/GO Method

1. INTRODUCTION

Maple Syrup Urine Disease (MSUD) is a rare autosomal recessive metabolic disorder affecting the metabolism of amino acids, which occurs due to a deficiency of the activity of the mitochondrial enzyme complex branched-chain l-2-ketoacid dehydrogenase (BCKD). The main symptom of MSUD is accumulation of the branched chain amino acids (BCAA) leucine (Leu), isoleucine (Ile) and valine (Val) in blood, urine and cerebrospinal fluid. This deficiency results in mental retardation if not detected soon after birth and l-leucine and/or its keto acid are considered to be the main neurotoxic metabolites in MSUD. So, identification and detection of the abnormal levels of these metabolites in urine samples are necessary to diagnosis and therapy of these pathologies [1].

MSUD screening methods based on gas chromatography-mass spectrometry (GC-MS) and tandem mass spectrometry (MS-MS) have been developed. MS-MS has been reported to be a powerful diagnostic tool in MSUD patients. GC-MS and MS-MS have high resolutions that enable them to be used to measure several amino acids simultaneously. However, these instruments are prohibitively expensive, and hospitals in developing countries cannot afford it [2].

In general, the high-performance liquid chromatography (HPLC) method has been devised to measure amino acid levels, but developed methods have disadvantages, which include complex sample preparation and long analysis time [3].

Mass screening emergency for MSUD in childhood demands to develop simple and inexpensive methods with a rapid and quantitative response. Quartz crystal nanobalance (QCN) is a sensing system based on the sorption of analyte on an adsorbent material [4]. The QCN comprises a thin vibrating AT-cut quartz wafer sandwiched between two metal excitation electrodes. When small amounts of mass are adsorbed at the quartz electrode surface, the frequency of the quartz is changed according to the well-known Sauerbrey equation [5]:

$$\Delta F = -2.26\times10^{-6} F_0^2 (\Delta m / A) \quad (1)$$

where ΔF is the measured frequency shift, F_0 the original oscillation frequency of the dry crystal, Δm the mass change, A the piezoelectrically active area of the excitation electrodes.

Due to some advantages including low cost, portabil-

ity and easy on-line analysis, the quartz crystal nanobalance (QCN) sensor is extensively used for the measurements of mass changes in a variety of chemical and biological studies, such as determination of volatile organic compounds [6,7], poisonous compounds [8] and immunoassay [9,10].

In some cases, the major drawback of the sensors based on QCN is a lack of selectivity since along with the analyte, other compounds usually interfere. In other words, there is no discrimination between the sources of the mass changes. To overcome this shortcoming, one approach is the pattern recognition technique that can be used for the data processing of the QCN signals for the simultaneous determination of mixtures of compounds. Multilinear regression (MLR), partial least squares (PLS) and net analyte signal (NAS) are examples of multivariate analytical techniques [11,12,13].

HLA/GO algorithm (Hybrid Linear Analysis presented by Goicoechea and Olivieri), one of the NAS-based methods, has been successfully used for resolving multicomponent mixtures. Goicoechea and Olivieri [14] have determined tetracycline in blood serum by using synchronous spectrofluorimetry through the HLA/GO algorithm. This algorithm has also been applied for the simultaneous determination of leucovorin and methotrexate, by spectrophotometric [15] and Sorbic (SOR) and benzoic (BEN) acids in fruit juice samples by using spectroscopic signals [16]. HLA/GO has also been applied to the determination of binary mixtures of amoxycillin and clavulanic acid by stopped-flow kinetic analysis [17].

To our knowledge, no study reported for the detection and determination of Leu, Ile and Val using QCN technique. In the present study, we report the simultaneous determination of Leu, Ile and Val in the solution containing some common urine analytes using polystyrene (PS) coated QCN. NAS is utilized to process the frequency data of the crystal at various times, based on different adsorption dynamics of Leu, Ile and Val on the PS coated QCN.

1.1. Theory Notation

An $I\times J$ data matrix R composed of the calibration responses of I samples at J times, a $J\times 1$ vector sk containing the pure adsorption profile of analyte k at unit concentration, and an $I\times 1$ vector ck of calibration concentrations of analyte k are the used matrices and vectors throughout the present work. The net analyte signal (NAS) for analyte k (r^*k) is given by the following equation:

$$r^*k = [I - R_{-k}(R_{-k})^+]r = P_{\mathrm{NAS},kr} \qquad (2)$$

where r is the adsorption profile of a given sample (when r is the profile sk of pure k at unit concentration, **Eq.(2)** becomes $s^*k = P_{\mathrm{NAS},k}s_k$), I is a $J\times J$ unitary matrix, R_{-k} is a $J\times A$ column space spanned by the adsorption profile of all other analytes except k (R_{-k}^+ is the pseudo-inverse of R_{-k} and A is the number of factors used to build the model, and$P_{\mathrm{NAS},k}$ is a $J\times J$ projection matrix which projects a given vector onto the NAS space.

The concentration of component k in an unknown sample is obtained from its adsorption profile (r) as

$$c_k = \frac{s_k^T \Pr}{s_k^T Ps_k} = \frac{s_k^T P \Pr}{s_k^T PPs_k} = \frac{(s_k^*)^T r_k^*}{\left\|s_k^*\right\|^2} \qquad (3)$$

The applied method in this research involves using the mean (uncentred) calibration profile. It is first obtained as

$$\bar{r}_{cal} = \frac{1}{I}\sum_{i=1}^{I} r_{i,cal} \qquad (4)$$

where $r_{i,\mathrm{cal}}$ is the profile for the ith calibration sample. Then the contribution of analyte k is subtracted from the data matrix R in the following way:

$$R_{-k} = R - \frac{c_k \bar{r}_{cal}^T}{\bar{c}_{k,cal}} \qquad (5)$$

where $\bar{c}_{k,cal}$ is the mean (uncentred) calibration concentration of analyte k. The calculation of net sensitivity (s^*_k) is then carried out with the following equation:

$$s_k^* = P_{NAS,k}\left[\frac{\bar{r}_{cal}^T}{\bar{c}_{k,cal}}\right] \qquad (6)$$

1.2. Selection of Time Window

In the present work, the selection of the optimum range of time window was made by calculating an error indicator (EI) as a function of a moving window for each prediction sample, using information of the NASRP (called "net analyte signal regression plot"). NASRP is a plot of the elements of the sample vector $\mathbf{r}^*_k$ versus those of $\mathbf{s}^*_k$ and should fit a straight line through the origin, with random residuals and slope ck. Large and correlated residuals in this plot reveal discrepancies between the measured profile (and thus in $\mathbf{r}^*k$) and the model and, possibly, bias in the estimated concentration. The expression for EI used in the present context is [17]:

$$EI = \frac{\left[s^2\left(1 + \frac{N^2 s^2}{4\left\|r^*\right\|^2}\right)\right]^{1/2}}{\left\|r^*\right\|} \qquad (7)$$

where s is the standard deviation from the best-fitted straight line to the NASRP (in a given adsorption region),

and *N* is the number of points in the latter plot.

2. MATERIALS AND METHODS

2.1. Reagents and Material

All reagents used in this experiment were of analytical grade. leucine (Leu), isoleucine (Ile), valine (Val) were from Sigma chemicals with analytical grade. Polystyrene (PS) was supplied by Tabriz Petrochemical Co., Iran.

2.2. Instrumentation

10 MHz AT-Cut quartz crystals with gold coating on both sides were commercially available from International Crystal Manufacturer (ICM, Oklahoma, USA). For QCN experiments a home made apparatus was used as described in our previous work [18].

2.3. Procedures

A solution casting method was used to coat the polymer over the quartz crystal electrode. Using a Hamilton micro liter syringe (Hamilton BonaduzAG, Switzerland), 4µL of PS/chloroform solution (0.3%, w/v) was dropped on top of the gold electrode of the quartz crystal. A thin layer of PS was obtained after solvent evaporation.

An orthogonal design was applied for the formation of calibration and prediction sets including Leu, Ile and Val. Orthogonal design is used in order to give the most information from the analytical system by using only a few samples. The calibration and prediction sets were prepared according to four-level orthogonal design.

The concentrations varied in the linear range of each compound (50–300 mgL^{-1} for Leu, 100-400 mgL^{-1} for Ile and Val). All solutions were filtered using a syringe filter (0.2 µm) before injecting to the cell. Milli-Q water was used to desorbed analyte and recover the electrode. All measurements were carried out at room temperature (25 °C).

3. RESULTS AND DISCUSSION

3.1. Determination of Pure Leu

The polymer-coated electrode was exposed to a constant concentration of aqueous Leu solution (50 mgL^{-1}). Typical responses for Leu are shown in **Figure 1**. The frequency of the crystal decreased due to the adsorption of analyte to the surface of polymer modified electrode according to **Eq.(1)**. The recorded responses showed that the electrode is sensitive to Leu. The frequency of the crystal was back shifted to its initial value by exposure to the Milli-Q water indicating the full description of analyte from the electrode surface (**Figure 1**).

As the concentration of analyte increased the magnitude of the response increased (**Figure 2**), the calibration

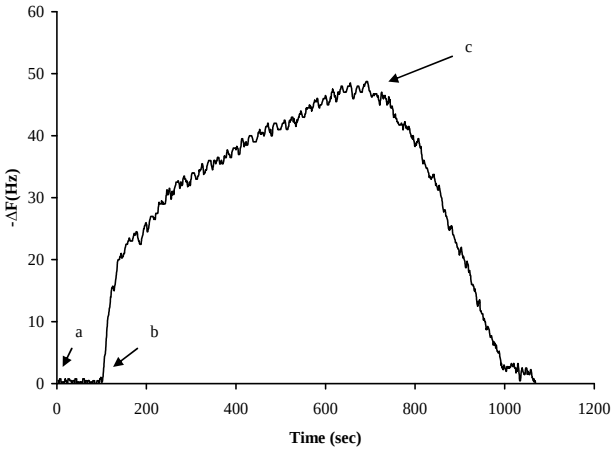


Figure 1. Typical frequency change of PS modified quartz crystal electrode recorded upon exposure to a Leu solution (100 mgL^{-1}). (a) Milli-Q water, (b) Leu solution (100 mgL^{-1}), (c) Milli-Q.

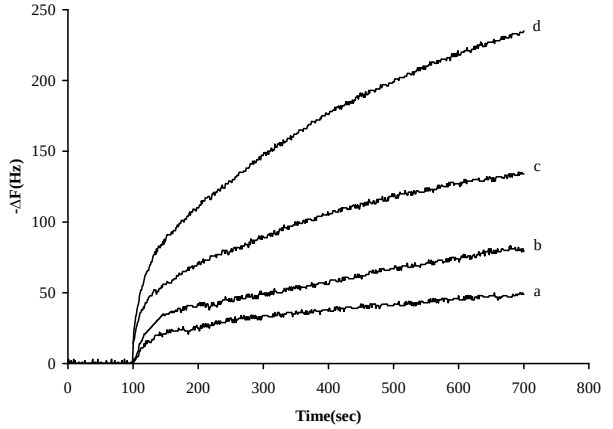


Figure 2. Frequency changes of PS modified quartz crystal electrode as a function of time exposed to various concentration of Leu solutions: (a) 50 mgL^{-1}, (b) 100 mgL^{-1}, (c) 200 mgL^{-1}, (d) 300 mgL^{-1}.

curve was constructed by plotting the frequency shifts against the concentration of Leu (**Figure 3**). The responses were linear against Leu concentrations in the range 50–300 mgL^{-1} and with linear regression coefficient of 0.9806 (n = 4).

3.2. Determination of Leu in the Presence of Ile and Val

Based on the above results, the QCN sensor coated with PS can be employed as pure Leu sensor. Since Leu, Ile and Val are considered as the agents that exist simultaneously in the most urine samples of MSUD patients, it is necessary to investigate the cross-sensitivity between Leu, Ile and Val. So, the frequency shifts were recorded for quartz crystal PS-coated electrode upon exposure Ile and Val. The linear range was obtained 100–400 mgL^{-1} for Ile and Val. The frequency shift obtained for quartz

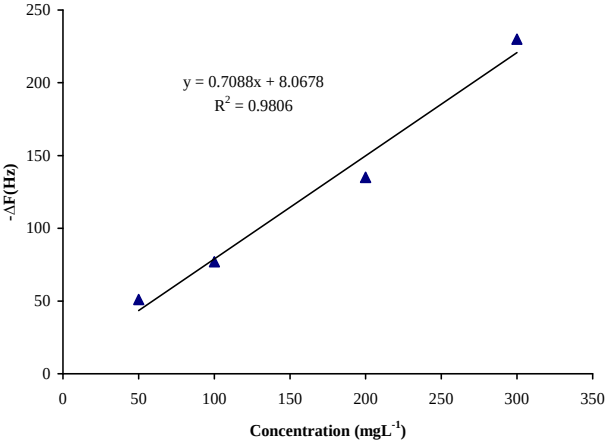


Figure 3. Calibration graph for Leu solutions exposed to PS modified quartz crystal electrode. Exposure time: 10 min.

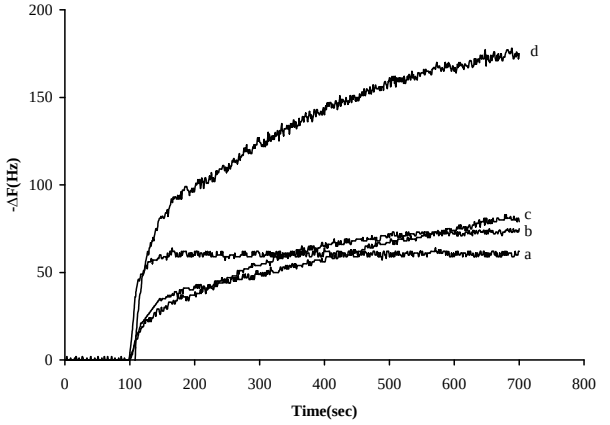


Figure 4. PS modified QCN sensor response upon exposure to: (a) Leu at the concentration of 100 mgL^{-1}, (b) Ile at the concentration of 160 mgL^{-1}, (c) Val at the concentration of 130 mgL^{-1}, (c) mixture of Leu (100 mgL^{-1})/Ile(160 mgL^{-1}) /Val(130 mgL^{-1}).

crystal PS-coated electrode upon exposure to Leu (100 mgL^{-1}), Ile (160 mgL^{-1}), Val (130 mgL^{-1}) and a mixture of Leu (100 mgL^{-1})/Ile (160 mgL^{-1})/Val (130 mgL^{-1}) were recorded (**Figure 4**). The concentrations were selected from the linear range of each compound. The obtained responses showed a significant change in the shape of the frequency–time curves of Leu with including Ile and Val (**Figure 4**). Then, net analyte signal-based HLA/GO method was considered to develop a model for selectively determination of Leu, Ile and Val compounds.

3.3. Optimization of HLA/GO Method

Selection of the optimum number of factors to be used within the HLA/GO algorithms allows one to model the system with the optimum amount of information. In HLA/GO analysis of the calibration set, the PRESS value (predicted error sum of squares) for prediction samples varies as a function of the number of factors. In the present work, cross-validation has been used to select the optimum number of factors for two time intervals, in the range comprised between 1 and 600 s.

The selection of the optimum time region is caused to increase the predictive ability of multivariate analysis by discarding the non-informative parts of adsorption profile from the original data.

The selection of the optimum time region for the analysis was carried out by evaluating the best predicted values for the prediction samples and the minimum error EI values. In this regard, using the optimized number of factors selected in each region, an EI was calculated for each prediction sample, using information of the NASRP.

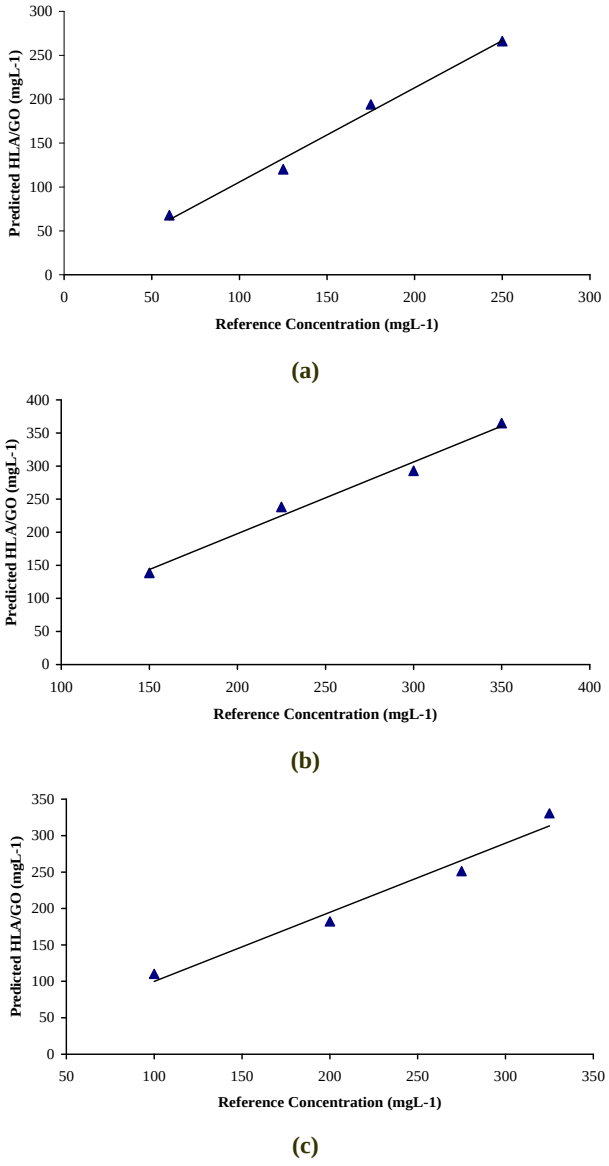


Figure 5. Predicted vs. actual concentrations for HLA/GO calibration models (a) Leu, (b) Ile and (c) Val.

Table 1. Optimization of the sensor range in the prediction of Leu, Ile and Val in the mixture by application of the NAS signal and evaluation of the EI. ([1]actual concentration; [2]predicted concentration)

Sample	Time Range	Factor	EI	Leu[1]	Leu[2]	Error	EI	Ile[1]	Ile[2]	Error	EI	Val[1]	Val[2]	Error
1	1-300	3	0.43	60	67.7	7.7	0.04	325	330.6	5.61	0.57	300	293.1	–6.8
	1-600	5	0.56		43.8	–16.5	0.04		346.1	21.1	0.57		317.1	17.1
2	1-300	3	0.42	125	120.1	–5.1	0.08	275	251.1	–13.86	0.06	350	365.0	15.06
	1-600	5	0.36		146.0	21.0	0.05		142.6	–22.23	0.77		370.9	20.9

The moving window was obtained by varying the time range. **Table 1** shows the ranges of time tested, the optimum number of factors for each region, the EI values calculated, and the predicted values for Leu, Ile and Val. The minimum EI value calculated using information of the NASRP indicates 1-300 s as the most adequate time region for the analysis in this case.

The optimized model was tested in the analysis of the prediction set and plots of c_{pred} versus c_{act} were constructed (**Figure 5**). As it can be seen, the plot showed very good linearity and the values of 0.9894, 0.9707 and 0.9854 were obtained as correlation coefficient for Leu, Ile and Val, respectively.

4. CONCLUSIONS

Leucine (Leu), isoleucine (Ile), valine (Val) were simultaneously determined using adsorption profile data recorded using PS-coated QCN sensor in conjunction with HLA/GO multivariate calibration method. Determination was based on frequency shifts of PS modified quartz crystal electrode due to the adsorption of Leu at the surface of modified electrode in the presence of Ile and Val. The responses were linear against Leu concentrations in the range 50-300 mgL^{-1} and with linear regression coefficient of 0.9806 (n = 4), respectively. The selection of optimum time ranges for each analyte separately were performed by getting the minimum EI, based on the minimization of the PRESS, as a function of a moving adsorption time window. The analysis of the prediction set was used to test the optimized model and plots of c_{pred} versus c_{act} showed very good linearity. The values of 0.9894, 0.9707 and 0.9854 were obtained as correlation coefficient for Leu, Ile and Val, respectively.

5. ACKNOWLEDGMENTS

We are most grateful the financial supports of this research project by the University of Tabriz.

REFERENCES

[1] Barschak, A. G., Marchesan, C., Sitta, A., Deon, M., Giugliani, R., Wajner, M., and Vargas, C. R., (2008) Maple syrup urine disease in treated patients: Biochemical and oxidative stress profiles, Clinical Biochemistry, **41**, 317–324.

[2] Deng, C. and Deng, Y., (2003) Diagnosis of maple syrup urine disease by determination of l-valine, l-isoleucine, l-leucine and l-phenylalanine in neonatal blood spots by gas chromatography-mass spectrometry, Journal of Chromatography B, **792**, 261–268.

[3] Bridi, R., Fontella, F. U., Pulrolnik, V., Braun, C. A., Zorzi, G. K., Coelho, D., Wajner, M., Vargas, C. R., and Dutra-Filho, C. S., (2006) A chemically-induced acute model of maple syrup urine disease in rats for neurochemical studies, Journal of Neuroscience Methods, **155**, 224–230.

[4] Lau, K. T., Micklefield, J., and Slater, J. M., (1998) The optimization of sorption sensor array for use in ambient conditions, Sens. Actuators. B, Chem., **50**, 69–79.

[5] Sauerbrey, G. Z., (1959) The use of quartz oscillators for weighing thin layers and for microweighing, Phys, **155**, 206–222.

[6] Mirmohseni, A. and Oladegaragoze, A., (2004) Determination of chlorinated aliphatic hydrocarbons in air using a polymer coated quartz crystal microbalance sensor Sensors and Actuators B, **102**, 261–270.

[7] Mirmohseni, A. and Oladegaragoze, A., (2003) Determination of Ammonia and Aliphatic amines in Air Using Poly(N-vinylpyrrolidone) Coated Quartz Crystal Microbalance Sensors and Actuators B, **89**, 164–172.

[8] Mirmohseni, A. and Alipour, A., (2002) Construction of a sensor for the determination of cyanide in industrial effluents: A method based on Quartz Crystal Microbalance Sensors and Actuators B, **84**, 245–251.

[9] Mirmohseni, A., Shojaei, M., and Farbodi, M., (2008) Application of a quartz crystal nanobalance to the molecularly imprinted recognition of phenylalanine in solution Biotechnology and bioprocess Engineering, **13**, 592–597.

[10] Shojaei, M., Mirmohseni, A., and Farbodi, M., (2008) Application of a quartz crystal nanobalance and principal component analysis for the detection and determination of histidine, Anal. Bioanal. Chem., **391**, 2875–2880.

[11] Zhu,W., Wei,W., Nie, L., and Yao, S., (1993) Anal. Chim. Acta, **282**, 535–541.

[12] Nyberg, H., (2008) Multivariate analysis applied to a test procedure for determining gun propelling charge weight Part II, Partial least squares analysis, Chemometrics and Intelligent Laboratory Systems, **92**, 118–124.

[13] Mirmohseni, A., Abdollahi, H., and Rostamizadeh, K., (2007) Net analyte signal-based simultaneous determination of ethanol and water by quartz crystal nanobalance sensor, Analytica Chimica Acta, **585**, 179–184.

[14] Goicoechea, H. C. and Olivieri, A. C., (1999) Enhanced synchronous spectrofluorometric determination of tetracycline in blood serum by chemometric analysis. Comparison of partial least-squares and hybrid linear analysis calibrations, Anal. Chem., **71**, 4361–4368.

[15] Espinosa-Mansilla, A., Meras, I. D., Gomez, M. J. R.,

Munoz de la Pena, A., and Salinas, F., (2002) Selection of the wavelength range and spectrophotometric determination of leucovorin and methotrexate in human serum by a net analyte signal based method Talanta, **58**, 255–263.

[16] Marsili, N. R., Sobrero, M. S., and Goicoechea, H. C., (2003) Spectrophotometric determination of sorbic and benzoic acids in fruit juices by a net analyte signal-based method with selection of the wavelength range to avoid non-modelled interferences, Anal. Bioanal. Chem., **376**, 126–133.

[17] Munoz de la Pena, A., Espinosa-Mansilla, A., Acedo Valenzuela, M. I., Goicoechea, H. C., and Olivieri, A. C., (2002) Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis, Anal. Chim. Acta, **463**, 75–88.

[18] Mirmohseni, A., Milani, M., and Hassanzadeh, V., (1999) Ion exchange properties of polypyrrole studied by electrochemical quartz crystal microbalance (EQCM), Polym. Int., **48**, 873–878.

