www.scirp.org/journal/jbise

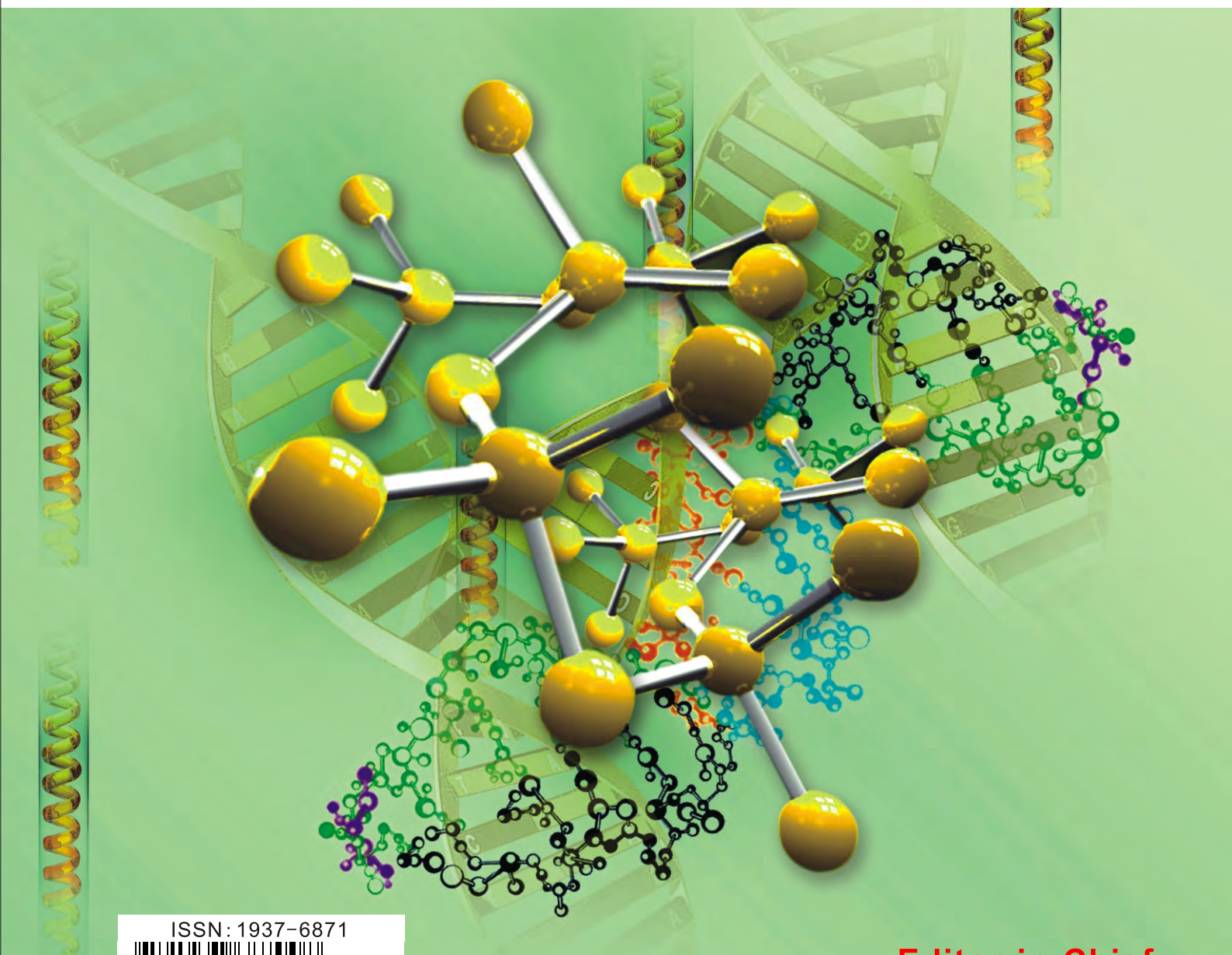
Scientific Research

JBiSE

ISSN: 1937-6871

Volume 3 Number 5 May 2010

Journal of Biomedical Science and Engineering



ISSN: 1937-6871

Editor-in-Chief
Kuo-Chen Chou

Journal Editorial Board

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

TABLE OF CONTENTS

Volume 3, Number 5, May 2010

Journal of Biomedical Science and Engineering (JBiSE)

Journal Information

SUBSCRIPTIONS

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

Subscription rates:

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

SERVICES

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

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

COPYRIGHT

PRODUCTION INFORMATION

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

J. Biomedical Science and Engineering, 2010, 3, 443-447 **JBiSE**
doi:10.4236/jbise.2010.35061 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Silent brain infarctions and leuko-araiosis in Chinese patients with first-ever acute lacunar strokes