J. Biomedical Science and Engineering, 2009, 2, 538-542 **JBiSE**
doi: 10.4236/jbise.2009.27078 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Design of low-offset low-power CMOS amplifier for biosensor application

Jin-Yong Zhang[1,2], Lei Wang[1], Bin Li[2]

[1]Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; [2]School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China.
Email: jy.zhang@sub.siat.ac.cn; wang.lei@siat.ac.cn

Received 19 June 2009; revised 10 July 2009; accepted 13 July 2009.

ABSTRACT

A compacted and low-offset low-power CMOS amplifier for biosensor application is presented in this paper. It includes a low offset Op-Amp and a high precision current reference. With a novel continuous-time DC offset rejection scheme, the IC achieves lower offset voltage and lower power consumption compared to previous designs. This configuration rejects large DC offset and drift that exist at the skin-electrode interface without the need of external components. The proposed amplifier has been implemented in SMIC 0.18-μm 1P6M CMOS technology, with an active silicon area of 100 μm by 120 μm. The back-annotated simulation results demonstrated the circuit features the systematic offset voltage less than 80 μV, the offset drift about 0.27 μV/℃ for temperature ranging from –30℃ to 100℃ and the total power dissipation consumed as low as 37.8 μW from a 1.8 V single supply. It dedicated to monitor low amplitude biomedical signals recording.

Keywords: Biomedical Integrated Circuit; CMOS Amplifier; Low-Offset and Low-Power; DC Offset Rejection; Biomedical Sensor

1. INTRODUCTION

Recently, there is increasing demand for portable and wearable devices to continuously monitor vital signals such as electroencephalography (EEG) and electrocardiography (ECG), blood pressure, etc. [1]. These devices usually contain various types of biosensors. CMOS amplifier is a crucial block at the front-end of these sensors, because most biomedical signals are characterized by their relative weak amplitude and low frequency, usually of few mV or less and the frequency below 1 kHz [2]. Meanwhile, these signals are often accompanied by large DC offset caused by skin-electrode interface. Therefore, amplifying such weak signals requires an amplifier with low-offset and low-offset drift, which is quite challenging without using any trimmed components.

There are some techniques have been developed to deal with the design challenges. Alternatively, auto-zeroing (AZ), correlated double sampling (CDS) and chopper stabilization techniques (CHS) [3,4,5,6] are utilized in sensor amplifier design to obtain DC offset rejection and high noise performance. However, these circuits have some disadvantages such as the employment of large capacitors, either off-chip or on-chip. Furthermore, these circuits add many CMOS switches that inevitably introduce switching noise, thermal noise, residual nonlinear switch errors and the CHS circuit consumes more power as circuit working in the chopping frequency. In fact these circuits were optimized for low flicker noise at the cost of higher bandwidth and worse thermal noise performance [7]. Trimming amplifier's components is another skill, but performance of this circuit is strongly related to the on-chip components matching and it increases the cost.

As a result, we turned to the CMOS amplifier design using continuous-time technique for high performance and low-cost solution. In this paper, an integrated continuous-time CMOS amplifier with low-offset voltage and low-power consumption was designed to meet the required biosensor. The proposed amplifier, designed in SMIC 0.18-μm CMOS technology, achieved less than 80 μV offset voltage and consumed only 37.8 μW under a 1.8 V supply. It is a good candidate for biosensor application.

2. CIRCUIT IMPLEMENTATION

Amplifier is an important block at the front-end of the biosensor system as in [8]. **Figure 1** shows the architecture of the integrated CMOS amplifier. It consists basically of three blocks, which are current reference, bias generator and low-offset amplifier core. A high precision

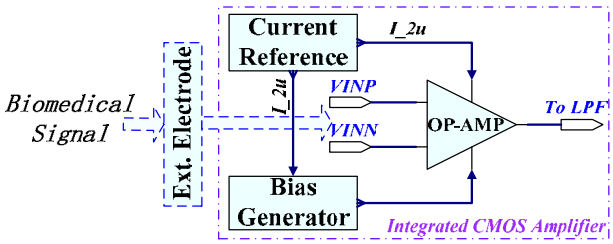


Figure 1. Architecture of the proposed integrated CMOS amplifier.

current reference was integrated in this design, it generated multiple branches of 2 μA temperature and supply independent current and were used to bias the amplifier and the bias generator. The complete schematic diagram of the proposed integrated CMOS amplifier is depicted in **Figure 2**.

2.1. Current Reference/Bias Generator

Minimizing the variation of the reference current and bias voltage for amplifier is crucial as well as achieving high performance in sensor systems. A novel compensation scheme for supply and temperature dependency of MOSFET-only current reference was presented in this design [9]. The complete schematic of the reference is shown in **Figure 2**. It includes start up circuit, self-based current generator, supply and temperature compensation circuit.

The operation principle is that if two current outputs having the same dependency on supply voltage and temperature are subtracted with proper weighting, the compensation output would be obtained. As demonstrated in **Figure 3**, two self-biased current references generated I_{out1} and I_{out2}, respectively.

$$I_{out1,2} = \frac{2}{\mu_p C_{ox} (W/L)_p} \cdot \frac{1}{R_s^2} \cdot (1 - \frac{1}{\sqrt{K_{1,2}}})^2 \quad (1)$$

Two current mirrors are adopted in this circuit to copy the currents I_{out1} and I_{out2} to get I_{m1} and I_{m2}. The size of the transistors and the resistance R_S are determined so the two current outputs I_{m1} and I_{m2} have the same supply dependency and different magnitude. Then, by subtracting I_{m2} from I_{m1}, the supply independent output current I_S could be obtained, but it is still a function of temperature. Through a simple analysis, the supply compensated output current I_S was given by:

$$I_{out1,2} = \frac{2}{\mu_p C_{ox} (W/L)_p} \cdot \frac{1}{R_s^2} \cdot \left[(1 - \frac{1}{\sqrt{K_1}})^2 - (1 - \frac{1}{\sqrt{K_2}})^2 \right] \quad (2)$$

For the negative temperature coefficients of resistor R_S and μ_P, the supply independent current I_S has a proportional-to-absolute-temperature (PTAT) characteristic. The drain current of PMOSFET, I_T, its temperature coefficient is also positive. Then the temperature compensated output current I_{REF} could be obtained by subtracting I_T from I_S. In bias circuit, the master biasing current and voltage of the complete amplifier were derived from the supply and temperature independent current reference.

2.2. Low-Offset Operational Amplifier

Offset in operational amplifier originates in both random and systematic manner [10]. The random offset comes from imperfect fabrication of identical devices. The systematic offset can be considered as errors in the design, it occurs due to the channel length modulation of transistors and the magnitudes of the offset voltages are different according to the input and output common-mode voltages [11]. In this design, a continuous-time asymmetrical differential input structure with active DC offset rejection circuit was implemented to minimize the systematic offset of the amplifier [12,13].

The principle of the active DC offset rejection technique is illustrated in **Figure 4**, considering the amplifier connected as a unity gain following configuration, where the input swing is nearly equal to the output swing. The common-mode level of the input and the output could be detected and amplified by the DC offset rejection circuit, and changed to the feedback signals for current sinks of the amplifier. This is a negative feedback network. By adjusting the current of the current sinks, the input and the output common-mode voltage would be maintained in same level to minimize the systematic offset.

As depicted in **Figure 2**, the circuit of low-offset

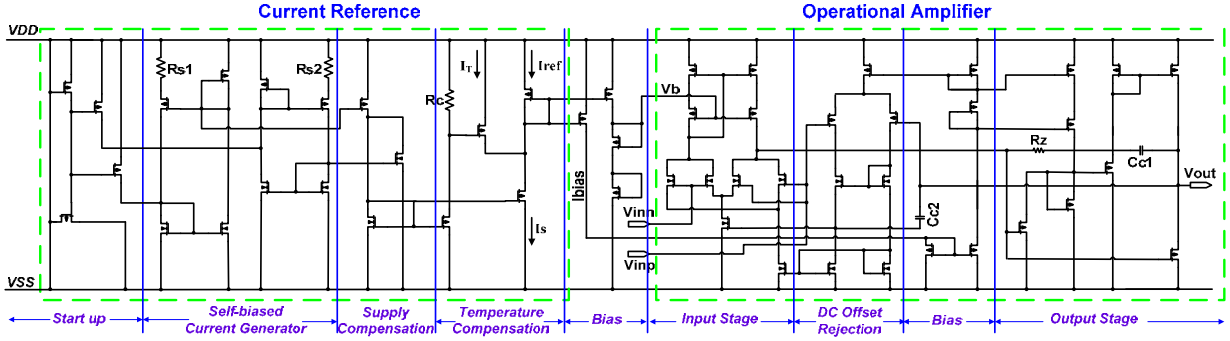


Figure 2. Complete schematic diagram of the integrated CMOS amplifier.

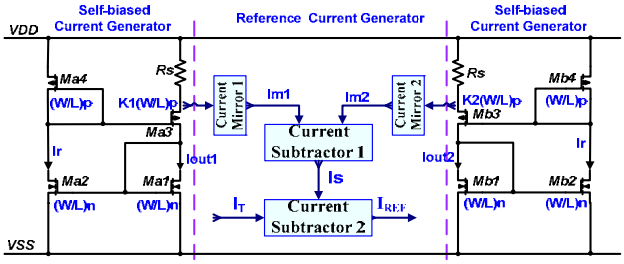


Figure 3. The basic principle of the current reference.

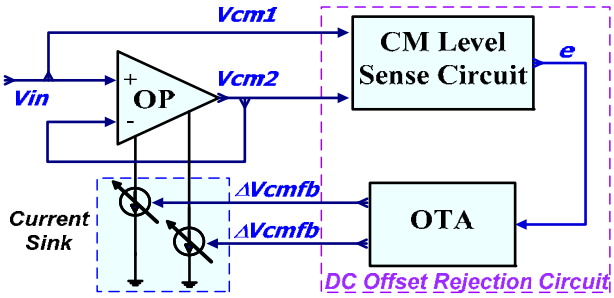


Figure 4. The principle of the DC offset rejection technique.

amplifier is divided into three parts: input stage, DC offset rejection circuit, and output stage. In input stage, the input MOS transistor pairs were designed as asymmetrical differential structure. Besides, the input transistors and active load transistors with appropriate dimensions were used in order to obtain good matching characteristic. In DC offset rejection circuit, a single stage OTA structure was adopted to amplify the difference between input and output common-mode level. Therefore, via cascading the NMOS pairs to obtain the large gain of the OTA. In output stage, the class-AB structure was designed to improve the power efficiency, open-loop gain and driving capability. RC Miller compensation and capacitor compensation techniques were used in this circuit [14].

Finally, a careful layout was planned to reduce process-related random offset: a) the symmetrical layout style was addressed through the entire layout, b) common-centroid cross-coupling layout strategy together with poly guard rings were adopted for critical devices and c) input pairs, active mirror loads and current sources that need to be matched were selectively grouped and arranged with dummies to minimize the effect of spacing-dependent parameter mismatch [15].

3. SIMULATION RESULTS AND DISCUSSIONS

This design has been implemented using the SMIC 0.18-μm CMOS 1P6M technology. **Figure 5** shows the complete layout of the integrated CMOS amplifier, with total silicon area of 100 μm by 120 μm. This chip has sent to be fabricated. We will test it with real-world physiological signals in near future.

3.1. Current Reference

The temperature drifts and supply regulation of the reference current are shown in **Figures 6 (a)** and **(b)**, respectively. The reference offered a current of 2 μA when adjusted to have a zero temperature coefficient at room temperature. It could be observed that an overall temperature coefficient of 0.625nA/°C is obtained between 0°C and 80°C, which corresponds to about 2.2% variation. The response of it is better for temperature from

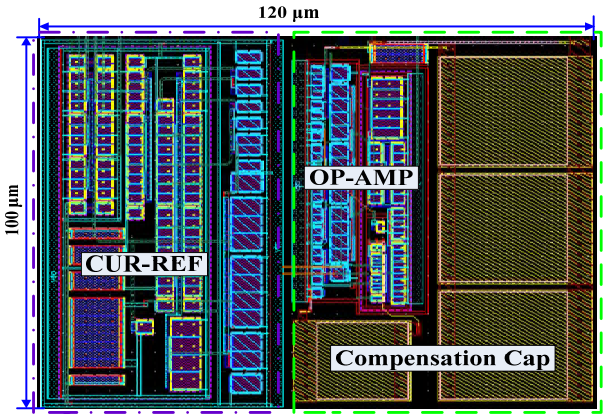
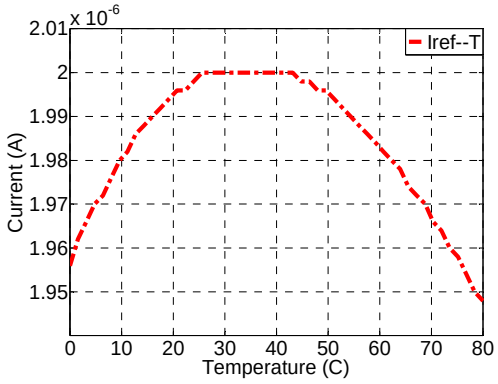
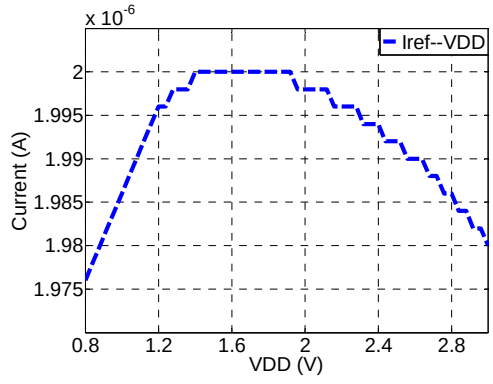


Figure 5. Layout view of the integrated CMOS amplifier.



(a)



(b)

Figure 6. Layout view of the integrated CMOS amplifier Simulated reference current dependence on temperature (a) and supply voltage (b).

20°C to 50°C with a temperature coefficient of 0.13n A/°C. The line regulation of reference current is about 0.55%/V when the supply voltage ranges from 0.8 V to 3 V.

3.2. Low-Offset Operational Amplifier

Figure 7 shows the AC responses of the integrated CMOS amplifier while driving a 3 pF capacitive load. It offered 60 dB open-loop gain, 63.5° phase margin, and 2.82 MHz unity gain bandwidth. DC sweep analysis of the amplifier connected in an inverting unity-gain configuration is shown in **Figure 8**. The simulation results showed good following characteristic between Vin and Vout, and the offset voltage less than 80 μV by averaging. **Figure 9** depicts the offset drifts of the amplifier over a wide temperature range from –30°C to 100°C. The mean offset drift is 0.24 μV/°C, it illustrated the integrated CMOS amplifier was able to sustain low offset voltage over a wide temperature range.

The performance achieved in this design was compared with other state-of-the-art designs for biomedical application. As listed in **Table 1**, it could be seen that our design offered comparable performances. The proposed

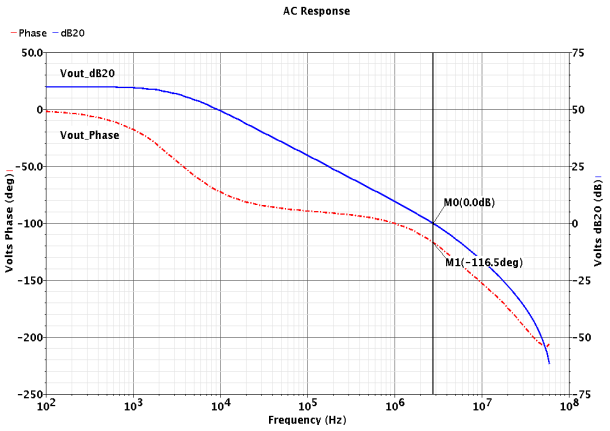
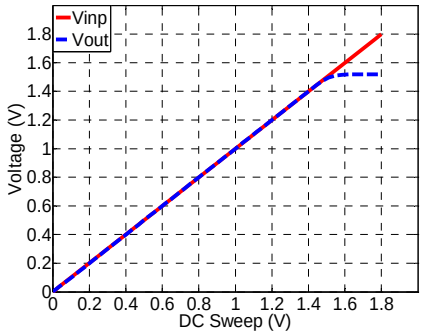
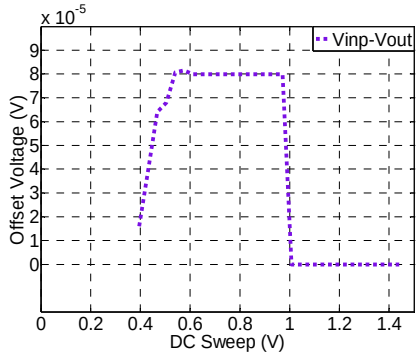


Figure 7. AC simulation results of the integrated CMOS amplifier.



(a)



(b)

Figure 8. The simulation results of the following characteristic (a)and offset tunning range (b).

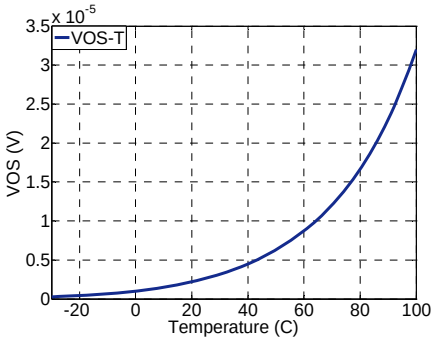


Figure 9. The offset variation with temperature.

Table 1. Performance comparison with other publications. (*instrumentation amplifier)

Parameter	Ref [5] 2002	Ref [16] 2004	Ref [17]* 2005	Ref [18] 2007	This work
CMOS Technology	0.7-μm	1.5-μm	0.35-μm	0.18-μm	0.18-μm
Supply	5 V	3 V	3~4 V	1 V	1.8 V (Typical)
Gain	60.6 dB	39.3 dB	56 dB	14 dB	60 dB
Phase Margin	—	—	—	—	63.5°
Unity Gain Bandwidth	5.5 kHz	2.7 KHz	130 KHz	2.7 KHz	2.82 MHz
CMRR	137 dB	—	100 dB (@60 Hz)	—	>85 dB (@DC-100kHz)
PSRR	—	50 dB	—	50 dB	>100 dB (@DC-10kHz)
Positive Slew Rate	—	0.64 V/μs	50 mV/μs	10 mV/μs	3.45 V/μs
Negative Slew Rate	—	0.64 V/μs	50 mV/μs	10 mV/μs	1.67 V/μs
Input Offset Voltage	88.7 μV	~811 μV	0.3 mV	1.7 mV	80 μV (max)
Offset Drift	—	—	—	—	0.27 μV/℃
Power Dissipation	11 mW	114.8 μW	72.6 μW	3.15 μW	37.8 μW
Core Area	5 mm^2	0.107 mm^2	0.2 mm^2	0.056 mm^2	0.012 mm^2

CMOS amplifier performed technical merits of low-offset voltage, reasonable low-power and with relative small die size, confirming the effectiveness and robustness of the proposed circuit architecture when using both circuit design technique and careful layout technique.

4. CONCLUSIONS

A low-offset low-power and compacted CMOS amplifier with continues-time active DC offset rejection design technique for biosensor applications is presented on-chip in this paper, without the need of trimming. To improve circuitry robustness over power supply and temperature, a high precision current reference was integrated in this design. The whole circuit occupies an area of 100 μm by 120 μm. The back-annotated simulation results suggested that this integrated CMOS amplifier can offer significantly enhanced metrics, in terms of the low-offset less than 80μV, the offset drift about 0.27 μV/℃ for temperature ranging from –30℃ to 100℃, and the total power dissipation approximately 37.8 μW at a single 1.8 V power supply. This integrated CMOS amplifier is particularly useful for a wide range of biosensor applications where a front-end preamplifier is required.

REFERENCES

[1] Yazicioglu, R. F., Merken, P., and Van Hoof, C., (2005) Integrated low-power 24-channel EEG front-end, Electronics Letters, **41(8)**, 457–458.

[2] Webster, J. G., (1998) Medical instrumentation: Application and design, 3rd Edition, Wiley, NewYork.

[3] Chan, C. H., Wills, J., LaCoss, J., Granacki, J. J., and Choma, J., (2007) A novel variable-gain micro-power band-pass auto-zeroing CMOS Amplifier, IEEE International Symposium Circuits and Systems, 337–340,.

[4] Permanod, B., Silawan, N., and Toumazou, C., (2007) Drift reduction in Ion-sensitive FETs using correlated double sampling, IET Electronics Letter, **43(16)**, 857–859.

[5] Panitantum, N. and Wongkomet, N., (2002) A low-offset high-CMRR switched-capacitor instrumentation amplifier for data acquisition applications, Circuits and Systems, **1**, 339–342.

[6] Hanasusanto, G. A. and Zheng, Y. J., (2007) A chopper stabilized pre-amplifier for biomedical signal acquisition, International Symposium on Integrated Circuits, 200–203.

[7] Li, L., (2007) High gain low power operational amplifier design and compensation techniques, A Dissertation of department of Electrical and Computer Engineering, Brigham Young University.

[8] Teo, T. H., Lim, G. K., David, D. S., Tan, K. H., Gopalakrishnan, P. K., kHz-s.l, and Singh, R., (2007) Ultra low-power sensor node for wireless health monitoring system, IEEE International Symposium on Circuits and Systems, 2363–2366.

[9] Yoo, C. A. and Park, J., (2007) CMOS current reference with supply and temperature compensation, Electronics Letters, **43(25)**, 1422–1424.

[10] Chan, P. K., Siek, L., Tay, H. C., and Su, J. H., (2000) A low-offset class-AB CMOS operational amplifier, in Proceedings, The 2000 IEEE International Symposium on Circuits and Systems, **3**, 455–458.

[11] Holzmann, P. J., Wiegerink, R. J., Gierkink, S. L. J., Wassenaar, R. F., and Stroet, P., (1996) A low-offset low-voltage CMOS op amp with rail-to-rail input and output ranges, 1996 IEEE International Symposium on Circuits and Systems, **1**, 179–182.

[12] Lim, K. T., Kim, S. J., and Kwon, O. K., (2003) The OP-amplifier with offset cancellation circuit, Electron Devices and Solid-State Circuits, 2003 IEEE Conference, 445–447.

[13] Gosselin, B., Ayoub, A. E., and Sawan, M., (2006) A low-power bioamplifier with a new active DC rejection scheme, in Proceedings, IEEE International Symposium on Circuits and System, 2185–2188.

[14] Hogervorst, R., Tero, J. P., Eschauzier, R. G. H., and Huijsing, J. H., (1994) A compact power-efficient 3 V CMOS rail-to-rail input/output operational amplifier for VLSI cell libraries, Solid-State Circuits, IEEE Journal, **29(12)**, 1505–1513.

[15] Christopher Saint and Judy Saint, IC Mask Design: Essential Layout Techniques, McGraw-Hill Professional, 2002.

[16] Mohseni, P. and Najafi, K., (2004) A fully integrated neural recording amplifier with DC input stabilization, Biomedical Engineering, IEEE Transactions, **51(5)**, 832–837.

[17] Maryam, S. B., Rakesh, K. L., and Dinesh, K., (2005) A low-power and compact analog CMOS processing chip for portable ECG recorders, Asian Solid-State Circuits Conference, 473–476.

[18] Putra, A., Teo, T. H., and Rajinder, S., (2007) Ultra low-power low-voltage integrated preamplifier using class-AB op-amp for biomedical sensor application, International Symposium on Integrated Circuits, 216–219.

J. Biomedical Science and Engineering, 2009, 2, 543-549 **JBiSE**
doi: 10.4236/jbise.2009.27079 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Retinal vasculature enhancement using independent component analysis

Ahmad Fadzil M. Hani[1]*, Hanung Adi Nugroho[1,2]**

[1]Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, Tronoh, Perak Darul Ridzuan, Malaysia; [2]Department of Electrical Engineering, Univeristas Gadjah Mada, Jl. Grafika 2, Kampus UGM, Jogjakarta, Indonesia.
Email: *fadzmo@petronas.com.my; **hanungadin@gmail.com

Received 22 June 2009; revised 16 July 2009; accepted 24 July 2009.

ABSTRACT

Retinal vasculature is a network of vessels in the retinal layer. In ophthalmology, information of retinal vasculature in analyzing fundus images is important for early detection of diseases related to the retina, e.g. diabetic retinopathy. However, in fundus images the contrast between retinal vasculature and the background is very low. As a result, analyzing or visualizing tiny retinal vasculature is difficult. Therefore, enhancement of retinal vasculature in digital fundus image is important to provide better visualization of retinal blood vessels as well as to increase accuracy of retinal vasculature segmentation. Fluorescein angiogram overcomes this imaging problem but it is an invasive procedure that leads to other physiological problems. In this research work, the low contrast problem of retinal fundus images obtained from fundus camera is addressed. We develop a fundus image model based on probability distribution function of melanin, haemoglobin and macular pigment to represent melanin, retinal vasculature and macular region, respectively. We determine retinal pigments makeup, namely macular pigment, melanin and haemoglobin using independent component analysis. Independent component image due to haemoglobin obtained is used since it exhibits higher contrast retinal vasculature. Contrast of retinal vasculature from independent component image due to haemoglobin is compared to those from other enhancement methods. Results show that this approach outperforms other non-invasive enhancement methods, such as contrast stretching, histogram equalization and CLAHE and can be beneficial for retinal vasculature segmentation. Contrast enhancement factor up to 2.62 for a digital retinal fundus image model is achieved. This improvement in contrast reduces the need of applying contrasting agent on patients.

Keywords: Contrast Enhancement; Independent Component Analysis; Medical Image Processing; Retinal Fundus Image

1. INTRODUCTION

Analyzing retinal fundus image is important for early detection of several diseases related to the retina, e.g. diabetic retinopathy. In diabetic retinopathy, retinal capillary occlusion occurs and accordingly causes enlargement of foveal avascular zone. Foveal avascular zone is the fovea where there is no blood vessels and located in the very centre of macula. Information of retinal vasculature is important to accurately determine the foveal avascular zone. However, digital color fundus images obtained from fundus camera suffer from several problems as can be seen from **Figure 1**. **Figure 1(a)** illustrates the problems of very low contrast and non-uniform illumination which can be seen at the area towards the edge of the image. **Figure 1(b)** shows the occurrence of noise which consists of impulse and Gaussian noises. Detection of the foveal avascular zone is even difficult due to very low image contrast of retinal vasculature against the background in the macular region.

A number of enhancement methods focused in the image spatial domain [2,3,4,5]. Histogram equalization with its modification is commonly used to enhance the image contrast [6]. However, histogram equalization tends to over-enhance the image and results in noisy appearance of the output image. One of the adaptive methods called contrast limited adaptive histogram equalization (CLAHE) worked well on the enhancement of retinal vasculature [7]. Iznita found that the contrast improvement using contrast limited adaptive histogram equalization on an image model ranges between 1.7 and 3 [8]. However, contrast limited adaptive histogram equalization creates artefacts in the enhanced image and the selection of contrast gain limit is image-dependent.

Other related works used the information of color taken from digital color images [9,10]. Colors observed

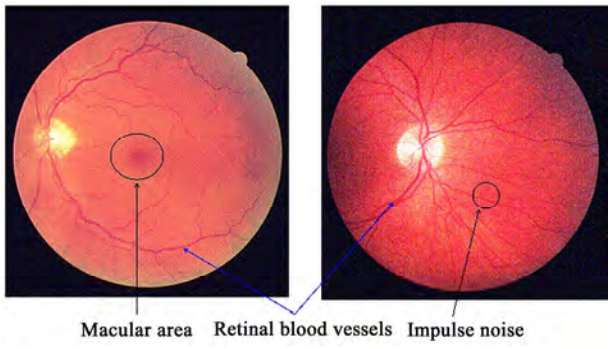


Figure 1. Digital fundus images obtained from fundus camera [1].

in the retinal image correspond to the architecture of retinal layer and the optical properties of the pigments [11,12]. Styles *et al*. developed a model using the concentrations of the five main absorbers found in the fundus layers, namely retinal haemoglobin, choroidal haemoglobin, choroidal melanin, retinal pigment epithetlium melanin and macular pigment [13]. This approach focuses more towards reconstructing rather than improving the contrast. Tsumura *et al.* showed that spatial distributions of melanin and haemoglobin from a skin color image can be separated using independent component analysis [9,14]. Nugroho *et al.* successfully applied a technique based on principal component analysis and independent component analysis to convert the RGB skin image into a skin image that represents skin areas due to melanin and haemoglobin only [10]. The above efforts focus on using independent component analysis to transform digital color image (RGB) into independent components that correspond to the biological makeup of the skin.

The objective of this work is to address the low contrast problem of retinal fundus images obtained from fundus camera when no contrasting agent is injected. A novel approach is presented to enhance the contrast of retinal vasculature by determining the retinal pigments, namely macular pigment, haemoglobin and melanin from fundus images. Distribution of haemoglobin is extracted from a fundus image to reveal retinal vasculature, which is a network of vessels in the retinal layer. Contrast of retinal vasculature obtained using this approach is compared to those from other enhancement methods, such as contrast stretching, histogram equalization and contrast limited adaptive histogram equalization to test the performance of this approach.

2. APPROACH

The approach taken in this research is as follows. First, a model of ocular fundus based on the light interaction is developed to describe the reflectance of the fundus. Second, a model of spectral absorbance of the retinal image is developed to show the components composing the observed colours in a digital fundus image. Third, independent component analysis based on the spectral absorbance of the model is applied to determine retinal pigments from fundus images. Finally, two fundus image models are developed to test performance of the proposed algorithm.

2.1. Ocular Fundus Model

Ocular fundus represents the structure of the back of the eyes that consists of multiple layers of tissue [13]. The ocular fundus image obtained from a fundus camera shows different intensity of reflectance. The reflectance depends on the wavelength, the structure of fundus' layers, the optical properties and quantities of retinal pigments in the ocular fundus. The incident light from a fundus camera can be reflected, absorbed, scattered or transmitted by the retinal tissues.

Generally, the structure of the eye can be classified into two main groups, namely ocular media and ocular fundus [15]. Ocular media consists of cornea, lens and vitreous. It is located between the ocular fundus and the observer. The ocular fundus consists of the retina, the retinal pigment epithelium, the choroid and the sclera. The reflectance of the fundus can be described in the terms of these layers [16]. **Figure 2** depicts a model of ocular fundus showing possible pathways of the reflected light.

2.2. Ocular Fundus Spectral Absorbance Model

The spectral absorbance image provides useful information to identify the absorbance components [14]. In this work, we focus on the distribution of retinal pigments, namely haemoglobin, melanin, and macular pigment, rather than on the fundus layers, to model spectral absorbance of the ocular fundus [17].

Basis of linear combination of the absorption coefficients of melanin, haemoglobin and macular pigment is modelled from three absorbances $\mu_a(\lambda_1)$, $\mu_a(\lambda_2)$ and

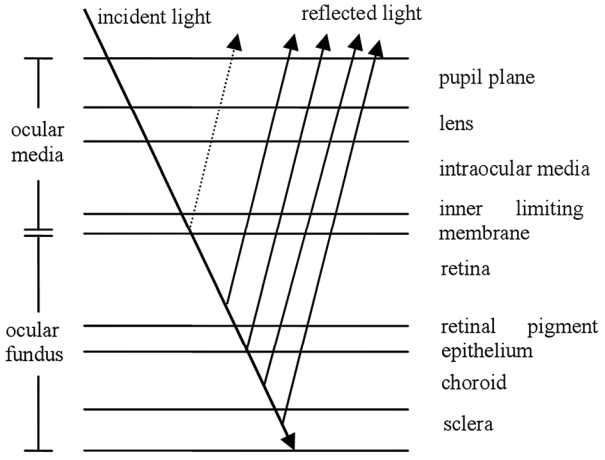


Figure 2. A model of ocular fundus showing pathways of reflected light.

$\mu_a(\lambda_3)$ at three wavelengths λ_1, λ_2 and λ_3. These wavelengths λ_1, λ_2 and λ_3 represent the red (R), green (G) and blue (B) color channels. Fundus spectral absorbance image shows spectral characteristics of the absorbance components in the ocular fundus. Two conditions are assumed when analyzing fundus spectral absorbances. First, the color observed in the fundus image is due to the distributions of melanin, haemoglobin and macular pigment. Second, the quantities of these components are spatially independent of each other. The spectral absorbance in the fundus image represents the linear combination of the absorption coefficients of melanin, haemoglobin and macular pigment.

Let $\boldsymbol{s}_{x,y}$ and $\boldsymbol{v}_{x,y}$ designate a three-dimensional (3-D) quantity vector and composite color vector on an image coordinate (x, y) of a digital color image. A mixing matrix **A** with $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ represents pure color vectors of the three components (haemoglobin, melanin, macular pigment) per unit quantity. It is assumed that a linear combination of mutually independent pure color vectors with the quantities of $s_{1x,y}$, $s_{2x,y}$ and $s_{3x,y}$ result in the composite color vectors of $\boldsymbol{v}_{1x,y}$, $\boldsymbol{v}_{2x,y}$ and $\boldsymbol{v}_{3x,y}$ on the image coordinate (x, y). The following equation illustrates the transformation matrix, where T denotes the transpose.

$$\boldsymbol{v}_{x,y} = \mathbf{A}\ \boldsymbol{s}_{x,y} \tag{1}$$

$$\boldsymbol{s}_{x,y} = [s_{1x,y},\ s_{2x,y},\ s_{3x,y}]^T \tag{2}$$

The pixel value of each channel corresponds to each element of the color vector. **Figure 3** depicts the spectral absorbances of the ocular fundus which consist of pure spectral vectors of melanin, haemoglobin and macular pigment.

2.3. ICA of Fundus Spectral Absorbance Image

Independent component analysis (ICA) is a technique to determine the original signals from mixtures of several independent sources [18,19]. The independent component analysis is modelled as

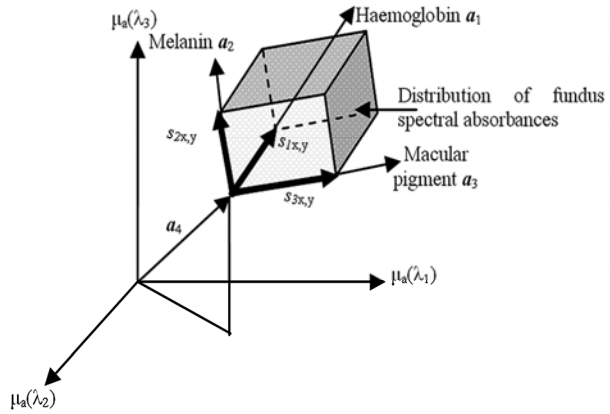


Figure 3. Model of spectral absorbance of the ocular fundus.

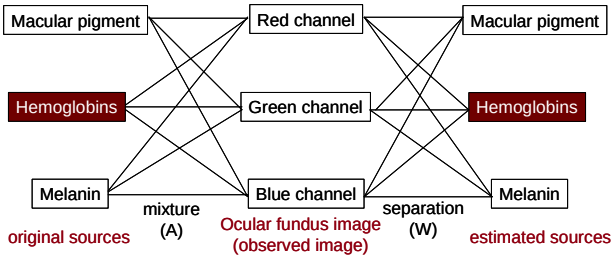


Figure 4. The problem of ICA in ocular fundus image.

$$\boldsymbol{v} = \mathbf{A}\boldsymbol{s} \tag{3}$$

with mixing matrix **A** and random vector $\boldsymbol{v}$, denoting the mixtures $v_1, v_2, \ldots, v_n$. Similarly, $\boldsymbol{s}$ random vector denotes the elements of $s_1, s_2, \ldots, s_n$. This model shows how the observed data $\boldsymbol{v}_n$ is generated by a process of mixing the components $\boldsymbol{s}_i$. The independent components cannot be directly observed and neither can the mixing matrix. Only the random vector $\boldsymbol{v}$ is being observed. Mixing matrix **A** and random vector $\boldsymbol{s}$ are estimated using $\boldsymbol{v}$. Subsequently, separating matrix **W** is used to find the independent component simply by

$$\hat{\boldsymbol{s}} = \mathbf{W}\boldsymbol{v}, \tag{4}$$

with $\hat{\boldsymbol{s}}$ is defined as estimated sources. The objective of independent component analysis is then to get $\hat{\boldsymbol{s}}$ as close as possible to $\boldsymbol{s}$, which is determined as original sources, by determining the optimum separating matrix **W**. Mutually independent components are determined as elements of vector $\boldsymbol{s}$ from the mixture of vectors in the image. A diagram is shown in **Figure 4** to illustrate the idea of using independent component analysis in separating the spatial distributions of melanin, haemoglobin, and macular pigment in the ocular fundus. Three color channels, namely red, green and blue channels, represent random vector $\boldsymbol{v}$ and are used to determine these independent components [17]. By applying the independent component analysis to the composite colour vectors in the image, the relative quantity and pure colour vectors of each independent component are determined with no prior information on the quantity as well as colour vector. In this case, the independent components represent the retinal pigments, i.e. melanin, haemoglobin and macular pigment. The quantities of the retinal pigments are presumed to be mutually independent for the image coordinate. The separating matrix **W** is defined to separate vector $\hat{\boldsymbol{s}}_{x,y}$ using the following equations.

$$\hat{\mathbf{s}}_{x,y} = \mathbf{W}\ \mathbf{v}_{x,y} \tag{5}$$

$$\hat{\mathbf{s}}_{x,y} = [\hat{s}_{1x,y},\ \hat{s}_{2x,y},\ \hat{s}_{3x,y}]^T \tag{6}$$

The extracted independent components $\hat{s}_{1x,y}$, $\hat{s}_{2x,y}$ and $\hat{s}_{3x,y}$ may be similar to $s_{1x,y}$, $s_{2x,y}$ and $s_{3x,y}$, respectively. The composite colour vector $\boldsymbol{v}_{x,y}$ is determined based on the logarithm transformation of the pixel intensities in

the color channels of red, green and blue. Logarithmic transformation is used to transfer reflectance spectra into spectral absorbance since spectral absorbance image provides useful information to identify the absorbance components [14].

$$[\mu_a(\lambda_1),\mu_a(\lambda_2),\mu_a(\lambda_3)]=[-\log(r_{x,y}),-\log(g_{x,y}),-\log(b_{x,y})] \quad (7)$$

here, the values of $r_{x,y}$, $g_{x,y}$ and $b_{x,y}$ correspond to pixel intensity in the color channels of red, green and blue respectively. The composite color vector is denoted as

$$\boldsymbol{v}_{x,y} = [\mu_a(\lambda_1), \mu_a(\lambda_2), \mu_a(\lambda_3)]^T \quad (8)$$

According to the model of spectral absorbance in the ocular fundus from **Figure 3**, the color density vector of the fundus can be stated as

$$\boldsymbol{v}_{x,y} = \mathbf{A}\, \boldsymbol{s}_{x,y} + \boldsymbol{a}_4 \quad (9)$$

where $\mathbf{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$ and $\boldsymbol{s}_{x,y} = [s_{1x,y}, s_{2x,y}, s_{3x,y}]^T$. Elements $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ of the mixing matrix $\mathbf{A}$ represents pure color vectors of the three components (haemoglobin, melanin, macular pigment) per unit quantity. It is assumed that a linear combination of mutually independent pure color vectors with the quantities of $s_{1x,y}$, $s_{2x,y}$ and $s_{3x,y}$ results in the composite color vectors of $\boldsymbol{v}_{1x,y}$, $\boldsymbol{v}_{2x,y}$ and $\boldsymbol{v}_{3x,y}$ on the image coordinate (x, y). Additionally, $\boldsymbol{a}_4$ is similar to noise in the ICA model. In this case, the model is assumed to be noise-free, therefore $\boldsymbol{a}_4$ can be neglected.

Several methods, such as fast fixed-point algorithm (FastICA) [20], joint approximate diagonalization of eigen-matrices (JADE) [21] and information- maximization (infomax) [22] have been proposed to solve the problem of independent component analysis. In ICA, the only assumption needed are: 1) the sources are statistically independent, 2) the probability densities of the sources are non-Gaussian, 3) the mixing of the sources into the observations is linear, and 4) the number of observations is larger than or equal to he number of sources [19]. The FastICA algorithm with symmetrical orthogonalization is used to get the estimated independent components because of its good accuracy and high computational speed for high dimensional data [20].

2.4. Fundus Image Model

A model of fundus image is developed to test the performance of independent component analysis in separating the distribution of macular pigment, hemoglobin and melanin. Mixture of three mutually independent components, i.e. macular pigment, hemoglobin and melanin is used to model a fundus image. As shown in **Table 1**, the statistical intensity description of macular pigment, haemoglobin and melanin in red, green and blue channels are taken from the 44 test images from FINDeRS [23]. A smaller region containing macular area is sampled to get the probability density function of the retinal pigments. In the macular region, retinal capillaries usually show a very low contrast between retinal blood vessels and the background. A clustering method using k-means based on the intensity of red, green, and blue channels of the macular pigment, haemoglobin and melanin is performed to classify the samples due to large value of standard deviation and intensity range of the sample. Based on the experiment, two numbers of clusters are found to be optimal to classify the samples.

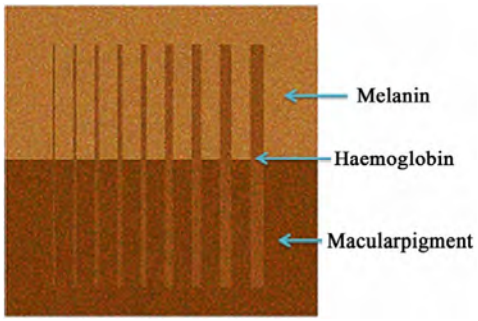
Table 1. Statistical intensity description of macular pigment, haemoglobin and melanin in red (R), green (G) and blue (B) channels.

	Macular pigment	Haemoglobin	Melanin
Mean R	97.14868	120.0417	156.5642
Standard deviation R	28.39299	27.16076	23.98799
Skewness R	0.714194	0.74299	0.07827
Kurtosis R	0.522469	0.259348	0.215941
Minimum R	46.26375	73.32877	102.7158
Maximum R	174.1571	195.1525	219.6025
Mean G	48.29849	62.24773	96.08492
Standard deviation G	13.82747	17.29835	19.95184
Skewness G	0.719616	0.376683	0.620011
Kurtosis G	1.577469	0.236853	1.59036
Minimum G	21.67789	32.06349	57.14166
Maximum G	95.15525	114.2881	164.2989
Mean B	7.971792	17.48195	35.1042
Standard deviation B	5.483867	11.73761	18.85215
Skewness B	1.989686	1.596864	1.31945
Kurtosis B	6.239178	3.978853	2.216477
Minimum B	1.992481	2.412698	12.66622
Maximum B	31.36347	63.40678	100.4873

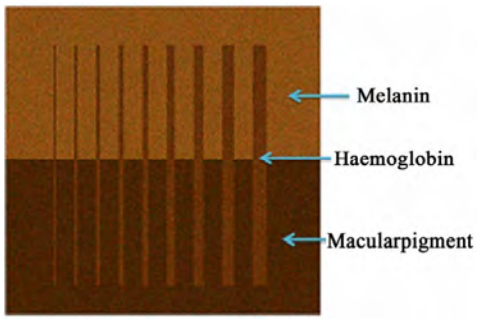
In **Figure 5**, two fundus image models to represent fair and dark fundus images with mixture of specified sample intensity distribution of macular pigment, haemoglobin and melanin in red, green and blue channels are shown. Using these models as the input, the independent component analysis should be able to separate these components into three outputs, namely macular pigment, haemoglobin and melanin.

3. RESULTS AND DISCUSSIONS

A fundus image model is firstly tested using independent component analysis to see performance of the algorithm. The inputs to the FastICA are three separate channels (i.e. red, green and blue channels) of a color fundus image model. As can be seen from **Figure 6**, the proposed algorithm successfully separates the compo-

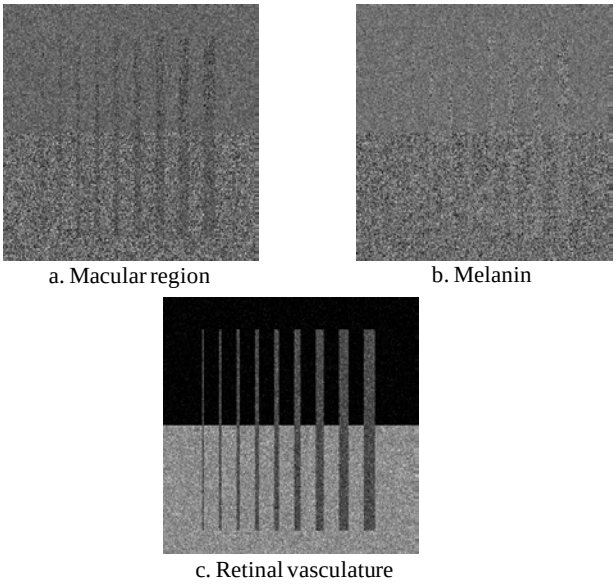


(a). Fair image model



(b). Dark image model

Figure 5. Fundus image model.



a. Macular region b. Melanin

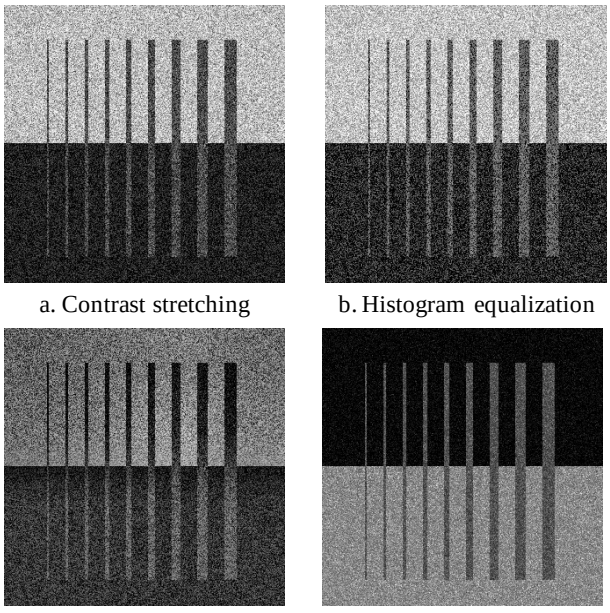
c. Retinal vasculature

Figure 6. Independent component analysis of dark fundus image model.

nents into three, namely macular pigment, haemoglobin and melanin. These three independent components represent macular region, retinal vasculature and melanin, respectively. In **Figure 6(a)**, the brighter area in lower part of the fundus image model is related to the macular region. In **Figure 6(b)**, the melanin is illustrated as the brighter area in upper part of the fundus image model. These two components can be clearly distinguished from the other component, which is related to retinal vasculature, since the retinal vasculature is almost invisible in the appearance of these two components (i.e. macular region and melanin). Furthermore, the retinal vasculature is clearly visualized in **Figure 6(c)**. As a result, independent component image due to haemoglobin obtained exhibits higher contrast retinal vasculature compared to that of the original image.

In this work, 44 retinal fundus images containing macular region are taken from FINDeRS database to model a retinal fundus image. The fundus image model undergoes several enhancement methods, such as contrast stretching, histogram equalization, contrast limited adaptive histogram equalization (CLAHE) to measure contrast improvement factor of these methods and compare to the proposed algorithm. A smaller region containing the macular area is taken to see the enhancement of retinal capillaries, which usually show a very low contrast between retinal blood vessels and the background. **Figure 7** shows green band of dark fundus image model undergoing several enhancement methods, i.e. contrast stretching, histogram equalization and CLAHE. Qualitatively, haemoglobin related ICA shows better enhancement because no artefacts is produced in the process. Nevertheless, the other three enhancement methods tend to increase the noise presence in the image as well as to produce artefacts.

From the fundus image model, the green band image shows the average contrast intensity of 24.60 and 16.80 for fair and dark image model, respectively. Using these values as a reference, the proposed algorithm using ICA



a. Contrast stretching b. Histogram equalization

c. Contrast limited AHE d. ICA (haemoglobin)

Figure 7. Dark fundus image model undergoes several enhancement methods.

with contrast enhancement factor of 1.47 and 2.62 for

fair and dark fundus image models shows a better improvement for both fair and dark fundus image models than that of the other three enhancement methods. Furthermore, as is shown in **Figure 8**, CLAHE with enhancement factor of 1.37 and 1.98 for fair and dark fundus image model, respectively, is still better than that of contrast stretching and histogram equalization. However, compared to that of CLAHE, the proposed algorithm produces no artefacts in the process.

Here an example of retinal image showing macular region is taken to see enhancement of retinal vasculature using the proposed algorithm. In a preliminary work using the above algorithm, it is found that non-uniform illumination in fundus images resulted in false detection of the retinal pigments [17]. This is because the algorithm responds to the spectral reflectance or absorbance of the retinal pigments in the image. Therefore, homomorphic filtering is performed prior to independent component analysis to reduce the problem of non-uniform illumination. Homomorphic filtering is used to reduce illumination which varies slowly in space and at the same time [24].