Peterus Thajeb[1,2,3,4*], Wen-Yuan Lee[2], Chung-Hung Shih[5], Teguh Thajeb[6], James Davis[3], Rosanne Harrigan[3], Linda Chang[3]

[1]Center for Stroke Care and Prevention, Cathay General Hospital Sijhih, Sijhih, Taiwan, China;
[2]Section of Neurosurgery and Neurology, China Medical University Hospital, Taipei, Taiwan, China;
[3]Biomedical Science, Graduate Division, John A Burns School of Medicine, University of Hawaii at Manoa, Hawaii, USA;
[4]Faculty of Medicine, National Yang-Ming University, Taipei, Taiwan, China;
[5]Institute of Respiratory Therapy, Taipei Medical University, Taipei, Taiwan, China;
[6]Department of Internal Medicine, Landseed Hospital, Pingjen, Taiwan, China.
Email: peterus@hawaii.edu; thajebp@hotmail.com

Received 3 January 2010; revised 12 January 2010; accepted 15 January 2010.

ABSTRACT

We report on silent brain infarction (SBI) and leuko-araiosis (LA) of 23 patients with clinically diagnosed "first-ever" acute ischemic lacunar stroke. The lacunar syndromes were pure motor hemiparesis (10), pure sensory syndrome (2), ataxic hemiparesis (3), dysarthria clumsy hand syndrome (3), and sensory-motor deficit (5). Nineteen out of the 23 patients presented with completed strokes on arrival to the hospital, and 4 (17%) developed evolving-stroke within 24 hours of stroke onset. A lacune corresponded to the acute stroke could be found in all patients on brain magnetic resonance imaging (MRI), and in 18 (78%) on brain computed tomography (CT). MRI showed additional subclinical or asymptomatic "silent brain infarctions or lacunes" (SBI) in 19 (83%) of 23 patients, and leuko-araiosis (LA) of moderate to severe degree (≥ grade 2) was present in 61% of patients although dementia was absent. Hypertension is the risk factor in 78% of cases followed by diabetes mellitus, smoking, and elevated plasma cholesterol level. Independence of the types of lacunar syndromes, patients with hypertension and diabetes mellitus are associated with high grade LA. None with normal blood pressure and plasma glucose had grade 3 or grade 4 LA ($p < 0.05$). In conclusion, evolving-stroke occurs in one-fifth of patients with "first-ever" lacunar infarct within the first 24 hours of stroke onset. SBI was found in 83% of cases. Hypertension and diabetes mellitus are associated with additional SBI and high grade LA. The severity of leuko-araiosis per se dictates the cerebrovascular risks.

Keywords: Computed Tomography; First-Ever Stroke; Lacune; Leuko-Araiosis; Magnetic Resonance Imaging; MRI; Silent Brain Infarction

1. INTRODUCTION

Elegant study on 1042 routine autopsied brains by Fisher in 1965 [1,2], and data from 2,859 necropsied cases reported by Tuszynski *et al.* [3] suggest that lacunes occur in between 6% [3] and 11% [1,2] of cases, respectively. Approximately, eighty percent of these lacunes had no history of stroke and neurological deficit, and thus called "asymptomatic" or "silent" lacunes [3,4]. During life, the diagnosis of asymptomatic or subclinical silent lacunes in the pre-era of MRI has been a difficult task for several reasons. First, ethically a neurological healthy subject would not justify having an intensive laboratory work-up for any possible reason of unexpected intracranial silent lesion. Second, in the earlier days, before the advent of MRI technique, a brain CT would not sensitive enough to detect a small lesion of less than 15 mm in a clinically overt lacunar stroke. Therefore, the actual incidence of subclinical silent lacunes in the general population, and in asymptomatic patients with cerebrovascular risk factors remains conjectural. Extensive use of brain MRI in the past decade had suggested that prevalence of MRI-based diagnosis of silent brain infarction (SBI) ranged from 8% [5] to 28% [6]. These SBIs included silent lacunes and silent non-lacunar infarctions [5-8]. A screening health examination in Japan on 246 neurologically normal adults revealed 13% had possible silent lacunar lesions on brain MRI [9]. But nothing is known about SBI in Chinese/Taiwanese with "first-ever" acute ischemic lacunar syndromes. This issue will be addressed herein.