Figure 9 shows an original color fundus image undergoing homomorphic filtering and its independent components estimated by the FastICA algorithm. The components represent the distribution of the pigments, namely macular pigment, haemoglobin and melanin. The brighter area in the centre of the first independent component (**Figure 9(b)**) represents the distribution of macular pigment. The second independent component (**Figure 9(c)**) shows the distribution of haemoglobin. It is indicated by the enhancement of retinal vasculature. The third independent component (**Figure 9(d)**) shows brighter area related to the distribution of melanin. This result is consistent with the location of melanin, which is fairly distributed in the retinal pigment epithelium and the choroid. Based on the assumption that the image is noise-free, independent component analysis is able to determine the retinal pigments. Moreover, as shown in **Figure 10**, a green band image undergoing CLAHE is compared to the haemoglobin-related component image after the intensity is being inverted to demonstrate that contrast enhancement is also achieved. In this work, CLAHE is also performed on the same images undergoing the proposed algorithm to compare the contrast improvement between these two methods. Having measured the contrast improvement factor on the fundus image model, the proposed algorithm consistently shows better visualization and enhancement compared to that of the CLAHE, which is commonly used as pre-processing for segmentation of retinal vasculature in fundus images. This improvement can be beneficial to improve the accuracy of retinal vasculature segmentation and reduce the need for injecting contrasting agent to the patients.

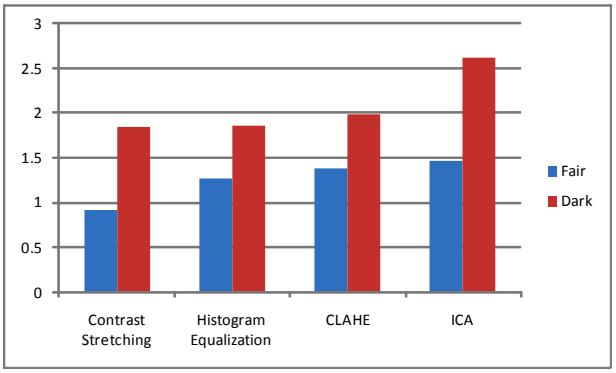


Figure 8. Contrast enhancement factor of retinal vasculature in fundus image model.

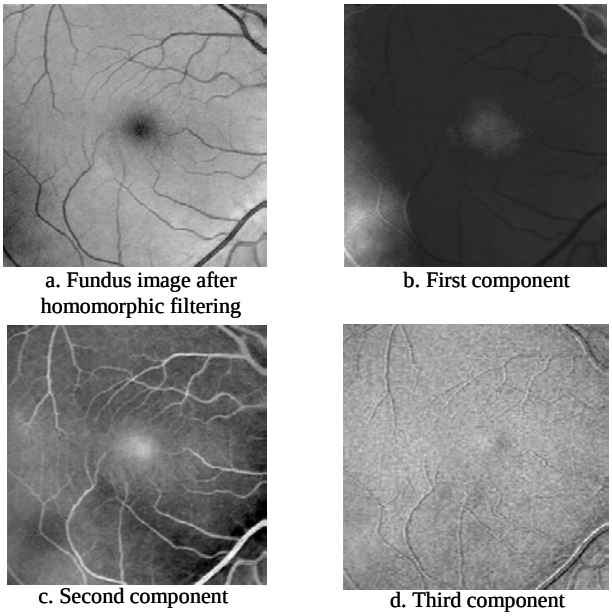


Figure 9. Independent component analysis of a retinal image containing macular region.

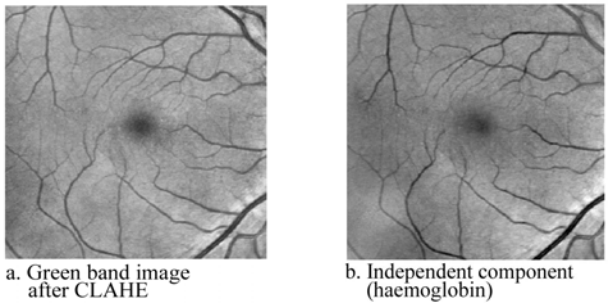


Figure 10. Comparison of contrast enhancement of retinal vasculature between CLAHE and ICA.

4. CONCLUSIONS

Analyzing retinal fundus images is usually difficult as they are of very low contrast. Low contrast between blood vessels and the background makes it difficult to

accurately determine retinal vasculature. Retinal vasculature can be used to determine existence of pathology, macular area and foveal avascular zone. Typical contrast enhancement methods usually create artefacts or introduce noise. Even though fluorescein angiography produces better contrast enhancement, it is not preferable due to its invasive nature of injecting contrasting agent.

In this work, the developed method based on the spectral absorbance model and independent component analysis enables us to determine the retinal pigments, namely haemoglobin, melanin and macular pigment. A fundus image model has been developed to test the performance of the proposed algorithm. As a result, retinal vasculature, macular pigment and melanin distribution can be determined from digital fundus image. Results show that this approach outperforms other non-invasive enhancement methods, such as contrast stretching, histogram equalization and CLAHE and can be beneficial for vessel segmentation. The algorithm produces no artefacts in the process. Using the haemoglobin component, the contrast between retinal blood vessels and the background can be enhanced with contrast enhancement factor up to 2.62 for a model of fundus image. This improvement in contrast reduces the need of applying contrasting agent on patients.

REFERENCES

[1] Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and Ginneken van B., (2004) Ridge-based vessel segmentation in color images of the retina, Medical Imaging, IEEE Transactions, **23**, 501–509.

[2] Niemeijer, M., Staal, J., Ginneken v. B., Loog M., and Abramoff M. D., (2004) Comparative study of retinal vessel segmentation methods on a new publicly available database, Medical Imaging: Image Processing, Milan, 648–656.

[3] Shimahara, T., Okatani, T., and Deguchi, K., (2004) Contrast enhancement of fundus images using regional histograms for medical diagnosis, SICE Annual Conference, **1**, 650–653.

[4] Sinthanayothin, C., Kongbunkiat, V., Phoojaruenchanachai, S., and Singalavanija, A., (2003) Automated screening system for diabetic retinopathy, Image and Signal Processing and Analysis, ISPA, Proceedings of the 3rd International Symposium on, 915.

[5] Wu, D., Zhang, M., Liu, J. C., and Bauman, W., (2006) On the adaptive detection of blood vessels in retinal images, Biomedical Engineering, IEEE Transactions on, **53**, 341–343.

[6] Yang, S., Oh, J. H., and Park, Y., (2003) Contrast enhancement using histogram equalization with bin underflow and bin overflow, in Image Processing, ICIP, Proceedings International Conference on, **1**, I–881–4.

[7] Fadzil, M. H. A., Izhar, L. I., Venkatachalam, P. A., and Karunakar, T. V. N., (2007) Extraction and reconstruction of retinal vasculature, Journal of Medical Engineering Technology, **31**, 435–442.

[8] Iznita, L., (2006) Analysis of retinal vasculature and foveal avascular zone for grading of diabetic retinopathy, in M.Sc, Thesis Electrical and Electronics Engineering Programme Bandar Seri Iskandar, Malaysia, Universiti Teknologi PETRONAS.

[9] Tsumura, N., Haneishi, H., and Miyake, Y., (1999) Independent-component analysis of skin color image, J. Opt. Soc. Am. A, **16**, 2169.

[10] Nugroho, H., Fadzil, M. H. A., Yap, V. V., Norashikin, S., and Suraiya, H. H., (2007) Determination of Skin Repigmentation Progression, Engineering in Medicine and Biology Society, EMBS Annual International Conference of the IEEE, 3442.

[11] Preece, S. J. and Claridge, E., (2002) Physics Based Medical Image Understanding of the Colouration of the Ocular Fundus with Application to Detection of Diabetic Retinopathy, Department of Computer Science, University of Birmingham, Edgbaston, Birmingham Technical Report CSR–04–08.

[12] Orihuela-Espina, F., Claridge, E., and Preece, S. J., (2003) Histological parametric maps of the human ocular fundus: preliminary results, Medical Image Understanding and Analysis, 133–136.

[13] Styles, I. B., Calcagni, A., Claridge, E., Orihuela-Espina, F., and Gibson, J. M., (2006) Quantitative analysis of multi-spectral fundus images, Medical Image Analysis: Special Issue on Functional Imaging and Modelling of the Heart (FIMH 2005), **10**, 578–597.

[14] Tsumura, N., Haneishi, H., and Miyake, Y., (2000) Independent Component Analysis of Spectral Absorbance Image in Human Skin, Optical Review, **7**, 479–482.

[15] Knighton, R. W., (1995) Quantitative reflectometry of the ocular fundus, Engineering in Medicine and Biology Magazine, IEEE, **14**, 43–51.

[16] Berendschot, T. T. J. M., DeLint, P. J., and Norren, D., (2003) Fundus reflectance-historical and present ideas, Progress in Retinal and Eye Research, **22**, 171–200.

[17] Ahmad Fadzil, M. H., Nugroho, H. A., Venkatachalam, P. A., Nugroho, H., and Izhar, L. I., (2008) Determination of Retinal Pigments from Fundus Images using Independent Component Analysis, in Biomed 4th Kuala Lumpur International Conference on Biomedical Engineering, Kuala Lumpur, 555–558.

[18] Comon, P., (1994) Independent component analysis, A new concept?, Signal Processing, **36**, 287.

[19] Hyvarinen, A. and Oja, E., (2000) Independent component analysis: algorithms and applications, Neural Networks, **13**, 411.

[20] Hyvarinen, E., (1999) Fast and robust fixed-point algorithms for independent component analysis, Neural Networks, IEEE Transactions on, **10**, 626–634.

[21] Cardoso, J. F. and Souloumiac, A., (1999) Blind beamforming for non Gaussian signals, IEE Proceedings-F, **140**, 362–370.

[22] Bell, A. J. and Sejnowski, T. J., (1995) An information-maximisation approach to blind separation and blind deconvolution, Neural Computation, **7**, 1129–1159.

[23] Ahmad Fadzil, M. H., Nugroho, H. A., Nugroho, H., and Izhar, L. I., (2009) Fundus Image Database for Non Invasive Diabetic Retinopathy Monitoring and Grading System (FINDeRS), Universiti Teknologi PETRONAS.

[24] Adelmann, H. G., (1998) Butterworth equations for homomorphic filtering of images, Computers in Biology and Medicine, **28**, 169–181.

J. Biomedical Science and Engineering, 2009, 2, 550-558 **JBiSE**
doi: 10.4236/jbise.2009.27080 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Analysis of correlated mutations, stalk motifs, and phylogenetic relationship of the 2009 influenza A virus neuraminidase sequences

Wei Hu

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

Received 13 October 2009; revised 27 October 2009; accepted 30 October 2009.

ABSTRACT

The 2009 H1N1 influenza pandemic has attracted worldwide attention. The new virus first emerged in Mexico in April, 2009 was identified as a unique combination of a triple-reassortant swine influenza A virus, composed of genetic information from pigs, humans, birds, and a Eurasian swine influenza virus. Several recent studies on the 2009 H1N1 virus utilized small datasets to conduct analysis. With new sequences available up to date, we were able to extend the previous research in three areas. The first was finding two networks of co-mutations that may potentially affect the current flu-drug binding sites on neuraminidase (NA), one of the two surface proteins of flu virus. The second was discovering a special stalk motif, which was dominant in the H5N1 strains in the past, in the 2009 H1N1 strains for the first time. Due to the high virulence of this motif, the second finding is significant in our current research on 2009 H1N1. The third was updating the phylogenetic analysis of current NA sequences of 2009 H1N1 and H5N1, which demonstrated that, in clear contrast to previous findings, the N1 sequences in 2009 are diverse enough to cover different major branches of the phylogenetic tree of those in previous years. As the novel influenza A H1N1 virus continues to spread globally, our results highlighted the importance of performing timely analysis on the 2009 H1N1 virus.

Keywords: Entropy; Co-mutation; Mutation; Mutual Information; Neuraminidase; Phylogenetic Analysis; Random Forest; Stalk Motif; Swine Flu

1. INTRODUCTION

There are three types of flu viruses, types A, B, and C. Type A viruses are the most pathogenic to humans. The influenza has on its surface two glycoproteins, hemagglutinin (HA) and neuraminidase (NA), based on which influenza is classified. There are 16 types of HA proteins and 9 types of NA proteins, which are named H1, H2, H3 and etc. For example, “bird flu” is H5N1 and “swine flu” is H1N1. The HA binds the virus to sialic acid receptors on the host cell surface. The NA protein facilitates the release of virions to infect other cells by removing sialic acid residues from the viral HA during entry and release from cells. The NA protein is a tetramer of four identical polypeptide chains anchored in the membrane of the virus. Its head domain is globular and supported by a long and thin stalk.

The “Spanish” influenza pandemic of 1918–1919 caused about 50 million deaths worldwide and about one-third of the world’s population was infected. One unique feature of the 1918 influenza pandemic was the simultaneous infection of humans and swine. Recent studies on the 1918 virus revealed that the genes encoding the HA and NA surface proteins of the 1918 virus were derived from an avian-like influenza virus shortly before the start of the pandemic [1].

In April 2009, a novel strain of the influenza A (H1N1) virus was discovered in patients from Mexico and the United States and it spread across the globe via human-to-human contact within a very short time. Because of the seriousness of this new flu virus, the World Health Organization (WHO) has officially declared the H1N1 virus a global pandemic. Genomic analysis of the 2009 influenza A (H1N1) virus suggested that it is closely related to common reassortant swine influenza A viruses isolated in North America, Europe, and Asia. Its NA sequences have 94.4% similarity at the nucleotide level with European swine influenza A virus strains from 1992 [2].

Flu drugs such as oseltamivir (Tamiflu®) and zanamivir (Relenza®) currently in use only target the NA proteins, and disrupt the capability of the virus to escape infected cells and move elsewhere to infect other healthy cells. Clinical reports suggested that the new virus is susceptible to the two drugs [3]. However, a growing concern is that more drug resistant mutants will emerge under the selection pressure of constant drug use. Re-

searchers from Rensselaer Polytechnic Institute [4] designed a new flu drug by targeting both the HA and NA genes of the virus as an effective way to treat the next mutation of H1N1 swine flu.

Two recent studies [5,6] provided insights into the interactions of flu drugs with NA of the 2009 H1N1 virus. One study [6] developed a 3D structure of 2009 H1N1 NA and compared it with the crystal structure of 2006 H5N1 NA and the structure of 1918 H1N1 NA. It found that the hydrophobic Try347 in H5N1 NA does not match with the hydrophilic carboxyl group of oseltamivir as in the case of H1N1 NA, which explains in part the reason why the H5N1 avian influenza virus is drug-resistant to oseltamivir.

Another study [5] found that the NA sequences of 2009 H1N1 are phylogenetically more closely related to European H1N1 swine flu and H5N1 avian flu rather than to the H1N1 counterparts in the America. It also investigated the sequence variations of 2009 H1N1 NA, using three sequences of NA, A/H1N1 /California/04/2009, A/H5N1/Vietnam/2004, and A/H1N1/Brevig Mission/1/18 (the 1918 Spanish flu). With multiple sequence alignment, they found that among the 387 residues of the NA domain, the 2009 H1N1 NA differs from the other two strains in 21 positions. The novel mutations of NA are mainly located at the protein surface and not near the binding pocket for currently used NA inhibitors. It is natural to explore whether there are any potential mutation sites that may interfere with the active sites in the near future.

In light of the possible emerging of new mutations of 2009 H1N1 that could lead to serious drug resistance, it is imperative to study the potential mutations and co-mutations of the current sequences of 2009 H1N1, because mutations tend to function in concert to achieve some biological purposes. In this study, we employed entropy and mutual information theory [7,8] to investigate this issue.

Because the NA stalk supports the head domain, its length can influence the function of NA. A special NA stalk motif with a 20-amino acid deletion in the 49th to 68th positions of the stalk region was first identified in H5N1 in 2000. There was a gradual increase of this special NA stalk motif in H5N1 isolates from 2000 to 2007, and it was in all 173 H5N1 human isolates from 2004 to 2007. The H5N1 virus carrying this special NA stalk motif has the highest virulence and pathogenicity in chicken and mice [9]. This finding prompted us to search for similar stalk motifs in the current 2009 H1N1 virus strains.

In summary, the goal of our study is to conduct a timely analysis of mutations, co-mutations, stalk motifs, and phylogenetic relationship of the 2009 H1N1 NA sequences available up to date. Such information can be valuable in further efforts to improve drug design and flu treatment.

2. MATERIALS AND METHODS

2.1. Sequence Data

Published NA sequences of 7251 influenza A virus were downloaded from the Influenza Virus Resource (http://www.ncbi/nlm.nih.giv/genomes/FLU/FLU.html) of the National Center for Biotechnology Information (NCBI) on Sept 13, 2009. We were mainly interested in the sequences in 2009, but also needed the sequences in 2008 and 2007 to provide comparison in the study of stalk motifs. There were 283 sequences of H1N1 and H5N1 and 52 of H3N2 in 2009. All the sequences used in the study were aligned with MAFFT [10].

2.2. Entropy and Mutual Information

In information theory [7,8], entropy is a measure of the uncertainty associated with a random variable. Let x be a discrete random variable that has a set of possible values $\{a_1, a_2, a_3,\ldots,a_n\}$ with probabilities $\{p_1, p_2, p_3,\ldots,p_n\}$ where $p(x=a_i)= p_i$. The entropy H of x is

$$H(x) = -\sum_i p_i \log p_i$$

The mutual information of two random variables is a quantity that measures the mutual dependence of the two variables or the average amount of information that x conveys about y, which can be defined as:

$$I(x, y) = H(x) + H(y) - H(x, y)$$

where $H(x)$ is the entropy of x, and $H(x,y)$ is the joint entropy of x and y. $I(x,y)=0$ if and only if x and y are independent random variables.

In current study, each of the n columns in a multiple sequence alignment of a set of NA sequences of N residues is considered as a discrete random variable x_i ($1 \le i \le N$) that takes on one of the 20 (n=20) amino acid types with some probability. $H(x_i)$ has its minimum value 0 if all the residues at position i are the same, and achieves its maximum if all the 20 amino acid types appear with equal probability at position i, which can be verified by the Lagrange multiplier technique. A position of high entropy means that the amino acids are often varied at this position. While $H(x_i)$ measures the genetic diversity at position i in our current study, $I(x_i, y_j)$ measures the correlation between residue substitutions at positions i and j.

Entropy and mutual information were applied to sequence analysis extensively. Mutual information was employed to identify groups of covariant mutation positions in the sequences of HIV-1 protease and to distinguish the correlated residue substitutions resulting from neutral mutations and those induced by multi-drug resistance [11]. Based on entropy a simple informational index was proposed in [12] to characterize the patterns of synonymous codon usage bias. In another

study, sequence data of 1032 complete genomes of influenza A virus (H3N2) during 1968-2006 were used to construct networks of genomic co-occurrence to describe H3N2 virus evolutionary patterns and dynamics. It suggested that amino acid substitutions corresponding to nucleotide co-changes cluster preferentially in known antigenic regions of HA [13]. Further, mutual information was used to construct site transition network based on 4064 HA1 of A/H3N1 sequences from 1968 to 2008, which was able to model the evolutionary path of the influenza virus and to predict seven possible HA mutations for the next antigenic drift in the 2009–2010 season [14]. Recently, entropy and mutual information were also applied to indentify critical positions and co-mutated positions on HA for predicting the antigenic variants [15].

2.3. Mutual Information Evaluation

In order to assess the significance of our mutual information values of residue pairs of NA, it is necessary to show that these values are significantly higher than those based on random sequences. For each residue position of NA, we randomly permuted the amino acids from different sequences at that position and calculated the mutual information of these random sequences. This procedure was repeated 1000 times. The P value was calculated as the percentage of the mutual information values of the permuted sequences that were higher than those of the sequences of NA.

2.4. Random Forest Clustering

Random Forest, proposed by Leo Breiman in 1999 [16], is an ensemble classifier based on many decision trees. The structure of a single tree could be easily altered by a small perturbation of data. Random Forest overcomes this problem by averaging across different decision trees. For many data sets, Random Forest produces a highly accurate classifier for supervised learning, comparable to Support Vector Machine, the state of the art machine-learning algorithm. It computes proximities between cases and this technique can be extended to unlabeled data, leading to unsupervised clustering. In [17] random forest clustering was applied to renal cell carcinoma.

To view the clusters formed by Random Forest, multidimensional scaling [18] was utilized to project high-dimensional data down into a low-dimensional space while preserving the distances between them. First the proximities between cases i and j form a symmetric and positive definite matrix {prox(i,j)}. Then a second positive definite and symmetric matrix {cv(i,j)} is constructed using the entries of {prox(i,j)}. Random Forest extracts a few largest eigenvalues of the cv matrix and their corresponding eigenvectors. The values of $\sqrt{e(i)}v(i)$ are referred to as the ith scaling coordinate, where e(i) and v(i) re the ith eigenvalue and eigenvector of matrix cv [19]. In this study, the first and second scaling coordinates were utilized to visualize the data.

2.5. Important Sites in NA

The N1 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 [12]. According to the numbering in [5], these residue positions 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, and 448–468.

3. RESULTS

3.1. Mutations and Co-mutations

The NA molecule is a homotetramer consisting of four identical polypeptide chains, each of about 470 amino acids. The exact number varies depending on the strain of the virus. The enzymatic domain of the NA is supported from the virus envelope by a polypeptide stalk of variable length. The major molecular determinants that are known to influence the functional activities of the NA protein are the enzyme active site, the stalk length, the sialic acid binding site, and potential glycosylation sites.

In this study, entropy analysis was applied to locate the positions that have elevated likelihood to develop mutations and those that already had mutated. The top 31 positions with high entropy are displayed in **Figure 1**. All except four, 248, 339, 340, and 454, were mutational positions discovered in [5]. These exceptional positions are of interest because they have the potential to mutate and position 248 is close to one active site 247. Notably, there were two clusters of mutations, one near position 286 and another near position 386 (**Figure 1**).

For the sake of overview and comparison, we also plotted in **Figure 1** the distinct entropy distributions of the N1 and N2 sequences deposited at the NCBI web site during 2009 so far. The sequences of N1 and N2 varied the most in the neighborhood of the stalk region between positions 36 and 76. In addition, the N1 sequences in H1N1 and H5N1 are experiencing a much greater genetic change than the N2 sequences in H3N2 this year as illustrated by their entropy (**Figure 1**).

Next, we seek to probe the potential mutations that may affect the drug binding sites from a greater distance. We calculated the mutual information of each possible residue pairs from 469 residues of NA, a total of 109746 pairs. The top 44 pairs (top 0.04% of all pairs) were selected, all with a P value of zero. Because we were interested in the mutations in the NA domain, only those pairs in that region were chosen, which gave us 11 pairs:

(149, 263), (149, 321), (263, 321), (228, 321), (188, 365), (189, 369), (221, 369), (189, 386), (149, 389), (263, 389), and (321, 389). Of these 11 pairs, two networks of co-mutations were uncovered (**Figure 2**), which were mapped to the homology-based 3D structure of N1 built in [5] (**Figures 3** and **4**). These two networks of co-mutations may form interaction chains to connect distant residues to the active sites.

The first network, consisting of positions 149, 263, 321, and 389, has a remarkable property that any one of them is highly correlated to all the other three. Position 149 is near the active site and the bound drug, therefore is of great importance and it is a part of the 150 loop region including positions 147, 148, 150, and 151. Calcium ions are important for the thermo stability and enzyme activity of influenza virus NAs. Three potential metal binding sites in each monomer of the tetramer were observed. The two mutation positions, 321 and 389, are located in the region of one such site at position 470 [20] (**Figure 4**).

The second network has positions 188, 189, 221, 365, and 369. Position 221 is near the three active sites 223, 225, and 228. Positions 365 and 369 are close to the active site 368 and the bound drug; therefore positions 188 and 189 may function together with 365 and 369 to influence the active site 368 and the bound drug (**Figures 5** and **6**). Position 221 is not a mutation site in the three-sequence alignment in [5], but it has high entropy.