2. PATIENTS AND METHODS

Patients recruited in this study were a subset from the

authors' consecutive stroke registry between January 1, 2007 and December 31, 2008. Subjects with a lacunar infarction due to small penetrating artery occlusion are often considered "minor" stroke by both laymen and many clinicians. Good recovery of symptoms and signs in certain proportion of cases leads to underestimation of the prevalence of lacunar infarction and silent brain infarction. In order to obtain a homogeneous study population, we address the question that how often a silent brain infarction (SBI) can be seen in patients with clinically "first-ever" lacunar stroke? Therefore, patients with non-lacunar infarction were excluded from the study. Only those fulfilling all diagnostic criteria set below for acute first-ever lacunar infarction and those with complete clinical data and available brain CT and MRI within 5 days of stroke onset will be analyzed. The diagnostic criteria of "first-ever" acute lacunar infarction were: 1) acute stroke with clinical manifestation of one of the 5 typical clinical lacunar syndromes described by Fisher [1,2]; 2) absence of past history of presumed stroke or transient ischemic attack (TIA), including amourosis fugax; 3) normal electroencephalography (EEG); 4) neuroimaging studies (brain MRI or CT or both) showing a corresponded acute brain infarction in the territory of the penetrating artery and of the size of a "lacune" (diameter of 15 mm or less); and 5) exclusion of those with presumed vascular dementia on mini-mental state examination (MMSE) [10] or modified mini-mental test (MMT) [11], or both, and on Hatchinski ischemic score [12]. Twenty-three patients thus recruited consisted of 6 women and 17 men. The age at stroke onset ranged from 51 to 85 years (mean 65 years).

MRI was performed by using either one of the following protocol. The old MRI machine (0.5 Tesla superconductive unit, MRI-50 A, manufactured by Toshiba Corporation) was a multiple-spin-echo T1-weighted image with TR 500 msec/TE 15 msec, T2-weighted images with TR 3000 msec/TE 120 msec, and proton density images (PDI) with TR 3000 msec/TE 30 msec. The lesions could be round or oval in shape. The single lesion that matched completely to the clinical presentation (one of the 5 lacunar syndromes) was called the corresponding lesion. Other unexpected additional (old) lesions or lacunes were called additional lesions or "silent" lacunes. A smaller "dot-like", or spotty high signal intensity on T2WI and PDI at the cerebal subcortical region that beyond the territory of the penetrating arteries, may or may not be seen on T1WI MRI or brain CT were considered "etat crible" [13]. The hyperintense poorly demarcated pathches of irregular border surrounding the lateral ventricles were designated as "leuko-araiosis" (LA) or periventricular "caps" or "rims" [14]. The grading of severity of LA was as follows: LA grade 0 (no leuko-araiosis), LA grade 1 (mild leuko-araiosis at the frontal horn subcortical white matter), LA grade 2 (moderate LA involving the frontal and posterior horn white matter), LA grade 3 (severe LA involving the whole rims of lateral ventricular white matter), and LA grade 4 (advanced LA, grade 3 LA plus involvement of the centrum semiovale). The new MRI machine has T1WI [TI 860.0/TR 2,200.0/TE 11.0], T2WI [TR 3,300.0/TE 113.0], FLAIR [TI 2,500.0/TR 9,000.0/TE 113.0], diffusion weighted images (DWI) [TR 3,600.0/TE 84.0], ADC map, and magnetic resonance angiogram (MRA) [TR 37.0/TE 7.0]. MRI definition of a lacune remains the same for T1WI and T2WI. Additional imaging parameters for acute lacunar infarction were lesion of high signal on T2FLAIR, and DWI, and low signal on ADC map. Asymptomatic or subclinical "silent" (old) cerebral infarctions were characterized by signal intensity of iso-or low-signal on DWI and iso-or high signal on ADC map.

Statistical analyses: Categorical Chi-square or Fisher's exact test (for $n \leq 5$) were used for statistical analyses. Probability of < 0.05 was considered statistical significance.

3. RESULTS

Mean age at onset was 65 years (range 51 to 85 years) with men to women ratio of 2.8. There was no statistical difference between gender and types of lacunar syndromes. Additional lesions suggesting silent cerebral lacunar infarctions were encountered in 83% of patients on MRI (**Table 1**) as compared to 43% on CT ($p < 0.005$). Hypertension was top one risk factor in 78% of patients. Other risk factors were diabetes mellitus (11 cases,) (48%), high cholesterol level (> 200 mg/dL) or/and low HDL-cholesterol level (10 cases), smoking (10 cases), hyperfibrinogenemia (> 400 mg/dL) (6 cases), high hematocrit (4), and hyperuricemia (> 7.5 mg/dL) (4 cases). The initial manifestations of lacunar syndromes were pure motor hemiparesis (PMH) in 10 patients, sensoy-motor deficit (SM) in 5, ataxic hemiparesis (AH) in 3, dysarthria clumsy hand syndrome (DCH) in 3, and pure sensory syndrome (PS) in 2. Nineteen of the 23 patients presented with completed strokes on arrival to the hospital. Four patients (17%) developed evolving stroke in the subsequent 24 hours of stroke onset. The frequencies of abnormalities found on brain CT and MRI were shown in **Table 1**. Hyperintense lesions on T2WI and PDI of brain MRI included the lacunes and etat cribles, were most commonly encountered at the corona radiata or periventricular white matter (PVWM), followed by internal