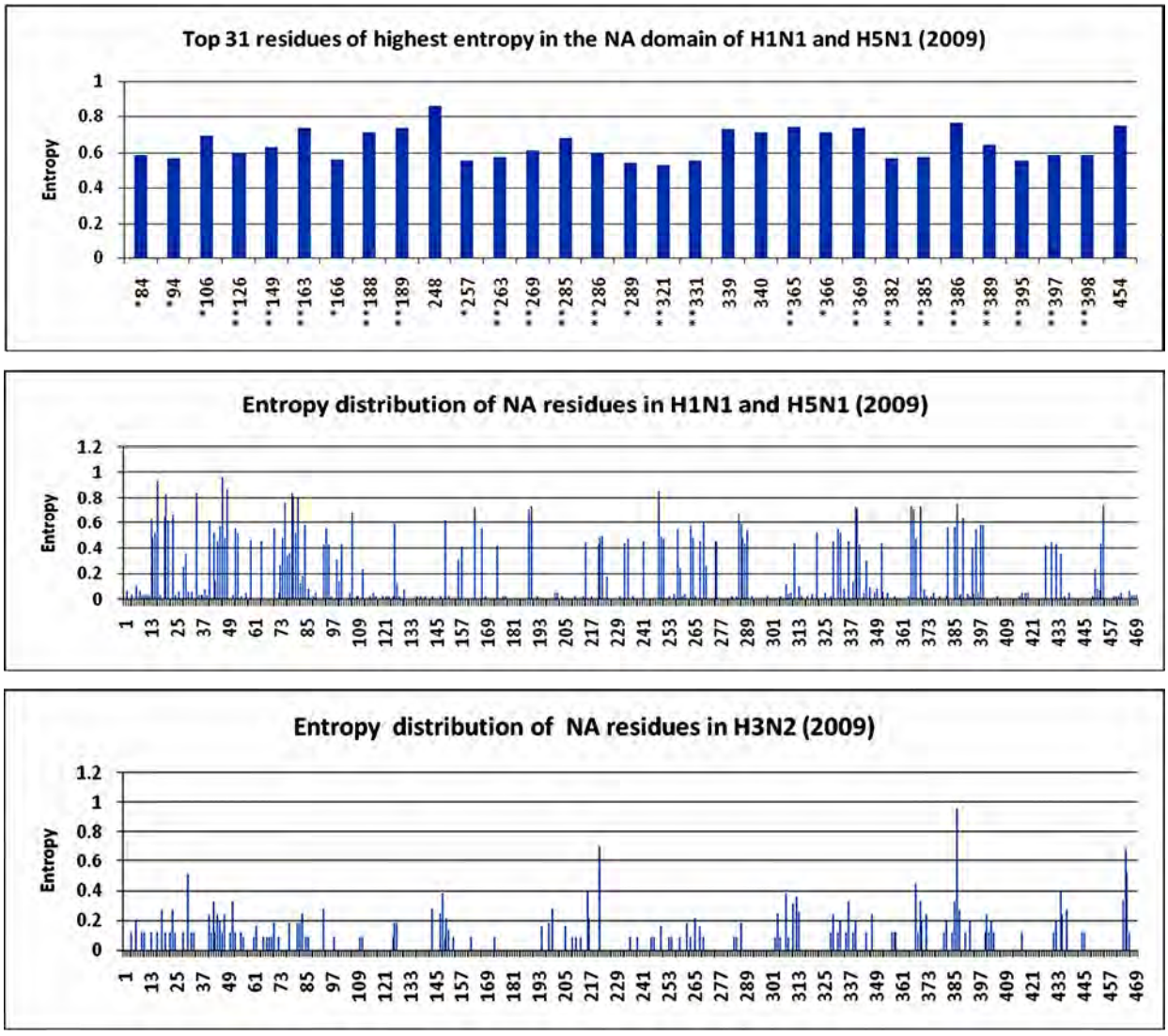


Figure 1. The top plot shows the top 31 residues of highest entropy in the NA domain (83–469) of H1N1 and H5N1 (2009). The residues that had one different amino acid than the two reference strains in [5] are marked with one asterisk, and those that had two different amino acids are marked with two asterisks. The middle and the bottom plots show the entropy of all residues in NA (1 – 469) of H1N1 and H5N1 (2009) and H3N2 (2009) respectively.

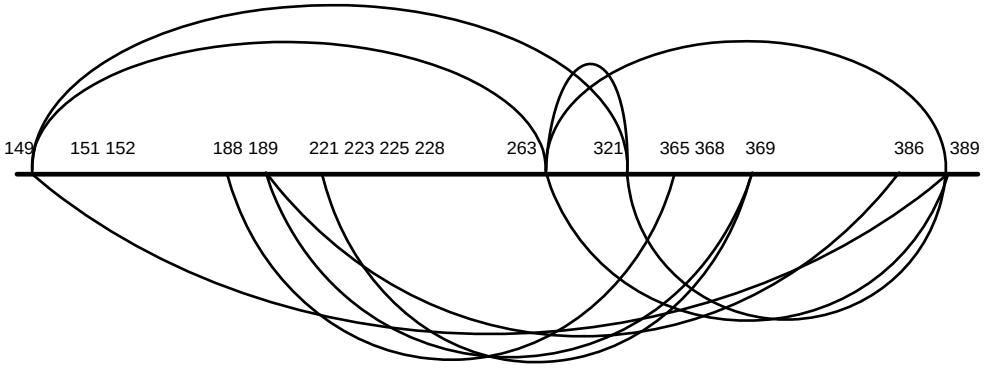


Figure 2. This plot shows the residues involved in the two networks of co-mutations with an arch to indicate the correlation between co-mutations. Three active sites 151, 152 and 368 are displayed next to their closest mutations.

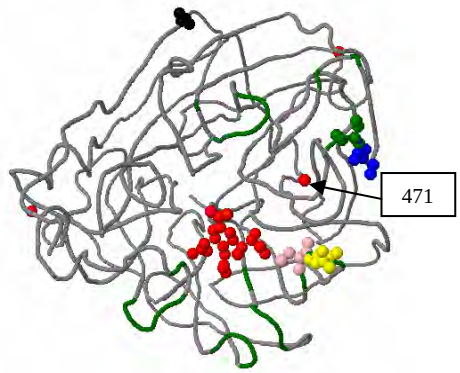


Figure 3. This plot shows in 3D structure four residues, 248, 339, 340, and 454, that have high entropy and have not mutated yet and one active site 247. Residue 248 is very close to active site 247. Calcium ion site 471 is marked to show its closeness to two residues 339 and 340. Residue 247 is in pink, 248 in yellow, 339 in blue, 340 in green, and 454 in black. The backbone of the antibody recognition sites is colored green and the bound drug (zanamivir) and three calcium ions are shown in red.

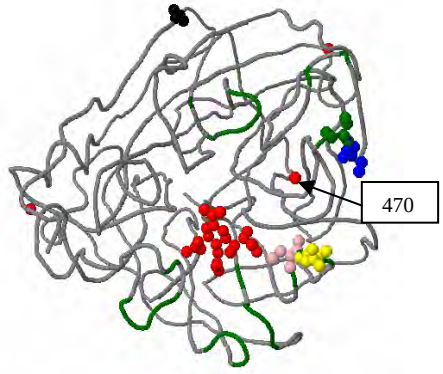


Figure 4. This plot shows in 3D structure the four residues, 149, 263, 321, and 389, in the first network of co-mutations. Residue 149 is in pink, 263 in yellow, 321 in blue, and 389 in green. Calcium ion site 470 is marked to illustrate its close position to residues 321 and 389.

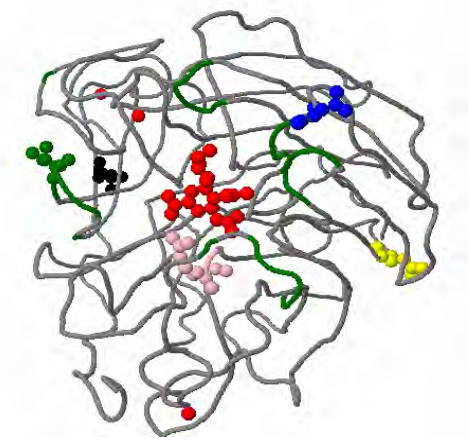


Figure 5. This plot shows in 3D structure the five residues, 188, 189, 221, 365, and 389, in the second networks of co-mutations. Residue 188 is in pink, 189 in yellow, 221 in blue, 365 in green, and 389 in black.

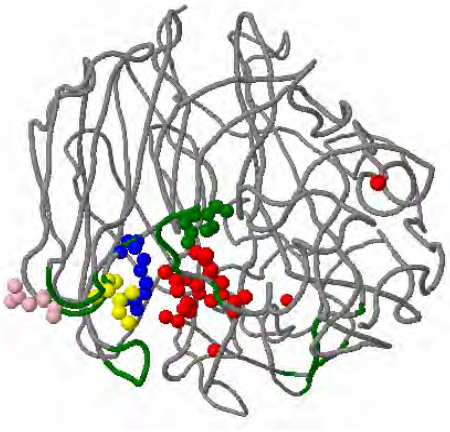


Figure 6. This plot shows in 3D structure the closeness of the four residues 221, 223, 225, and 228. One residue 221 is a part of the second network of co-mutations and the other three residues 223, 225, and 228 are active sites. Residue 221 is in pink, 223 in yellow, 225 in blue, and 228 in green.

This is a position of interest, because it is close to the active sites 223, 225 and 228 and co-mutates with a cluster of mutation positions 365, 366, and 369. The latter cluster also encloses another active site 368 (**Figures 2**, **5** and **6**).

3.2. Stalk Motifs

In this study, our intention was to discover the NA stalk motif patterns of 2009 H1N1and H5N1, which is the focus of this work, and those of 2008 and 2007 to place our findings in the right historical context. As noted in the previous section, the N1 strains in 2009 are experiencing a rapid genetic variation in the stalk region compared with other regions of N1, which could be reflected in the different stalk motifs appearing this year.

We extracted the sub-sequences consisting of positions from 36 to 79 in the NA sequences and discarded those that contained no amino acids. All the different motifs found are displayed in **Table 1**. In 2007, all NA stalk motifs in H1N1 and H5N1 had three different types referred to as types 1, 2, and 3. In 2008, type 3 stalk motif disappeared and type 4 appeared. In 2009, types 1, 2 and 4 persisted. In all these three years, types 1 and 2 persisted. Type 2, referred to as the special stalk motif in [9], was in H5N1 in all three years and type 1 was in H1N1 in all three years.

The important change in 2009 was that type 1, a common stalk motif in H1N1, was more prevalent in H5N1 and type 2, a common stalk motif in H5N1, was in H1N1 for the first time, a dramatic exchange of the two different motifs between these two subtypes of flu viruses. The special stalk motif was dominant in H5N1 in 2007 and 2008, but it was no longer the case in 2009. The new type 4 was more evident in H5N1 in 2009. These new patterns or exchange of different patterns of stalk motifs reminded us again of the fast evolutionary nature of the flu virus and the need for timely analysis of its data. The occurrence of the special stalk motif in 2009 H1N1, which may bring increased virulence to the current swine flu epidemic, is worthy of further attention and surveillance. These alterations in the stalk region of NA in H1N1 and H5N1 could also be reflected in the phylogenetic analysis conducted in the next section.

3.3. Phylogenetic Analysis

Even though the phylogenetic tree of a small number of NA sequences was constructed before [5]. With new H1N1 NA sequences being deposited at the NCBI web site regularly, it is constructive to perform phylogenetic analysis on these new sequences. For easy comparison, software MEGA [21] was used to reproduce the phylogenetic tree in **Figure 3** in [5] with the same NA sequences, which had eight sequences of 2009 H1N1 (numbered from 1 to 8) available at the time of the research (as of April 29th) in [5] and 44 different representative sequences of H1N1 or H5N1(numbered from 9 to 52) in previous years (left plot of **Figure 7**). We employed Random Forest to cluster these sequences to get a different view of their phylogenetic relationship (**Figure 8**), where a number is used to represent a sequence due to the limited space in that plot. The association of these numbers with their sequences can be found in **Figure 7**, where a number is printed before each flu subtype such as “2 H1N1 California 09 2009” and “6 H1N1 Texas 05 2009”. Random Forest clustering revealed that all eight 2009 H1N1 sequences, which formed their own single cluster, were similar to others only in the second scaling coordinate, but not in the first. We reasoned this was because all eight sequences were from the same country and the structures of these clusters might be different if the plentiful sequences deposited recently at the NCBI web site were used.

We took note of some minor but interesting differences between the phylogenetic tree in the left plot of **Figure 7** and Random Forest-based clusters of the same sequences in the left plot of **Figure 8**. In **Figure 8**, the eight sequences 1, 2, 3, …, 8 were clustered in one cluster away from all the other sequences. This cluster was close to two groups of sequence numbers in the second scaling coordinate. The first group consisted of sequences

Table 1. Different NA stalk motifs

36th ---------- Stalk region ----------79th	Subtype	Year	Number of strains shared this motif	Motif type number
HS IQTGSQNHTGICNQRIITYENSTWVNHTYVNINNTNVVAG KD	H1N1	2007	196	1
HS IQTGNQHQAEP----------ISNTNFLTE KA	H5N1	2007	189	2
-- ---------NQNQVEP----------ISNTNFLTE KA	H5N1	2007	1	3
HS IQIGSQGYPETCNQSVITYENNTWVNQTYINISNTNLIGG QA	H5N1	2007	2	1
HS IQTGSQNNTGICNQRIITYENSTWVNHTYVNINNTNVVAG ED	H1N1	2008	147	1
-- ----------NHTYVNINNTNVVAG ED	H1N1	2008	1	4
HS IQTGNQCQAEP----------ISNTKFLTE KA	H5N1	2008	70	2
HS IQLGNQNQIETCNQSVITYENNTWVNQTYVNISNTNFAAG QS	H1N1	2009	157	1
HS IQTGNQCQDEP----------ISNTKFLTE KA	H1N1	2009	21	2
----------SVITYENNTWVNQTYVNISNTNFAAG QS	H1N1	2009	39	4
HS INTGNQHQAEP----------ISNANFLTE KA	H5N1	2009	1	2
HS IQLGNQNQIETCNQSVITYENNTWVNQTYVNISNTNFAAG QS	H5N1	2009	20	1

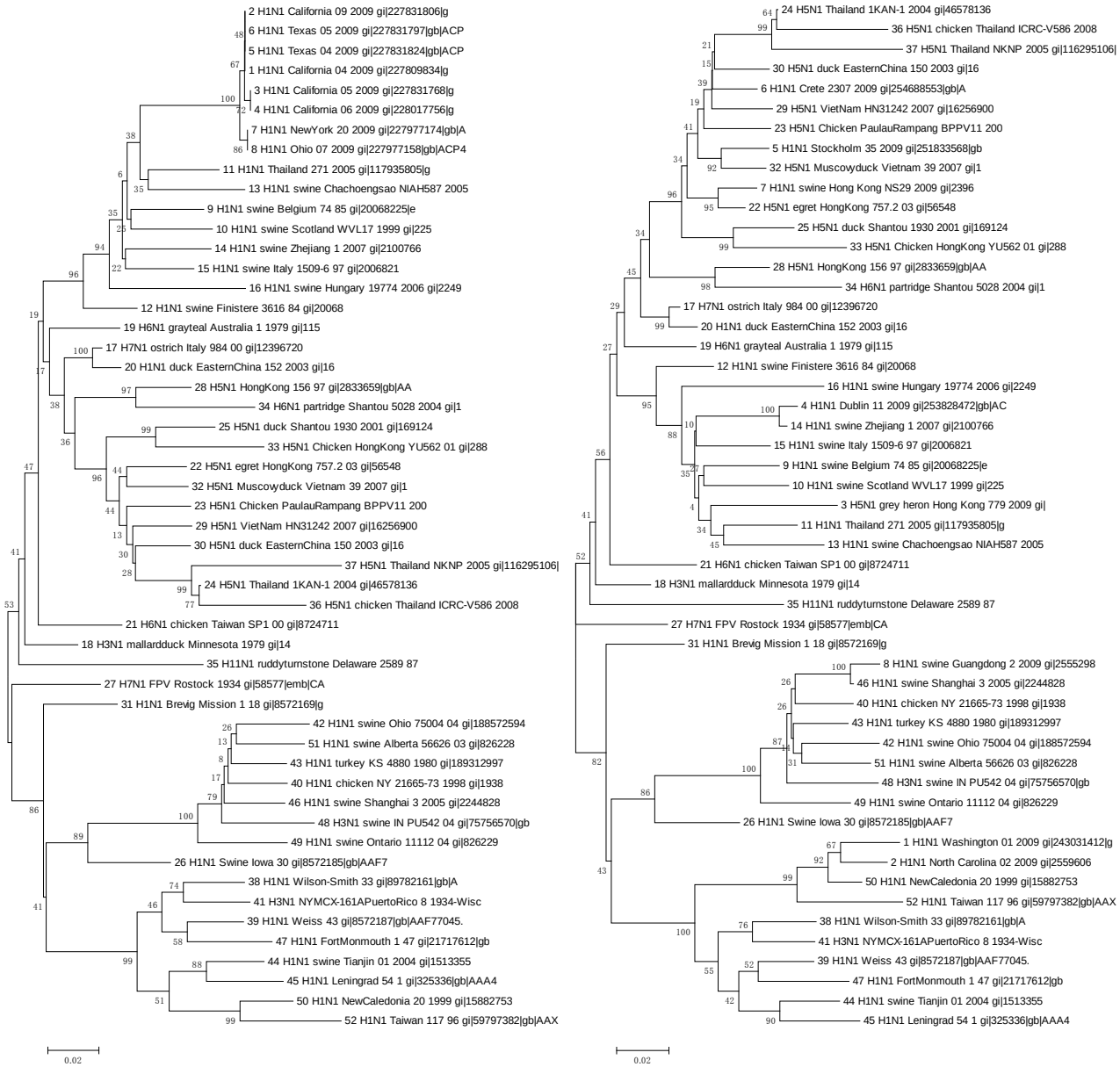


Figure 7. Left plot: a reproduced phylogenetic tree of the NA protein sequences of the N1 subtype family in [5]. Right plot: A phylogenetic tree of the same sequences used in the left plot except that the first eight sequences were replaced with eight new representative sequences from 283 N1 sequences in 2009.

13, 14, 16, 10, and 11, and the second had sequences 38, 35, and 31. The phylogenetic tree displayed the close relationship between the first group and the group of sequences 1, 2,…., 8, but did not do so for the second group. The first group had H1N1 swine Chachoengsao 2005 (13), H1N1 swine Zhejiang 2007 (14), H1N1 swine Hungary 2006 (16), H1N1 swine Scotland 1999 (10), and H1N1 Thailand 2005 (11). The second group had H1N1 Wilson Smith 33 (38), H11N1 Delaware 1987 (35), and H1N1 Brevig Mission 1918 (31). These two groups were all H1N1 subtype and the second group inherited from the ancient strains in the first group, and that is why they were similar and clustered together by Random Forest.

To get a new view of the clusters of the NA sequences of 2009 H1N1 and H5N1 available up to date, eight new representative sequences were selected from 283 sequences in the same year with cd-hit [22] to replace the eight NA sequences, numbered from 1 to 8, used in the left plot of **Figures 7** and **8**. The new sequences were H1N1 Washington 2009 (1), H1N1 North Carolina 2009 (2), H5N1 Hong Kong 2009 (3), H1N1 Dublin 2009 (4), H1N1 Stockholm 2009 (5), H1N1 Crete 2009 (6), H1N1 swine Hong Kong 2009 (7), and H1N1 swine Guang-

dong 2009 (8). In the Random Forest-based clusters in the right plot of **Figure 8**, there were three major branches with the center made of ancient NA sequences. Sequences 1, 2, and 8, sequences 3 and 4, and sequences 5, 6, and 7 were on each branch separately, which was also similarly reflected in the phylogenetic tree in the right plot of **Figure 7**. Sequences 1 and 2 were close to sequences 50 and 52 and sequence 8 was close to sequences 46 and 43. Sequences 3 and 4 were close to sequences 18, 16, 10, 13, 12, 9, 15, 11, and 14, which were mainly European swine strains. Sequences 5 and 6 were close to sequences 30 and 32 and sequence 7 was close to sequences 22, 25, and 33. All eight representative sequences were evenly spread in various clusters in the right plot of **Figure 8** and the phylogenetic tree in the right plot of **Figure 7**, which illustrated that as the 2009 H1N1 strains continue to evolve, diverse genetic makeup of the sequences would develop.

4. CONCLUSIONS

The recent outbreak of the novel 2009 H1N1 influenza has raised global concerns regarding its virulence and pandemic potential. Several recent studies on the 2009 H1N1 virus used small datasets to conduct analysis. With new sequences available up to date, we were able to extend the previous research in three areas using entropy, mutual information, and Random Forest. The first was finding two networks of co-mutations that may potentially affect the current flu-drug binding sites on NA. The second is discovering a special stalk motif, which was dominant in the H5N1 strains in the past, in the 2009 H1N1 strains for the first time. Due to the high virulence of this motif, the second finding is significant in our current research on 2009 H1N1. The third was updating the phylogenetic analysis of the NA sequences of 2009 H1N1 and H5N1, which demonstrated that, in clear contrast to previous findings, the N1 sequences in

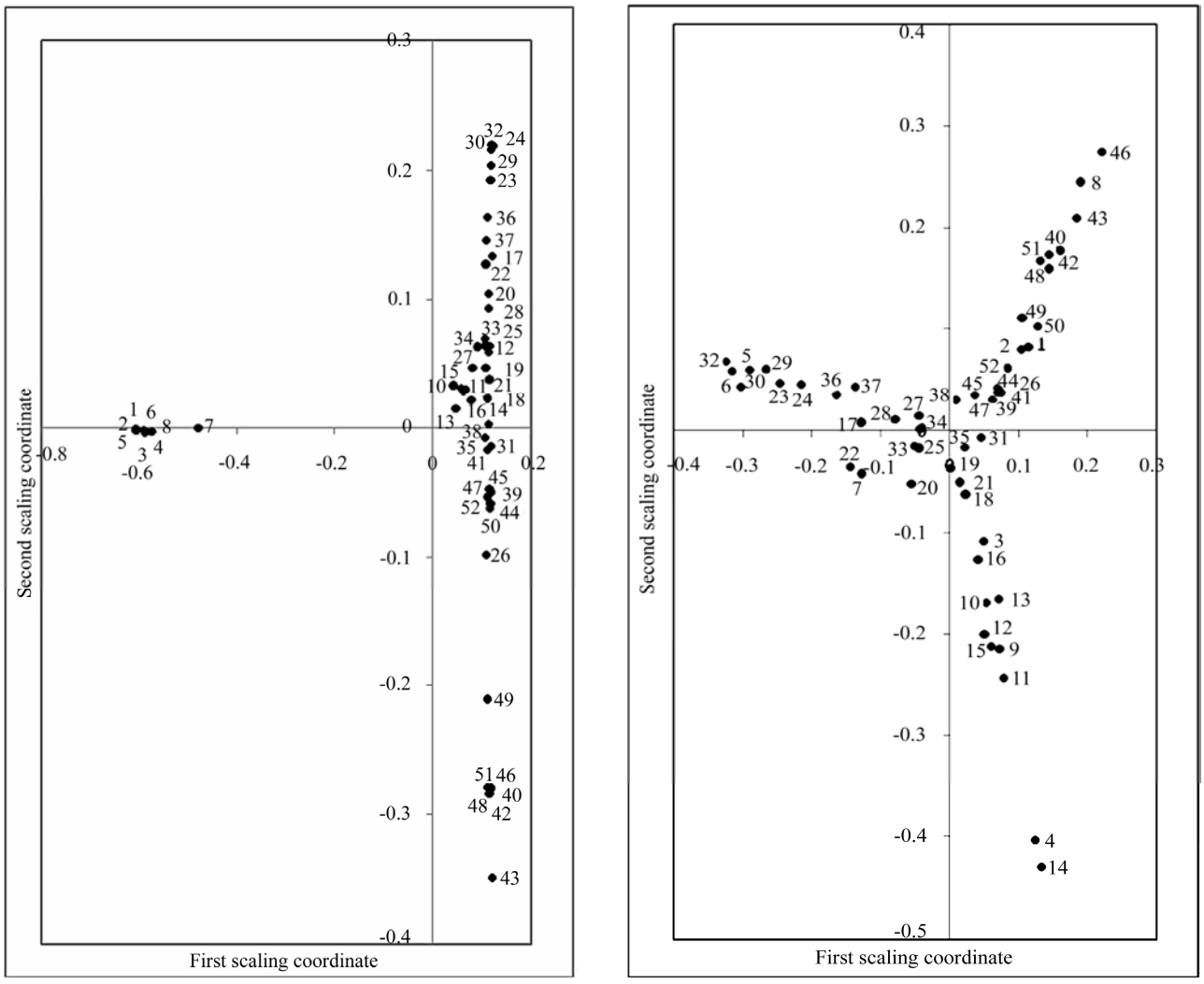


Figure 8. Left plot: Random Forest-based clusters of the same NA sequences used in the left plot of **Figure 7**. Right plot: Random Forest- based clusters of the same NA sequences used in the right plot of **Figure 7**.

2009 are diverse enough to cover different major branches of the phylogenetic tree of those in previous years. As the novel influenza A H1N1 virus continues to spread globally, our results highlighted the value of performing timely analysis on the 2009 H1N1 virus.

5. ACKNOWLEDGMENTS

We thank Houghton College for its financial support.

REFERENCES

[1] Taubenberger, J. K. and Morens, D. M., (2006) 1918 Influenza: the mother of all pandemics, Emerg. Infect. Dis., **12(1)**, 15–22.

[2] Trifonov, V., Khiabanian, H., and Rabadan, R., (2009) Geographic dependence, surveillance, and origins of the 2009 influenza A (H1N1) virus, N. Engl. J. Med., **361**, 115–119.

[3] Centers for Disease Control and Prevention (CDC), (2009) Update: Drug susceptibility of swine-origin influenza A (H1N1) viruses, MMWR Morb Mortal Wkly Rep 2009, **58**, 433–435.

[4] Weïwer, M., Chen, C. C., Kemp, M. M., and Linhard, R. J., (2009) Synthesis and biological evaluation of non-hydrolyzable 1,2,3-triazole-linked sialic acid derivatives as neuraminidase inhibitors, European Journal of Organic Chemistry, **16**, 2587.

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

[6] Wang, S. Q., Du, Q. S., Huang, R. B., Zhang, D. W., and Chou, K. C., (2009) Insights from investigating the interaction of oseltamivir (Tamiflu) with neuraminidase of the 2009 H1N1 swine flu virus, Biochemical and Biophysical Research Communications, **386(3)**, 432–6.

[7] Cover, T. A. and Thomas, J. A., (1991) Elements of information theory, John Wiley and Sons, NewYork.

[8] MacKay, D., (2003) Information theory, inference, and learning algorithms, Cambridge University Press.

[9] Zhou, H. B., Yu, Z. J., Hu, Y., Tu, J. G., Zou, W., Peng, Y. P., Zhu, J. P., Li, Y. T., Zhang, A. D., Yu, Z. N., Ye, Z. P., Chen, H. C., and Jin, M. L., (2009) The special neuraminidase stalk-motif responsible for increased virulence and pathogenesis of H5N1 influenza A virus, PLoS One, **4(7)**, e6277.

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

[11] Liu, Y., Eyal, E., and Bahar, I., (2008) Analysis of correlated mutations in HIV-1 protease using spectral clustering, Bioinformatics, **24(10)**, 1243–1250.

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

[13] Du, X. J., Wang, Z., Wu, A. P., Song, L., Cao, Y., Hang, H. Y., and Jiang, T. J., (2008) Networks of genomic co-occurrence capture characteristics of human influenza A (H3N2) evolution, Genome. Res., **18**, 178–187.

[14] Xia, Z., Jin, G. L., Zhu J., and Zhou, R. H., (2009) Using a mutual information-based site transition network to map the genetic evolution of influenza A/H3N2 virus, Bioinformatics, **25(18)**, 2309–2317.

[15] Huang, J. W., King, C. C., and Yang, J. M., (2009) Co-evolution positions and rules for antigenic variants of human influenza A/H3N2 viruses, BMC Bioinformatics, **10(Suppl 1)**, S41.

[16] Breiman, L., (2001) Random forests, Machine Learning, **45(1)**, 5–32.

[17] Shi, T., Seligson, D., Belldegrun, A. S., Palotie, A., and Horvath, S., (2005) Tumor classification by tissue microarray profiling: Random forest clustering applied to renal cell carcinoma, Mod. Pathol., **18(4)**, 547–57.

[18] Cox, T. F. and Cox, M. A. A., (2001), Multidimensional scaling, Chapman and Hall.

[19] http://www.stat.berkeley.edu/~breiman/RandomForests/.

[20] Xu, X. J., Zhu, X. Y., Dwek, R. A., Stevens, J., and Wilson, I. A., (2008) Structural characterization of the 1918 influenza virus H1N1 neuraminidase, Journal of Virology, **82(21)**, 10493–10501.

[21] Kumar, S., Nei, M., Dudley, J., and Tamura, K., (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences, Brief Bioinformatics, **9**, 299–306.

[22] Li, W. and Godzik, A., (2006) Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics, **22**, 1658–1659.

J. Biomedical Science and Engineering, 2009, 2, 559-563 **JBiSE**
doi: 10.4236/jbise.2009.27081 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

MRTF-A transactivates COMT gene and decreases the anti-tumor effects of tamoxifen

Zhi-Peng Liu*, Xue-Gang Luo*, Shu Guo, Jian-Xin Wang, Xin Zhang, Nan Wang, Yong Jiang, Tong-Cun Zhang

Key Laboratory of Industrial Microbiology, Ministry of Education, College of Biotechnology, Tianjin University of Science and Technology, Tianjin, China; *The authors contributed equally to this work; Correspondence to: Pro. Tong-Cun Zhang.
Email: tony@tust.edu.cn

Received 29 June 2009; revised 21 August 2009; accepted 1 September 2009.

ABSTRACT

Myocardin-related transcription factors A (MRTF-A) is a myocardin-related transcription factor that have been found strongly activated CarG box–containing genes through its direct binding to serum response factor (SRF). In the present study, the MRTF-A expression vector was constructed. The MTT assay showed that transfection of MRTF-A could significantly decrease the anti-tumor effect of tamoxifen on MCF-7 breast cancer cells. The bioinformatics analysis found that the CarG element existed in the promoter region of COMT gene of many familiar vertebrates, including of human, rhesus macaque, chimpanzee, etc. The results of RT-PCR assay further showed that MRTF-A could enhance the transcription level of COMT. These results are the first to indicate that COMT might be a target gene which could be regulated by MRTF-A/SRF, and such transactivation event might be involved in the process of tamoxifen resistance.

Keywords: MRTF-A; Tamoxifen; COMT; Breast Cancer

1. INTRODUCTION

Breast cancer is the most common cancer diagnosed in women in the world. Seventy percent of diagnosed breast cancers express Estrogen Receptor alpha (ER*a*) and are likely to be hormone-responsive. The most common therapy for ER*a*–positive breast cancers has employed the use of selective estrogen receptor modulators (SERMs) such as tamoxifen. As an adjuvant therapy in breast cancer, tamoxifen improves overall survival, and its widespread use is thought to have made a significant contribution to the reduction in breast cancer mortality seen over the last decade. However, although many patients benefit from tamoxifen, the resistance is an important clinical problem [1,2].

Myocardin-related transcription factors (MRTFs); including myocardin, MRTF-A/MKL1/MAL, and MRTF-B/MKL2, comprise a family of related transcriptional coactivators. MRTFs drives transcription through interaction with the ubiquitous transcription factor serum response factor (SRF), which acts on a responsive element $CC(A/T)_6GG$ (known as CarG box) that is commonly found in many gene promoters. Myocardin is expressed specifically in cardiac and smooth muscle cells and activates muscle genes associated with the differentiation of these cell types. MRTF-A and MRTF-B are expressed in a broad range of cell types [3,4].

Previous publications have suggested that MRTF-A might plays a role in development of mammary myoepithelial cells and breast cancer [5,6]. In the present paper, we further found that MRTF-A could decrease the anti-tumor effect of tamoxifen on MCF-7 human breast cancer cells, and the transcriptional activation of catechol-O-methyltransferase (COMT), a phase II metabolising enzyme of tamoxifen, might be involved in this process.

2. METERIALS AND METHODS

2.1. Regents and Cell Culcure

Tamoxifen (Wako Pure Chemicals, Osaka, Japan) was dissolved in Phosphate Buffered Saline (PBS). 1kb DNA ladder marker was purchased from Solarbio Science & Technology Co., Ltd (Beijing, China). Human liver cell lines L02 and breast cancer cell line MCF-7 was cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, Paisley, UK) containing 10% fetal bovine serum (FBS; Gibco).

2.2. Construction of MRTF-A Expression Vector

Total RNA of the LO2 cells was isolated using the method described in molecular cloning. First-strand cDNA was synthesized from the total RNA using M-MLV reverse transcriptase (Promega, Madison, USA) and oligo (dT)18 (Sangon, Shanghai, China). The full-ength MR-

TF-A gene was amplified by polymerase chain reaction (PCR) using the cDNA and the following primer pair: 5'-CAAGGTACCATGCCGCCTTTGAAAAG -3' (forward) and 5'- CCCGAATTCAGCCAGAGAGCTA- CAAGC -3' (reverse). PCR was performed at 94°C for 5 min, then 28 cycles at 94°C for 45 s, at 60°C for 45 s, and at 72°C for 3 min; extension was carried out at 72°C for 10 min. The PCR product (2824 bp) was double- digested with kpnI and EcoRI (Takara, Kyoto) and inserted into the pcDNA3.1 (+) mammalian expression vector (Invitrogen, Carlsbad, Canada). The recombinant construction was analyzed by restriction-enzyme digestion and sequencing to determine reading frame orientation and confirm sequence fidelity, and the positive recombinant plasmid was named pcDNA-MRTF-A.

2.3. Cell Transfection

Transient transfection of pcDNA-MRTF-A plasmid into the MCF-7 cells using Lipofectamin^{TM} 2000 transfection reagents (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. As a negative control, the empty vector pcDNA3.1 (+) (mock) was transfected in parallel, and to assess the transfection efficiency, the pEGFP-C3 plasmid was also transfected as a positive control simultaneously.

2.4. Evaluation of the Inhibition Rate of Tamoxifen on MCF-7 Cells

The MCF-7 human breast cancer cells were plated in 96-well plates at a density of 8×10^3 cells/well and transfected with the pcDNA-MRTF-A plasmid or pcDNA3.1 (+). After 24h, the medium was removed and replaced with fresh medium containing tamoxifen at concentrations of 5, 10 and 20 μM. Forty-eight hours later, MTT assay was performed. In brief, MTT (5 mg/ml) was added to the wells (20 μl/well). The plate was incubated in a cell incubator for 4 h, then the supernatant was removed and 150 μl of dimethyl sulfoxide was added into each well. After incubation for 10 min, the absorbance of each well was measured using a microplate reader (Bio-Rad) with a wavelength of 570 nm, with the reference wavelength set at 630 nm. Absorbance (A) was directly proportional to the number of viable cells. All assays were performed using six replicates. The inhibition rate was calculated as follows:

$$\text{Inhibition rate}(\%)=(1-\frac{A_{\text{experimental}}}{A_{\text{control}}})\times100$$

2.5. Bioinformatics Screen of the CarG Box in the Promoter of COMT Gene

The genomic sequence of COMT gene was extracted from the Database of GenBank. Their transcription start sites were then located by using the University of California at Santa Cruz (UCSC) genome brower, and the sequence from 2000 bp upstream of the transcription sites to 200 bp downstream of the transcription start sites (–2000~+200) was extracted by using the UCSC database assemblage, and then the potential CarG box was analyzed using the primer premier 5.0 software [7].

Table 1. Primers used in RT-PCR analysis. (a. F: forward primer. R: reverse primer)

Gene	Primers[a]	Product (bp)
GAPDH	F: 5'-ATTCAACGGCACAGTCAAGG-3' R: 5'-GCAGAAGGGGCGGAGATGA-3'	213
MRTFA	F: 5'-ACCGTGACCAATAAGAATGC-3' R: 5'-CCGCTCTGAATGAGAATGTC-3'	269
COMT	F: 5'-GGATTTCGCTGGCGTGAAG-3' R: 5'-TCCACCACCTCCCTGTATTCC-3'	306

2.6. RT-PCR Analysis

Total RNA of the transfected cells was extracted, and then the potential residual genomic DNA was eliminated with RNase-free Dnase I (Bio Basic Inc, Ontario, Canada). First-strand cDNA was synthesized as described above. For PCR amplification, primers specific for the cDNA of the MRTF-A, COMT gene and the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene were used (**Table 1**). PCR was performed at 94°C for 5 min, then 25 cycles at 94°C for 30 s, at 54°C for 30 s, and at 72°C for 45 s; extension was carried out at 72°C for 10 min. PCR products were electrophoretically separated in 1.5% agarose gels and visualized by ethidium bromide staining. The densities (D) of the bands were analyzed with Quantity One software and relative mRNA levels were deduced from the ratio of the mean values of MRTF-A or COMT to that of GAPDH. Cells transfected with the empty plasmid pcDNA3.1 (+) was used as blank control. The assay was performed using three replicates. The relative mRNA level was defined as:

$$\text{Relative mRNA level}=\frac{D_{\text{target gene}}}{D_{\text{GAPDH}}}$$

2.7. Statistical Analysis

The data from the above mentioned experiments were expressed as mean ± SD. The statistical significance of differences was determined using Student's t test. The minimal level of significance was $P < 0.05$.

3. RESULTS

3.1. Construction of MRTF-A Expression Vector

The cDNA encoding human MRTF-A was obtained by RT-PCR from Human liver cell lines L02 and cloned into the pcDNA3.1 (+) vector. By the methods of restriction digestion and sequence analyses, we confirmed that the length, position and orientation of the inserted MRTF-A gene were all correct, suggesting that the re-

combinant pcDNA-MRTF-A plasmid was constructed successfully (**Figure 1**).

3.2. Cell Transfection

The endotoxin-free plasmids were extracted and transfected into the MCF-7 cells using Lipofectamin™ 2000 transfection reagents. To assess the transfection efficiency, 0.5μg pEGFP-C3 plasmid was also transfected as a positive control simultaneously. As shown in **Figure 2**, the MCF-7 cell transfected with pEGFP-C3 exhibited bright fluorescence, indicating that the transfection efficiency in the experiment was high enough to perform the following experiments.

3.3. Effect of MRTF-A on the Anti-Tumor Activity of Tamoxifen

To investigate the effect of MRTF-A on the anti-tumor activity of tamoxifen, the MCF-7 breast cancer cells were transfected with the MRTF-A or the mock plasmid, and then treated with tamoxifen at different doses for 48 h. The inhibition rate was determined by MTT assay. As inhibition rates of cells transfected with MRTF-A were shown in **Figure 2**, at 5, 10 and 20 μM tamoxifen, the all significantly lower than the cells transfected with mock plasmid (p<0.05).

3.4. Bioinformatics Screen of the Potential Carg Box in the Promoter of COMT Gene

To investigate whether COMT was a potential target gene of MRTF-A, the sequence from 2000 bp upstream of the transcription sites to 200 bp downstream of the transcription start sites (–2000~+200) of COMT gene of familiar vertebrate was analyzed. As shown in **Table 2**, the CarG-like element existed in many vertebrates, including of human, rhesus macaque, chimpanzee, etc. This result indicated that COMT might be a target gene which could be regulated by MRTF-A/SRF.

3.5. Effect of MRTF-A on the Transcription Level of COMT

To further detect whether the transcription of COMT could be regulated by MRTF-A, the RT-PCR analysis was performed to detect the effect of MRTF-A on the

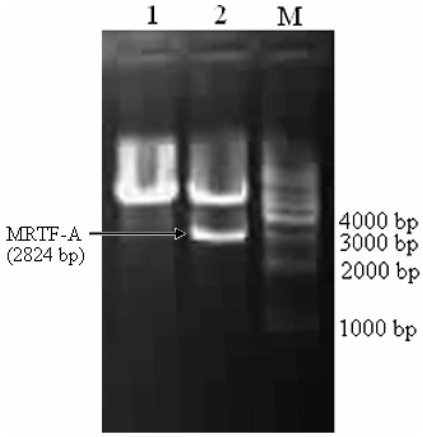


Figure 1. Identification of recombinant pcDNA-MRTF-A plasmid by restriction endonuclease. Lane 1: pcDNA3.1(+) plasmid double-digested with kpnI and EcoRI; Lane 2: pcDNA-MRTF-A double- digested with kpnI and EcoRI; Lane M: 1kb DNA ladder marker.

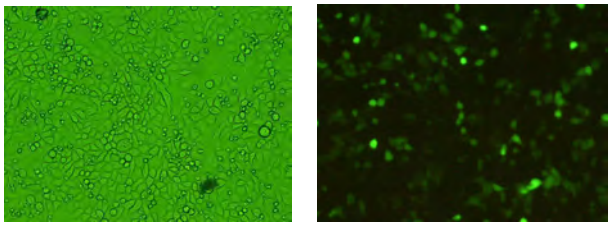


Figure 2. Cell image of MCF-7 transfected with PEGFP-C3-C3 plasmid. The left one was the normal micrograph and the right one was the fluorescence micrograph in the same visual field.

Table 2. The carg box in the promoter region of comt gene.

Organism		GenBank ID	CarG seq	Position
Sciencific name	**Common name**			
Homo sapiens	Human	NT_011519	CCTTTTATGG	-1105
Macaca mulatta	Rhesus macaque	NW_001095157	CCTTTATTGG	-933
Pan troglodytes	Chimpanzee	NW_001230944	CCTTTTATGG	-1187
Mus musculus	Laboratory mouse	NT_039624	CCTATTGGGG	-834
Rattus norvegicus	Rat	NW_047358	CCTGTTTTGG	-616
Bos taurus	Cattle	NW_001493554	CCAATACTGG	-9
Gallus gallus	Chicken	NW_001471459	CCTAAAGTGG	-1175
			CCTTATCAGG	-21
			CCTAAAAGGG	-1582
			CCCTAAAAGG	-1583
Monodelphis domestica (Opossum), Ornithorhynchus anatinus (duck-billed platypus), Bubalus bubalis (river buffalo), Canis lupus familiaris (Dog), Felis catus (Cat), Ovis aries (Sheep), Sus scrofa (Pig), Danio rerio (Zebrafish)		The typical CarG box was not found in the promoter of COMT gene. Alternatively, the COMT gene still remained unclear		

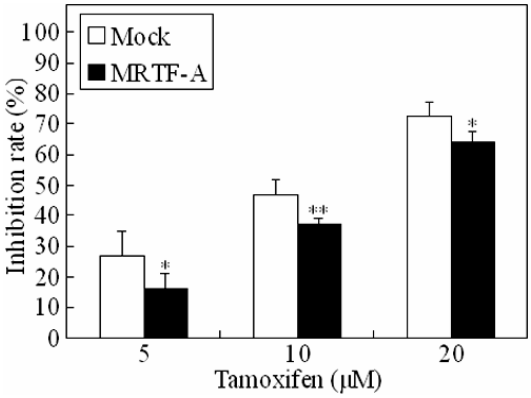


Figure 3. Decrease in anti-tumor of tamoxifen on MCF-7 breast cancer cells following transfection of MRTF-A. *p<0.05; **P<0.01 (compared with the mock).

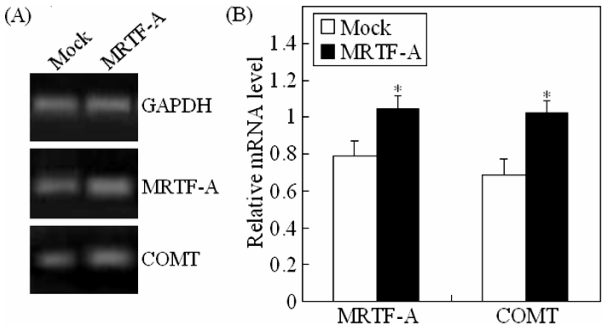


Figure 4. Promotion of COMT mRNA expression transfected with MRTF-A. (A) Electrophoretica graph of the RT-PCR assay; (B) The relative mRNA level determinded with the ratio of the band density of MRTF-A or COMT to that of GAPDH. *p<0.05; **P<0.01 (compared with the mock).

mRNA level of COMT. As shown in **Figure 3**, the relative mRNA level of MRTF-A and COMT of cells transfected with MRTF-A were both significantly higher than those of the mock cells, suggesting that MRTF-A could promote the transcription level of COMT.

4. DISCUSSION

In 1980s, tamoxifen was approved by the Food and Drug Administration (FDA) as an adjuvant therapy for the breast cancer. Despite the benefits of tamoxifen therapy, almost all tamoxifen-responsive breast cancer patients develop resistance to therapy. The drug resistance to tamoxifen therapy can take many forms, and one of the important possible mechanisms may be the metabolic activation [2,8,9].

The metabolic activation of tamoxifen involves the transformation of tamoxifen into the 4-OH-tamoxifen and 3, 4-di-OH-tamoxifen (catechol). It has been suggested that the anticancer activity of tamoxifen may be due to its 4-hydroxylated metabolite. However, because 4-OH-tamoxifen could be subsequently hydroxylated into the 3, 4-di-OH-tamoxifen, the 4-hydroxylated metabolite is usually at low observable levels in mammals. The 3, 4-di-OH-tamoxifen is a suitable substrate of COMT, which always catalyzes the transfer of a methyl group from S-adenosyl-methionine (SAM) to one of the phenolic hydroxyl groups in a variety of catechols including catechol estrogens and catecholamine neurotransmitters. In the presence of SAM, the 3, 4-di-OH-tamoxifen would be catalyzed into monomethoxy 3, 4-di-OH-tamoxifen [10,11,12].

MRTF-A is a member of Myocardin-related transcription factors family, which transactivates the promoters containing consensus CarG box. Compared with myocardin, which is expressed specifically in cardiac and smooth muscle cells, MRTF-A is expressed in a broad range of cell types and has more extensive transactivating functions than myocardin [13]. Previous studies have showed that the metabolizing function of COMT within the coronary artery may be an important determinant of the cardiovascular protective effects of circulating estradiol, and the COMT expression may be regulated in the myometrium to control the local action of estrogen [14,15,16]. In the present study, the bioinformatics analysis found that the CarG element widely existed in the promoter region of COMT gene of many familiar vertebrates, especially in human, rhesus macaque and chimpanzee, and the results of RT-PCR further showed that MRTF-A could enhanced the transcription level of COMT. These results indicated that COMT might be a target gene which could be regulated by MRTF-A/SRF, and such transactivation might be involved in the process of tamoxifen resistance. Further studies addressing the detail relationship between MRTF-A, COMT and tamoxifen and the mechanism should provide fundamental insights into the function of MRTF-A in the development and overcome of the breast cancer.

5. ACKNOWLEDGMENTS

This study was financially supported by the Tianjin natural science foundation (No. 07JCYBJC16400).

REFERENCES

[1] Deroo, B. J. and Korach, K. S., (2006) Estrogen receptors and human disease, J. Clin. Invest. Ann. Arbor., **116**, 561–570.

[2] Ring, A., and Dowsett, M., (2004) Mechanisms of tamoxifen resistance, Endocr-relat cancer, Bristol., **11**, 643–658.

[3] Wang, D. Z., Li, S., Hockemeyer, D., Sutherland, L., Wang, Z., Schratt, G., *et al*, (2002) Potentiation of serum response factor activity by a family of myocardin-related transcription factors, Proc. Natl. Acad. Sci., USA, Washington, **99**, 14855–14860.

[4] Teg Pipes, G. C., Creemers, E. E., and Olson, E. N., (2006) The myocardin family of transcriptional coactivators: Versatile regulators of cell growth, migration, and myogenesis, Gene dev, Cold Spring Harbor, **20**, 1545–1556.

[5] Sun, Y., Boyd, K., Xu, W., Ma, J., Jackson, C. W., Fu, A.,

et al, (2006) Acute myeloid leukemia-associated mkl1 (mrtf-a) is a key regulator of mammary gland function, Mol. Cell Biol., Washington, **26**, 5809–5826.

[6] Li, S., Chang, S., Qi, X., Richardson, J. A., and Olson, E. N., (2006) Requirement of a myocardin-ralated transcription factor for development of mammary myoepithelial cells, Mol. Cell Biol., Washington, **26**, 5797–5808.

[7] Sun, Q., Chen, G., Streb, J. W., Long, X., Yang, Y., Stoeckert, C. J., *et al*, (2006) Defining the mammalian CarGome, Genome. Res., Cold Spring Harbor, **16**, 197–207.

[8] Jordan, V. C., (2003) Is tamoxifen the rosetta stone for breast cancer, J. Natl. Cancer. Inst. Bethesda, **95**, 338–340.

[9] Hilsenbeck, S. G., Friedrichs, W. E., Schiff, R., O'Connell, P., Hansen, R. K., Osborne, C. K., *et al*, (1999) Statistical analysis of array expression data as applied to the problem of tamoxifen resistance, J. Natl. Cancer. Inst. Bethesda, **91**, 453–459.

[10] Mannisto, P. T. and Kaakkola, S., (2008) Catechol-O-methyltransferase (COMT): Biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors, Pharmacol. rev. Bethesda, **51**, 593–628.

[11] Singh, M. N., Stringfellow, H. E., Walsh, M. J., Ashton, K. M., Paraskevaidis, E., Abdo, K. R., *et al*, (2008) Quantifiable mRNA transcripts for tamoxifen-metabolising enzymes in human endometrium, Toxicology, Limerick, **249**, 85–90.