Table 1. Frequencies of abnormalities found on brain CT and MRI of 23 patients with "first-ever" lacunar infarction.

	Number (%) of Patients	
Detectable abnormalities	Brain CT	Brain MRI
Corresponding single lesion of acute infarction	18 (78)	23 (100)
Additional lesions	10 (43)	19 (83)*
Leuko-araiosis	2 (9)	14 (61)*

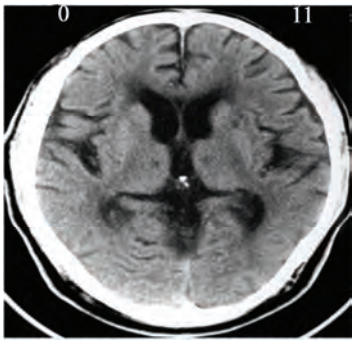
* $p < 0.05$

capsule, basal ganglia, thalamus, and brain stem (**Table 2** and **Figure 1**). Severity of leuko-araiosis in different subgroups was shown in **Table 3**. Regardless of the types of lacunar syndromes, patients with hypertension or diabetes mellitus were likely having moderate to severe LA (grade 2 to 3) (13 of the 18 patients) (72%, in contrast to 20% for those without these risk factors). Combination of hypertension and DM strengthened this effect from severe to advanced LA (grade 3 to 4). None of patient with normotension and euglycemia had LA > grade 2 ($p < 0.05$).

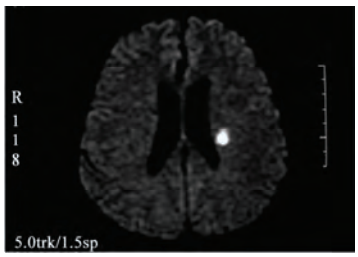
Table 2. Frequency and locations of additional silent lacunes and etat crible in 23 patients with various types of acute first-ever brain lacunar infarctions.

Location of Silent Number (%) of Patients with Positive CT/MRI					
Lacunes	PMH (10)	SM (5)	AH (3)	DCH (3)	PS (2)
Corona radiata*	4 / 7	2 / 3	0 / 1	0 / 2	1 / 2
Internal capsule	3 / 5	2 / 2	1 / 2	2 / 3	0 / 1
Basal ganglia	1 / 4	0 / 2	1 / 1	0 / 0	0 / 0
Thalamus	0 / 2	2 / 3	0 / 1	1 / 1	1 / 1
Brainstem	1 / 3	0 / 2	1 / 2	0 / 1	0 / 0

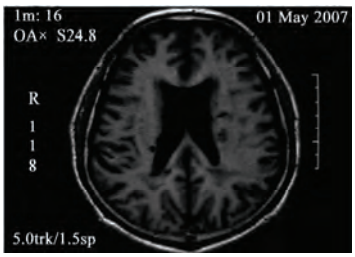
* $p < 0.05$



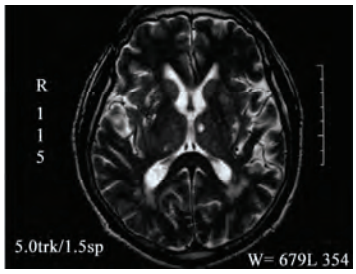
(a)



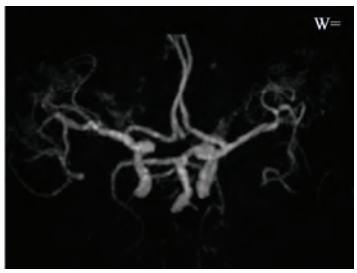
(b)



(c)



(d)



(e)

Figure 1. (a) Brain CT of an acute lacunar infarction involving the left thalamus, and old silent lacunes at left anterior basal ganglion, and right thalamus; (b-e) Brain MRI of another patient with acute lacunar infarction involving the left internal capsule; (b) Diffusion-weighted images (DWI) [TR 3,600.0/TE 84.0]; (c) T1WI; (d) T2WI show additional multiple old "silent brain infarctions" involving bilateral paraventricular white matter/corona radiata, basal ganglia, and thalami. MR angiogram; (e) shows mild atherosclerotic change with segmental narrowing of the anterior and middle cerebral arteries (white arrows); the internal carotid arteries and basilar artery are relatively normal.