[12] Dehal, S. S. and Kupfer, D., (1995) Evidence that the catechol 3, 4-dihydroxytamosifen is a proximate intermediate to the reactive species binding covalently to proteins, Cancer Res. Baltimore, **56**, 1283–1290.

[13] Cen, B., Selvaraj, A., Burgess, R. C., Hitzler, J. K., Ma, Z., Morris, S. W., *et al*, (2003) Megakaryoblastic leukemia 1, a potent transcriptional coactivator for serum response factor (SRF), is required for serum induction of SRF target genes, Mol Cell Biol., Washington, 23, 6597–6606.

[14] Dubey, R. K., Gillespie, D. G., Zacharia, L. C., Barchiesi, F., Imthurn, B., and Jackson, E. K., (2003) CYP450- and COMT-derived estradiol metabolites inhibit activity of human coronary artery SMCs, Hypertension, Dallas, **41**, 807–813.

[15] Dubey, R. K., Tofovic, S. P., and Jackson, E. K., (2004) Cardiovascular pharmacology of estradiol metabolites, J. Pharmacol. Exp. Ther., Bethesda, **308**, 403–409.

[16] Wentz, M. J., Jamaluddin, M., Garfield, R. E., Al-Hendy, A., (2006) Regulation of catechol-O-methyltransferase expression in human myometrial cells, Obstet. Gynecol., Hagerstown, **108**, 1439–1447.

J. Biomedical Science and Engineering, 2009, 2, 564-573 **JBiSE**
doi: 10.4236/jbise.2009.27082 Published Online November 2009 (http://www.SciRP.org/journal/jbise/).

Signal averaging for noise reduction in anesthesia monitoring and control with communication channels

Zhi-Bin Tan[1], Le-Yi Wang[2*], Hong Wang[3]**

[1,2]Department of Electrical and Computer Engineering, Wayne State University, Detroit, USA; [3]Department of Anesthesiology, Wayne State University, Detroit, USA.
Email: [1]au6063@wayne.edu; [2]lywang@wayne.edu; [3]howang@med.wayne.edu

Received 29 May 2009; revised 14 July 2009; accepted 14 July 2009.

ABSTRACT

This paper investigates impact of noise and signal averaging on patient control in anesthesia applications, especially in networked control system settings such as wireless connected systems, sensor networks, local area networks, or tele-medicine over a wide area network. Such systems involve communication channels which introduce noises due to quantization, channel noises, and have limited communication bandwidth resources. Usually signal averaging can be used effectively in reducing noise effects when remote monitoring and diagnosis are involved. However, when feedback is intended, we show that signal averaging will lose its utility substantially. To explain this phenomenon, we analyze stability margins under signal averaging and derive some optimal strategies for selecting window sizes. A typical case of anesthesia depth control problems is used in this development.

Keywords: Anesthesia Depth; Anesthesia Monitoring; Anesthesia Control; Signal Averaging; Noise Reduction; Open and Closed Loop Systems; Communications; Networked Systems

1. INTRODUCTION

To maintain an adequate depth of anesthesia without compromising patient's health, an anesthesiologist usually works as a multi-task feedback controller to roughly regulate the drugs titration while observing a variety of patient outcomes. Automatic anesthesia controller design aims to automatically regulate anesthesia levels by taking account on several physiological measurements and then frees up anesthesiologists for more important tasks in operation. Closed-loop control of anesthesia has been a goal of many researchers since the middle of 20th century. With the emergence of BIS monitor in late 1990s, the interests in closed-loop control of depth of hypnosis is renewed, the most notable works are seen in [1,2,3,4]. In an operation room, a wide range of medical devices are connected together or connected to patient through cables for measuring, monitoring and diagnosis. The cable clutter interferes with patient care, creats hazards for clinical staff and delays transport and positioning. To improve the clinical room efficiency and safty, it has been suggested to replace those cables with wireless connections [5].

While anesthesia patient vital signs such as anesthesia depth index, blood pressure, heart rate etc. are transmitted through a noisy wireless channel in a wide area, those transmitted signals will be corrupted by the transmission noise. It is well understood that within most algorithms that reduce effects of random noises on signals and systems, some types of signal averaging are used [6,7]. This is mainly because the laws of large numbers and central limit theorems provide a foundation for noise reduction. The rationale is that when averaging is applied, noises diminish in an appropriate sense. This fundamental understanding leads to algorithms in filtering, signal reconstruction, state estimation, parameter estimation, system identification, and stochastic control. The signal averaging can be used effectively when remote monitoring and diagnosis are involved. On the other hand, signal averaging introduces dynamic delays. Such delays will have detrimental effects on closed-loop systems, even destabilizing the system. Consequently, signal averaging encounters a fundamental performance limitation in feedback systems. To explain this phenomenon, we analyze stability margins under signal averaging and derive some optimal strategies for selecting window sizes. A typical case of anesthesia depth control problems is used in this development.

This paper is organized as follows. Section 2 discusses patient modeling and control in anesthesia appli-

[*]Research of this author was supported in part by the National Science Foundation under ECS-0329597 and DMS-0624849, and Michigan Economic Development Council.
[**]Research of this author was supported in part by Michigan Economic Development Council and Wayne State University's Research Enhancement Program.

cations. A typical case is presented with a detailed patient model derived from clinical data. A feedback control is designed to achieve closed-loop system stability, on the basis of state observers and pole placement control. Signal averaging and its effectiveness on open-loop and closed-loop applications are demonstrated in Section 3. We show that while extending filter windows can improve noise attenuation in open-loop systems, it can de-stabilize a closed-loop system, implying a fundamental performance limitation. The idea of using fast sampling is discussed.

Theoretical foundation of our performance analysis is presented in Section 4. It is shown that this can be transformed into a calculation of gain margin of a modified system. Performance limitation is analyzed that leads to an optimal selection of filter windows. It is shown that for a given sampling rate, even optimally designed averaging filters can only have very limited benefits in reducing noise effects. It is shown that noise reduction ratio is proportional to the sampling interval, providing a means of obtaining noise reduction with communication resources. These findings are applied to anesthesia control problems in Section 5. Finally, Section 6 summarizes some issues that are related but not resolved in this paper.

2. PATIENT MODELS AND FEEDBACK CONTROL

Real-time anesthesia decisions are exemplified by general anesthesia for attaining an adequate anesthetic depth (consciousness level of a patient), ventilation control, etc [3,4,8,9]. One of the most critical requirements in this decision process is to predict the impact of the inputs (drug infusion rates, fluid flow rates, ventilator mode, etc.) on the outcomes (consciousness levels, blood pressures, heart rates, airway pressures, and oxygen saturation, etc.). This prediction capability can be used for control, display, warning, predictive diagnosis, decision analysis, outcome comparison, etc.

2.1. Patient Models

The core function of this prediction capability is embedded in establishing a reliable model that relates the drug or procedure inputs to the outcomes in real-time and in individual patients. Due to significant deviations in physical conditions, ages, metabolism, pre-existing medical conditions, and surgical procedures, patient dynamics demonstrate nonlinearity and large variations in their responses to drug infusion. A basic information-oriented model structure for patient responses to drug infusion was introduced in [10,11,12]. Propofol (a common anesthesia drug) titration is administered by an infusion pump. The patient's anesthesia depth is measured by a BIS (Bi-Spectrum) monitor [13,14]. The monitor provides continuously an index in the range of [0,100] such that the lower the index value, the deeper the anesthesia state. Hence, an index value 0 will indicate "brain dead" and 100 will be "awake".

To establish patient models for monitoring and control, clinical data were collected. One of these data sets is used in this paper. The anesthesia process lasted about 76 minutes, starting from the initial drug administration and continuing until last dose of administration. Propofol was used in both titration and bolus. Fentanal was injected in small bolus amount three times, two at the initial surgical preparation and one near incision. Analysis shows that the impact of Fentanal on the BIS values is minimal. As a result, it is treated as a disturbance and not explicitly modeled in this example. The drug infusion was controlled manually by an experienced anesthesiologist. The trajectories of titration (in μg/sec) and bolus injection (converted to μg /sec) during the entire surgical procedure were recorded, which are shown together with the corresponding BIS values in **Figure 1**.

The patient was given bolus injection twice to induce anesthesia, first at $t = 3$ minute with 20 mg and then at $t = 5$ minute with 20 mg. They are shown in the figure as 10000 μg /sec for two seconds, to be consistent with the titration units. The surgical procedures were manually recorded. Three major types of stimulation were identified: 1) During the initial drug administration (the first 6 minutes), due to set-up stimulation and patient nervousness. 2) Incision at $t = 45$ minute for about 5 minutes duration. 3) Closing near the end of the surgery at $t = 60$ minute.

The data from the first 30 minutes are used to determine model parameters and functional forms. For estimating the parameters in the patient block, the data in the interval where the bolus and stimulation impact is minimal (between $t = 10$ to $t = 30$ minutes) are used. The patient model parameters were identified through Least-Squares estimation method [15].

Under a sampling interval $T = 1$ second, which is the standard data transfer interval for the BIS monitor, the combined linear dynamics was estimated. The patient model with propofal infusion rate as the input and BIS measurement as the output was identified as

$$P(z) = \frac{0.01872z^2 - 0.08813z + 0.09016}{z^5 - 1.159z^4 + 0.7501z^3 - 0.5989z^2 + 0.2984z - 0.2678} \quad (1)$$

The actual BIS response is then compared to the model response over the entire surgical procedure. Comparison results are demonstrated in **Figure 2**. Here, the inputs of titration and bolus are the recorded real-time data. The model output represents the patient response very well. In particular, the model captures the

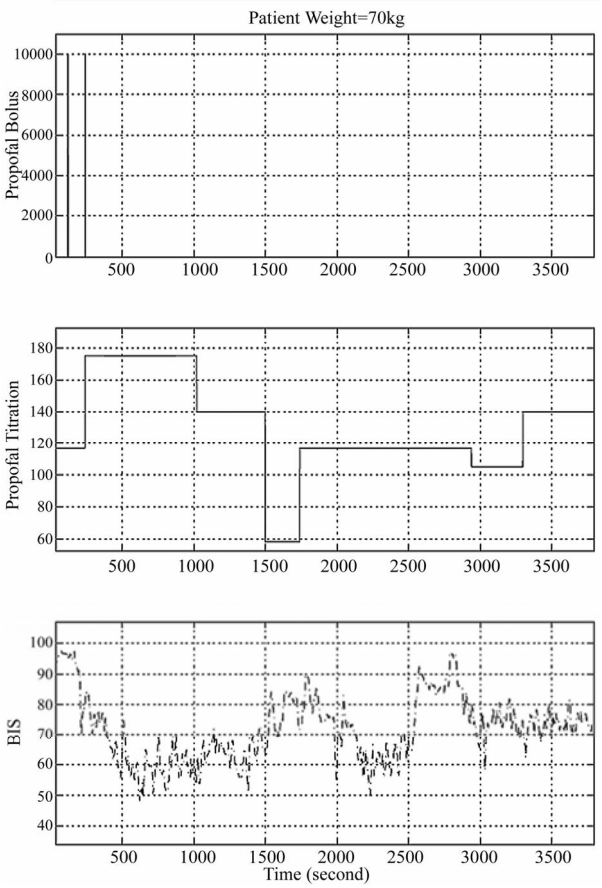


Figure 1. Actual patient responses.

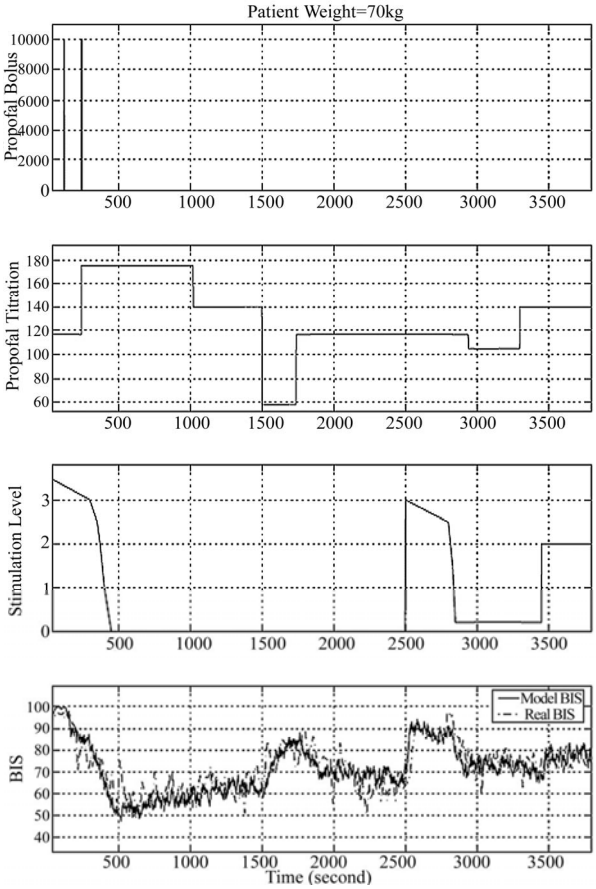


Figure 2. Patient model responses.

key trends and magnitudes of the BIS variations in the surgical procedure. This indicates that the model structure contains sufficient freedom in representing the main features of the patient response.

2.2. Feedback Control

Usually to eliminate steady-state error in tracking control, an integrator is inserted into the system

$$C(z)=\frac{1}{z-1}$$

A stabilizing feedback controller is then designed for the patient model (1) by using a full-order observer and pole placement design, leading to

$$F(z)=\frac{1.234z^5+0.6298z^4-3.644z^3+3.67z^2-1.981z+0.2479}{z^6-2.341z^5+2.284z^4-0.7252z^3-0.4057z^2+0.5714z-0.08343}$$

These system components result in a combined open-loop system

$$G(z)=F(z)C(z)P(z) \tag{2}$$

3. SIGNAL AVERAGING AND CONTROL PERFORMANCE

We will use a typical anesthesia control problem to understand impact of communication channels and utility of signal averaging on anesthesia monitoring and control. There are different window functions for signal averaging, such as uniform windows, exponential windows, etc. They are different only in their forms, but most conclusions for system analysis or error bounds are usually valid for all window types. As a result, we shall use the exponential windows to carry out our analysis. A signal averaging by exponential decaying weighting of rate $0<\alpha<1$ is

$$h_k=(1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}x_i \tag{3}$$

whose transfer function is

$$F_\alpha(z)=\frac{(1-\alpha)z}{z-\alpha} \tag{4}$$

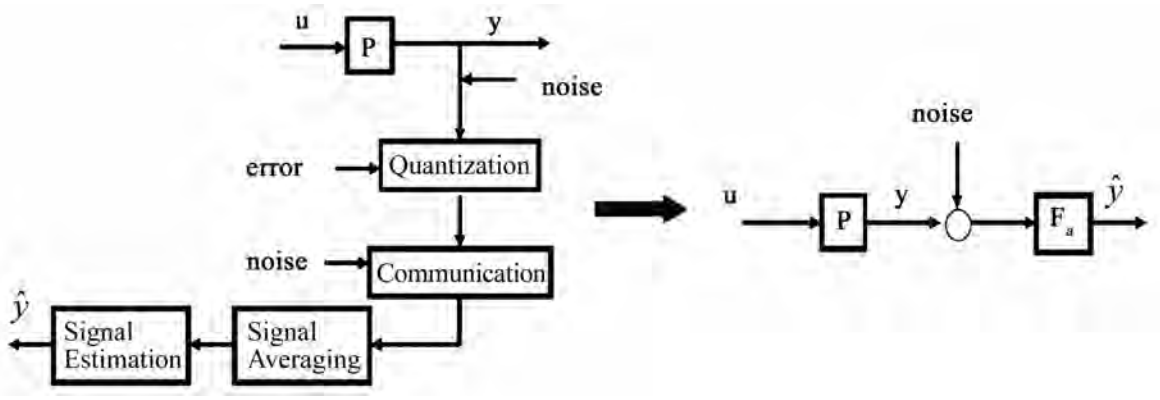


Figure 3. Signal filtering in open-loop system.

3.1. Open-Loop Systems

In wireless-based monitoring and diagnosis applications, the system is running in open-loop. In this case quantization errors and communication noises can be grouped as an additive noise to the patient output *y*. When signal averaging is applied to reduce noise effects, the resulting system can be represented by the block diagram in **Figure 3**.

Figure 4 illustrates impact of filtering on open loop systems. In open loop applications, filtering will not de-stabilize the system. Consequently, one may choose a window of long horizon to reduce the effects of noise. It is apparent that the longer the averaging window, the less the noise effect on the signal. However, it is also observed that signal averaging slows down system's response to the input. In other words, filtering introduces a dynamic delay. This delay has very important implication in the closed-loop applications.

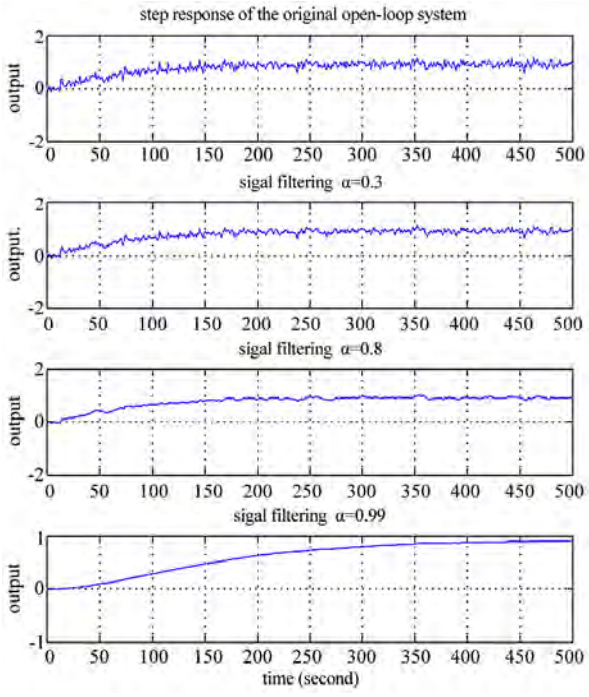


Figure 4. Effects of signal averaging on open-loop systems.

3.2. Closed-Loop Systems

On the other hand, if feedback control for anesthesia management decisions is intended, signal filtering becomes part of a closed-loop system. When signal averaging is applied, the averaging filter $\boldsymbol{F}_a$ is inserted into the system, resulting in a modified closed-loop system shown in **Figure 5**.

The close-loop system equations are:

$$y_k = Ge_k, \quad e_k = r_k - F_\alpha(y_k + d_k) \tag{5}$$

Then

$$y_k = Gr_k - GF_\alpha y_k - GF_\alpha d_k$$

and

$$y_k = H_r r_k + H_\alpha d_k \tag{6}$$

where,

$$H_r = \frac{G}{1+F_\alpha G}; \quad H_\alpha = \frac{-F_\alpha G}{1+F_\alpha G} \tag{7}$$

Figure 6 illustrates impact of filtering on closed loop systems. Although signal filtering can reduce the noise effect of the signal, it introduces a dynamic delay which has detrimental effects on the closed-loop system. The plots confirm that when filtering window is long the filter can destabilize the closed-loop system. Even when the filtering window size is small, its effectiveness is not very obvious. This example suggests that in closed-loop applications signal filtering has limited effectiveness. This understanding will be used to introduce new methods to reduce noise effects in such applications.

3.3. Re-Sampling

The plant in this case is identified as a 5th order difference equation in (1). The system can be well approxi-

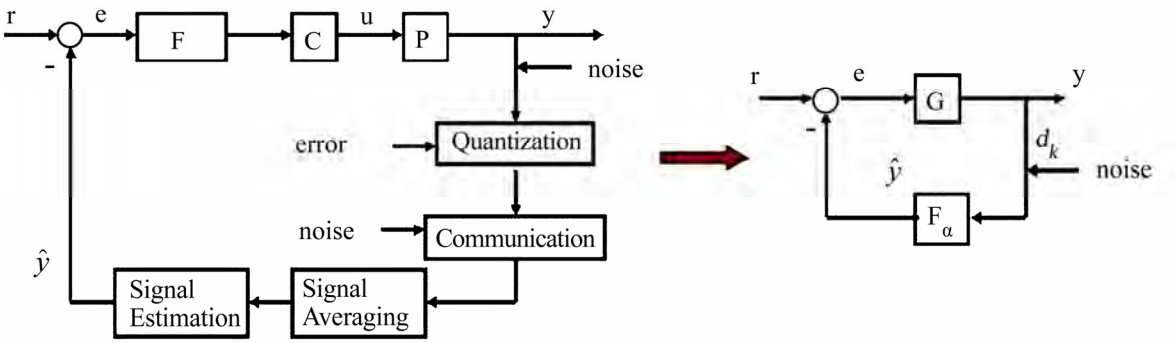


Figure 5. System modules and their equivalent representation.

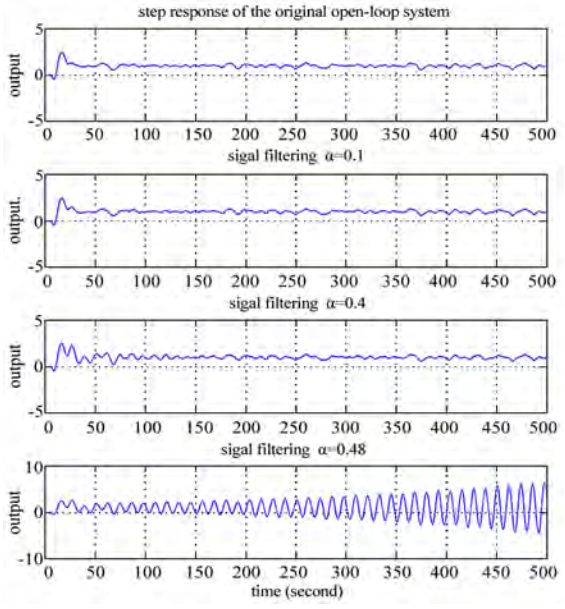


Figure 6. Effects of signal averaging on closed-loop systems.

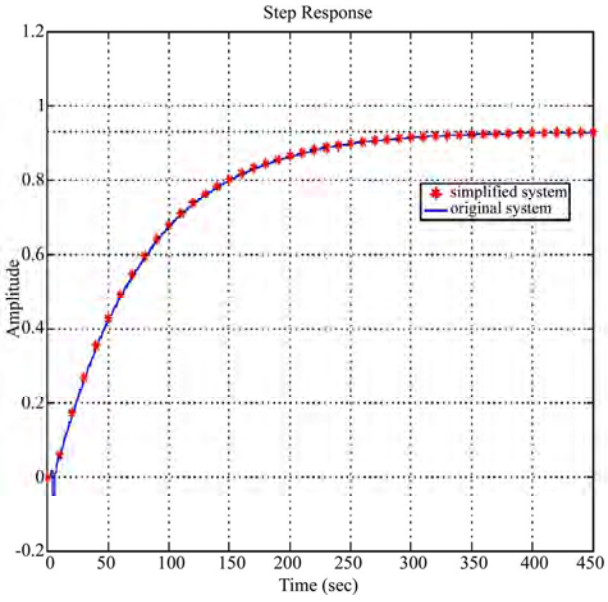


Figure 7. Step responses of the original system and the simplified system.

mated by a continuous-time system that consists of a pure time delay and a first-order dynamics, sampled with sampling interval T=1 second. Let a continuous-time system be

$$P(s) = e^{-5s}\frac{0.93}{73s+1} \quad (8)$$

The step responses of the original system (1) and the simplified system P(s) are shown in **Figure 7**.

This approximation allows us to use smaller sampling intervals to re-sample the output of the system. The benefits of re-sampling will become clear after some theoretical analysis in the next section.

4. ANALYSIS OF STABILITY AND PERFORMANCE

Definition 1 *The stability margin against exponential averaging, abbreviated as α-margin and denoted by* $\alpha_{max}(G)$, *is the largest* $0 \le \alpha \le 1$ *such that for all* $0 \le \alpha < \alpha_{max}(G)$, *the close-loop system (6) is stable and the system is unstable if* $\alpha > \alpha_{max}(G)$. *If the close-loop system is stable for all α, we denote* $\alpha_{max}(G) = 1$.

Suppose that the input to the filter is a noise corrupted constant

$$x_k = \theta + d_k$$

An exponential window of rate $0 < \alpha < 1$ is applied to this signal and its output is

$$h_k = (1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}x_i = \theta + (1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}d_i = \theta + \varepsilon_k$$

where

$$\varepsilon_k = (1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}d_i$$

If d_i is i.i.d. (independent and identically distributed)

with $Ed_i = 0$ and $Ed_i^2 = \sigma^2$, then $E\varepsilon_k = 0$ and

$$E\varepsilon_k^2 = \frac{1-\alpha}{1+\alpha}\sigma^2$$

Consequently, using h_k as an estimate of θ can reduce errors by $1-\alpha 1+\alpha$. We will call α as the decaying rate. Convergence (in the mean square sense) is achieved when $\alpha \to 1$: $\lim_{\alpha\to 1} E\varepsilon_k^2 = 0$.

Consider an exponential filter

$$F(s) = \frac{1}{\lambda s + 1} \tag{9}$$

whose impulse response is

$$f(t) = \frac{1}{\lambda} e^{-t/\lambda}, t \geq 0 \tag{10}$$

Note that

$$\int_0^\infty f(t)dt = 1$$

Now, the input-output relationship of this filter is

$$y(t) = \int_{-\infty}^{t} f(t-\tau)x(\tau)d\tau$$

$$= \frac{1}{\lambda}\int_{-\infty}^{t} e^{-(t-\tau)/\lambda}x(\tau)d\tau$$

When the signals are sampled with sampling interval $\boldsymbol{T}$, we denote $x_k = x(kT)$ and $y_k = y(kT)$. If α is related to λ and $\boldsymbol{T}$ by $\alpha = e^{-T/\lambda}$, we have

$$\lim_{T\to 0} \frac{T}{\lambda(1-\alpha)} = 1$$

For small $\boldsymbol{T}$, $y(t)$ is approximated by

$$y_k = y(kT) = \frac{1}{\lambda}\int_{-\infty}^{t} e^{-(t-\tau)/\lambda}x(\tau)d\tau$$

$$\approx \frac{T}{\lambda}\sum_{i=-\infty}^{k}(e^{-T/\lambda})^{k-i}x_i = \frac{T}{\lambda(1-\alpha)}(1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}x_i \approx (1-\alpha)\sum_{i=-\infty}^{k}\alpha^{k-i}x_i$$

In other words, for system analysis, we may approximate the discrete-time filter in (3) by its continuous-time counterpart in (9). These relationships between discrete-time averaging and continuous time averaging will be used to derive stability margins.

4.1. Stability Margin against Exponential Averaging

The relationship between α and λ will allow us to focus on stability analysis in continuous time systems and then transform the results to the discrete-time filters. This is stated in the following theorem.

Theorem 1 *If the exponential stability margin in the continuous-time domain is* λ_{max}, *then*

$$\lim_{T\to 0} \frac{T}{-\ln \alpha_{max}} = \lambda_{max}$$

Proof: This follows from the relationship

$$\alpha = e^{-T/\lambda}$$

We now concentrate on calculation of λ_{max}.

Definition 2 *The stability margin against exponential averaging for the continuous-time closed-loop system, abbreviated as continuous exponential A-margin and denoted by* $\lambda_{max}(G)$, *is the smallest* $\lambda > 0$ *under which the closed-loop system becomes unstable. If the closed-loop system remains stable for all* $\lambda > 0$, *we denote* $\lambda_{max}(G) = \infty$.

Suppose $G(s) = N(s)/D(s)$ where $n(s)$ and $d(s)$ are polynomial functions of *s* and coprime (that is, $N(s)$ and $D(s)$ do not have common zeros). Then λ_{max} is the largest $\lambda > 0$ before the closed-loop system becomes unstable. Consider the characteristic equation of the closed-loop system

$$1 + F_\lambda(s)G(s) = 1 + \frac{1}{\lambda s + 1}\frac{N(s)}{D(s)} = 0$$

or

$$\lambda s D(s) + D(s) + N(s) = 0 \tag{11}$$

which leads to

$$1 + \lambda\frac{sD(s)}{D(s)+N(s)} = 0 \tag{12}$$

This expression leads to the following conclusion.

Theorem 2 *The exponential A-margin* $\lambda_{max}(G)$ *of* $G(s)$ *is the gain margin of*

$$H(s) = \frac{sD(s)}{D(s)+N(s)} \tag{13}$$

We make several interesting observations from (12). First, from (11), λ_{max} may be calculated by using the Routh-Hurwitz test. Second, (12) is in a standard form for using root locus technique. So, we may plot the root locus of the system (13) (it is an improper system) and detect the λ value that reaches marginal stability, which will be λ_{max}. The root locus plot starts at the poles of system (13) which are precisely the poles of the closed-loop system without the averaging filter. Since

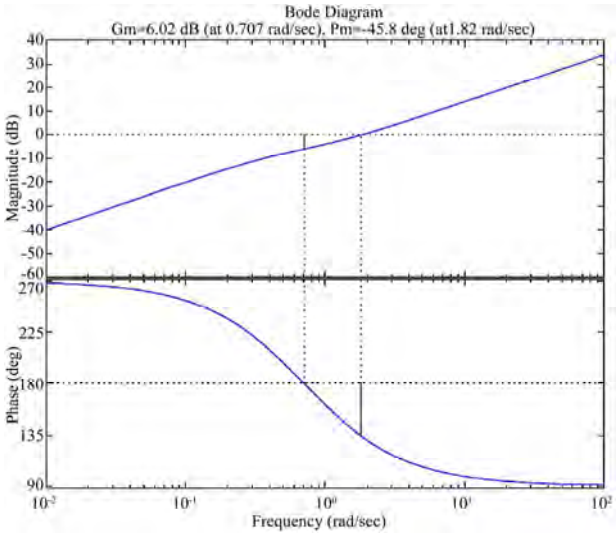


Figure 8. Using bode plots to obtain the gain margin.

the closed-loop system is stable, for small λ the closed-loop system with the filter will remain stable. The root locus plot moves towards the zeros of system (13) which are the poles of the open-loop system. Hence, if the open-loop system is unstable, the exponential A-margin is always finite.

Example 1 Suppose $G(s) = (s+2)/(s-1)$. Then,

$$H(s) = \frac{sD(s)}{D(s)+N(s)} = s^2 - s2s + 1$$

The gain margin can be obtained by using the Matlab function "margin" (which gives $\lambda_{max} = 2$) or by plotting the bode plot as shown in **Figure 1** which gives $\lambda_{max} = 6.02\,dB = 2$. Alternatively, from

$$(2s+1) + \lambda(s^2 - s) = \lambda s^2 + (2-\lambda)s + 1 = 0$$

we can calculate $\lambda_{max} = 2$ by the Routh-Hurwitz method.

4.2. Performance Analysis

Within the A-margin, what is the benefit of signal averaging? On one hand, signal averaging can reduce noise effect. On the other hand, averaging introduces delays and reduces closed-loop system performance. Consequently, an optimal choice of averaging becomes an issue.

Similarly, the continuous time close-loop system equations are:

$$y = \frac{G}{1+F_\lambda G} r - \frac{F_\lambda G}{1+F_\lambda G} d \qquad (14)$$

Here, we denote

$$H_\lambda = \frac{-F_\lambda G}{1+F_\lambda G} \qquad (15)$$

If d is a white noise, noise attenuation aims to reduce the L^2 norm of H_λ. Naturally, for optimal noise reduction, we should select

$$\eta = \inf_{0<\lambda<\lambda_{max}} \|H_\lambda\|_2 \qquad (16)$$

Example 2 For the system in Example 1, when λ takes values 0,0.1,…0.9, the corresponding H^2 norms for the closed-loop system H_λ are

λ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
η	6.22	7.00	7.87	9.00	10.49	12.59	15.73	20.96	31.44	62.85

The monotone increase of the L^2 norms indicates that for this system, averaging cannot reduce noise impact on the output. As a result, there should be no averaging for this system.

Example 3 For another example, consider a system

$$G(s) = \frac{s^2 + 2s - 1}{s^2 - s + 4}$$

The closed-loop system's characteristic equation is

$$\lambda sD(s) + D(s) + N(s)$$
$$= \lambda s^3 + (2-\lambda)s^2 + (1+4\lambda)s + 3 = 0$$

It can be calculated by the Routh-Hurwitz method that $\lambda_{max} = 1.366$. The H^2 norm of H_λ as a function of λ is plotted in **Figure 3**. The optimal averaging occurs at $\lambda = 0.59$ with the norm $\|H_{0.59}\|_2 = 2.5263$.

From the relationship $\alpha = e^{-T/\lambda}$, for small sampling interval T,

$$\alpha = e^{-T/\lambda} = e^{-T/0.59} = e^{-1.7T}$$

is the optimal rate for averaging in the discrete-time domain. For example, if $T = 0.01$, we obtain $\alpha = 0.983$.

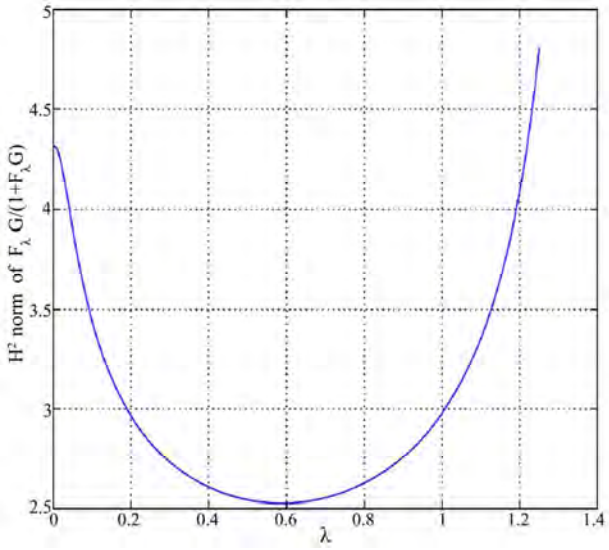


Figure 9. Optimal averaging rate.

4.3. Fast Sampling for Disturbance Attenuation

Although the optimal L^2 performance η in (16) cannot be improved for the continuous-time system, noise attenuation in the sampled system can be further improved.

We first establish a relationship between the L^2 norm of the continuous-time system and the l^2 norm of its sampled system. Suppose that the disturbance sequence d_k passes through a ZOH of interval T to become $d(t)$. The continuous-time system H is stable with impulse response $h(t)$. Then,

$$y(t) = \int_0^t h(t-\tau)d(\tau)d\tau$$

Suppose d_k is a pulse sequence, $d_0 = 1$, and $d_k = 0$, $k \neq 0$. Then, $d(t) = 1$, $0 \le t < T$, and $d(t) = 0$, otherwise. Under this input

$$y(t) = \int_0^T h(t-\tau)d\tau$$

Hence, the sampled values of $y(t)$, which form the pulse response of the sampled system, become

$$y_k = y(KT) = \int_0^T h(kT-\tau)d\tau$$

For small T this can be approximated by

$$y_k = Th(kT)$$

We note that for small T,

$$\|H\|_2^2 = \int_0^\infty h^2(t)dt \approx T\sum_{k=0}^\infty h^2(kT)$$

Consequently, if we use $\tilde{H}$ and $\tilde{h}_k$ to denote the sampled system and its impulse response, we have $\tilde{h}_k = Th(kT)$ and

$$\|\tilde{H}\|_2^2 = \|\tilde{h}_k\|_2^2 = T^2\sum_{k=0}^\infty h^2(kT) = T\|H\|_2^2$$

From

$$y_k = \sum_{i=0}^k \tilde{h}_{k-i}d_i$$

if d_k is i.i.d., mean zero and variance σ^2, then

$$\sigma_k^2 = Ey_k^2 = \sum_{i=0}^k\sum_{j=0}^k \tilde{h}_{k-i}Ed_id_j\tilde{h}_{k-j}$$

$$= \sigma^2\sum_{i=0}^k \tilde{h}_{k-i}^2 \le \sigma^2\|\tilde{h}_k\|_2^2 = \sigma^2T\|H\|_2^2$$

In fact,

$$\sigma_{max}^2 = \sup_k \sigma_k^2 \approx \sigma^2T\|H\|_2^2$$

If $\|H\|_2^2$ is optimized, then $\|H\|_2^2 = \eta^2$ as in (16). Consequently, the noise reduction ratio can be expressed as

$$\tilde{\eta}^2 = T\eta^2 \quad (17)$$

This is a relationship between noise reduction in the sampled system and the optimal L^2 norm of the continuous-time system. This analysis concludes that using faster sampling (smaller T) can reduce the noise effects.

5. CONTROL WITH SIGNAL AVERAGING

Findings from Section 4 provide some useful design guidelines. 1) Signal averaging is beneficial in reducing noise effects. 2) Effectiveness of signal averaging in closed-loop systems varies substantially with the filter windows or decaying rates. There is an optimal decaying rate at which signal filtering becomes most effective. 3) When the filter window is optimally selected, further noise attenuation can only be achieved by increasing the sampling rates. 4) Increasing sampling rates incurs higher bandwidth requirements for communications. When channel bandwidths are limited, there is a fundamental performance limit for noise attenuation. This is a unique feature for closed-loop systems. In open-loop applications, convergence can be obtained by applying signal averaging over a very long horizon. However, this cannot be applied to closed-loop systems since long windows of filtering destabilize the feedback system.

5.1. Anesthesia Applications

We now apply these understandings to anesthesia control systems. The open-loop transfer function in (2) can be derived as

$$G(z) = \frac{N(z)}{D(z)}$$

with

$$N(z) = 0.02311z^7 - 0.09699z^6 - 0.01243z^5 + 0.4466z^4 - 0.689z^3 + 0.5101z^2 - 0.2005z + 0.02235$$

and

$$D(z) = z^{12} - 4.5z^{11} + 9.248z^{10} - 11.48z^9 + 9.576z^8 - 5.684z^7 + 2.528z^6 - 0.7518z^5 - 0.2721z^4 + 0.6608z^3 - 0.507z^2 + 0.2003z - 0.02234$$

Since the open loop system is unstable, the stability margin of the closed-loop system with inserted averaging window is always limited. The closed-loop system's stability concerns have already been depicted in **Figure**

6. The closed-loop system's H^2 norm, which defines the system's ability in noise attenuation, is shown in **Figure 10**.

Then, the closed-loop system's step response is simulated when the filter is optimally selected and shown in **Figure 11**. The system inherent sampling rate is T=1 second.

While re-sampling is performed with T=1, the H^2 norm of the closed-loop system will be reduced further. For reduced sampling intervals, improvements of noise attenuation are illustrated in **Figure 12**.

5.2. Discussions

It can be seen from **Figure 10** that the optimal filter decaying rate is $\alpha_{opt} = 0.1300$ with the corresponding H^2 norm 9.0872 when T=1. The closed-loop system is stable and it's step response has much fluctuation in steady state. From the relationship, $\alpha_{opt} = e^{-T/\lambda_{opt}} = e^{-1/\lambda_{opt}}$, we obtain $\lambda_{opt} = 0.49$. This leads to the optimal choice of

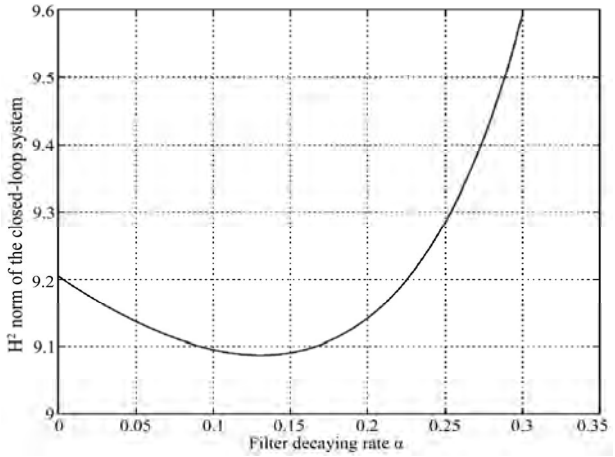


Figure 10. Closed-loop system performance vs. filter decaying rates.

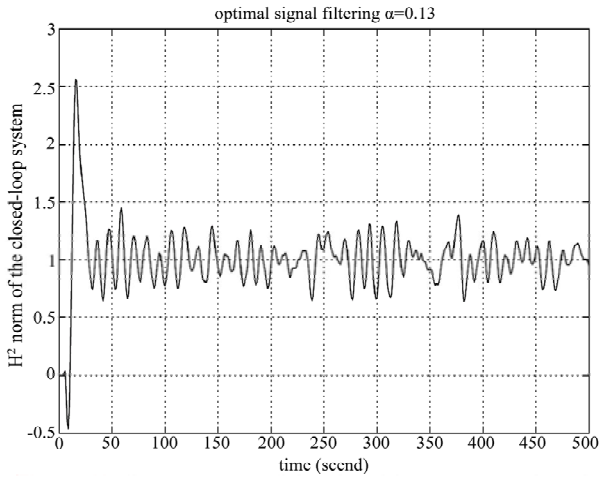


Figure 11. Step response of the closed-loop system when the filter is optimally selected, and sampling interval T=1.

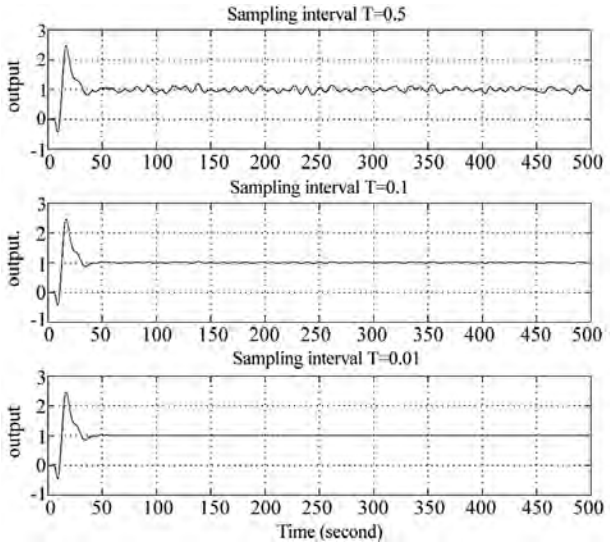


Figure 12. The closed-loop system performance for reduced sampling intervals.

decaying rate when the sampling interval T is reduced from 1 as

$$\alpha = e^{-T/\lambda_{opt}} = e^{-T/0.49} = e^{-2.04T}.$$

When sampling rate is increased to 1T, the H^2 norm of the closed-loop system will be reduced to 9.0872T as established in (17). **Figure 12** illustrates the step responses of the closed-loop system with sampling interval T=0.5, T=0.1 and T=0.01 second respectively. The steady state fluctuation of the step response is decreasing with the reduced sampling intervals.

6. CONCLUSIONS

The impact of communication channels on feedback control in anesthesia applications in wireless based systems was investigated in this paper. Such systems involve communication channels which are corrupted by noises and have limited bandwidth resources. Signal averaging is the fundamental method in dealing with stochastic noises and errors. It is used effectively in reducing noise effects when only remote monitoring and diagnosis are involved. However, the case is different when feedback is intended.

Our results show that the decaying rate of the averaging window has significant impact on the performance of the close-loop system. When α is larger than some value, the close-loop system becomes unstable. A concept of stability margins against exponential averaging is introduced. Its calculation can be performed by either the Routh-Hurwitz method or the root-locus method on a modified system. Furthermore, the strategy for choosing the optimal decaying rate is derived. Our results conclude that fast sampling must be used for improving noise reduction after optimal filter design. The analysis and design method is applied to anesthesia patient con-

trol problems.

Our analysis is conducted on the basis of the linear systems. Actually, anesthesia patient models contain nonlinearity. Our future work will consider analysis of nonlinear systems.

REFERENCES

[1] Nunes, C. S., Mendonca, T., Lemos, J. M., and Amorim, P., (2007) Predictive adaptive control of the bispectral index of the EEG (BIS): Exploring electromyography as an accessible disturbance, Mediterranean Conference on Control and Automation, Athens-Greece.

[2] Gentilini, A., Rossoni-Gerosa, M., and Morari, M., *et al*, (2001) Modeling and closed-loop control of hypnosis by means of bispectral index (BIS) with isoflurane, IEEE Trans. on Biomedical Engineering, **48**, 874−889,.

[3] Dong, C., Kehoe, J., Henry, J., Ifeachor, E. C., Reeve, C. D., and Sneyd, J. R., (1999) Closed-loop computer controlled sedation with propofol, Proc. of the Anaesthetic Research Society, 631.

[4] Zhang, X. S., Roy, R. J., and Huang, J. W., (1998) Closed-loop system for total intravenous anesthesia by simultaneously administering two anesthetic drugs, Proc. of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology, 3052–3055.

[5] Goldman, J. M., (2006) Medical device connectivity for improving safety and efficiency, American Society of Anesthesiologists Newsletters, **70(5)**, http://www.asahq.org/Newsletters/2006/05−06/goldman05_06.html.

[6] Talbot, S. L. and Boroujeny, B. F., (2008) Spectral method of blind carrier tracking for OFDM, IEEE Transactions on Signal Processing, **56(7)**.

[7] Bataillou, E., Thierry, E., Rix, H., and Meste, O., (1995) Weighted averaging using adaptive estimation of the weights, Signal Processing, **44**, 5l−66.

[8] Eisenach, J. C., (1999) Reports of Scientific Meetings—Workshop on Safe Feedback Control of Anesthetic Drug Delivery, Anesthesilogy, **91**, 600–601.

[9] Linkens, D. A., (1992) Adaptive and intelligent control in anesthesia, IEEE Control Systems Magazine, 6–11.

[10] Wang, L. Y. and Wang, H., (2002) Control-oriented modeling of BIS-based patient response to anesthesia infusion, Internat. Conf. Math. Eng. Techniques in Medicine and Bio. Sci., Las Vegas.

[11] Wang, L. Y. and Wang, H., (2002) Feedback and predictive control of anesthesia infusion using control-oriented patient models, Internat. Conf. Math. Eng. Techniques in Medicine and Bio. Sci., Las Vegas.

[12] Wang, L. Y., Wang, H., and Yin, G., (2002) Anesthesia infusion models: Knowledge-based real-time identification via stochastic approximation, 41st IEEE Cont. and Dec. Conf., Las Vegas.

[13] Gan, T. J., *et al*., (1997) Bispectral index monitoring allows faster emergence and improved recovery from propofol, Alfentanil, and Nitrous Oxide Anesthesia, Anesthesiology, **87**, 808–815.

[14] Rosow, C. and Manberg, P. J., (1998) Bispectral index monitoring, Annual of Anesthetic Pharmacology, **2**, 1084–2098.

[15] Ljung, L. and Söderström, T., (1983) Theory and Practice of Recursive Identification, MIT Press, Cambridge, MA.

Call for Papers

The 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE 2010)

June 18-20, 2010 Chengdu, China

The 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE 2010) will be held from June 18th to 20th, 2010 in Chengdu, China. You are welcome to share your recent advances and achievements in all aspects of bioinformatics and biomedical engineering on the conference. **And all accepted papers in iCBBE 2010 will be published by IEEE and indexed by Ei Compendex and ISTP.**

Topics

Bioinformatics and Computational Biology

- Protein structure, function and sequence analysis
- Protein interactions, docking and function
- Computational proteomics
- DNA and RNA structure, function and sequence analysis
- Gene regulation, expression, identification and network
- Structural, functional and comparative genomics
- Computer aided drug design
- Data acquisition, analysis and visualization
- Algorithms, software, and tools in Bioinformatics
- Any novel approaches to bioinformatics problems

Biomedical Engineering

- Biomedical imaging, image processing & visualization
- Bioelectrical and neural engineering
- Biomechanics and bio-transport
- Methods and biology effects of NMR/CT/ECG technology
- Biomedical devices, sensors and artificial organs
- Biochemical, cellular, molecular and tissue engineering
- Biomedical robotics and mechanics
- Rehabilitation engineering and clinical engineering
- Health monitoring systems and wearable system
- Bio-signal processing and analysis
- Biometric and bio-measurement
- Biomaterial and biomedical optics
- Other topics related to biomedical engineering

Special Sessions

Biomedical imaging
Biostatistics and biometry
The information technology in bioinformatics
Environmental pollution & public health

Sponsors

IEEE Eng. in Medicine and Biology Society, USA
Gordon Life Science Institute, USA
University of Lowa, USA
Wuhan University, China
Sichuan University, China
Journal of Biomedical Science and Engineering, USA

Important Dates

Paper Due: Oct.30, 2009
Acceptance Notification: Dec.31, 2009
Conference: June 18-20, 2010

Contact Information

Website:http://www.icbbe.org/2010/
E-mail: submit@icbbe.org

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

Editor–in–Chief

Editorial Board

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

ISSN 1937-6871 (Print), 1937-688X (Online)

Journal of Biomedical Science and Engineering, 2009, 2.
Published Online November 2009 in SciRes (http://www.scirp.org/journal/jbise).

TABLE OF CONTENTS

Volume 2, Number 7, November 2009

9771937687001010