4. DISCUSSION

Previous association studies on various risk factors such as hypertension, diabetes mellitus, cardiovascular disease, atrial fibrillation, carotid stenosis, and carotid intima-media thickness, and MRI-diagnosis of silent brain infarctions (SBI) were contradictory [4,5,9]. Regional cerebral blood flow in bilateral frontotemporal regions of brain measured by Xenon-133 inhalation method has been shown to reduce remarkably in neurologically normal adults with brain MRI evidence of silent lacunes [9]. But only less than half of the silent lacunes were seen on T1WI [9]. The autopsied data showed higher (81 %) frequency of SBI [3]. Recent advance in MRI technology improves the detection rate of SBI [4]. SBI was found more frequently than expected in subjects with no history of transient ischemic attack or stroke [4]. Hypertension was present in 72% of subjects with SBI [4,6] that was in accord with our study. Prevalence of SBI even steeply increases after age of 65 years [4,6-8]. However, hypercholesterolemia, hyperglycemia (fasting

plasma glucose ≥ 110 mg/dL), and high hematocrit (Hct ≥ 50%) or hypercoagulability were not risk factors of silent lacunes in several studies, but they were reported to be risk factors in another study [8,15]. Low-grade inflammation of cerebral vessel can be another risk factor of SBI or lacunes [16]. Among 40 patients with lacunes, 30% of cases showed single lesion on MRI [17] and these patients may have neuropsychological abnormality in certain aspect of cognitive domain, even though cognitive function in general remains normal [17]. Our data showed that simple neuropsychological tests such as MMSE, MMT, and Hachinski's score remained normal in patients with first-ever lacune, regardless whether they had or had not SBI [10-12]. A recent study working on more sophisticated neuropsychological tests performed at 1 month after acute lacunar infarction showed that 57.5% of patients had mild cognitive impairment (MCI) that was reflected on lower MMSE score (mean 28.4) [17]. MMSE score of 28 remains at the normal range of our healthy controls [11]. Furthermore, Grau and colleagues [17] included subjects with so called "atypical lacunes" in their study, and this part constituted more than a half of silent brain lacunes located in the striatothalamic region of the brain [17]. Atypical lacunes and pure motor hemiparesis accounted for most of the subjects with MCI [17]. Thalamic lesion constituted 44% of PS, 25% of PMH, and 25% of atypical lacune [17]. Nothing was known about unilaterality or bilaterality of the thalamic lesion and exact locations of the involvement of the different thalamic nuclei. A "typical thalamic lacune" usually located at the lateral tier of thalamic nucleus and frequently manifested with PS. However, a medially located thalamic infarction can be as small as a lacune, either unilateral or bilateral, and sometimes may be mistakenly considered as a "thalamic lacune" or an "atypical lacune". They are actually an ischemic stroke involving the paramedian mesencephalic arteries (PMAS) rather than the deep penetrating arteries of lacunar infarction [18]. PMAS is a branched or a trunk disease that frequently manifests itself with cognitive impairment, neurobehavioral dysfunction, and ocular signs [18]. It is a variant of the top-of-basilar artery syndrome. Therefore, more delicate test for various domains of cognition and higher brain functions might be needed to detect trivial abnormality in patients with lacunar stroke.

In addition, severe periventricular white matter (WM) change with higher grade (≥ 2) of leuko-araiosis correlates with the presence of hypertension (13/18 patients, in contrast to 1/5 in subjects with normal blood pressure) (**Table 3**). This is consistent with the previous studies [4-7]. Austrian Stroke Prevention Study (ASPS) looking at a community-based population on elderly normal subjects over 3-year follow-up [5] have shown that diastolic blood pressure and early-confluent white matter hyperintensities (WMH) at baseline MRI predicted best the white matter hyperintensity progression. This was also confirmed by a 6-year ASPS follow-up study [19]. Moreover, the WMH has been shown to associate with lower forced expiratory volume in 1 second (FEV1) and lower income of less than $ 50,000 per year in CHS [6]. WMH progression did not associate with cognitive functioning [5], punctuate WM lesions are considered benign and not progressive, whereas early-confluent WM abnormalities are progressive within 6-year follow-up [19]. This is in contradiction to the CHS study that concluded that WM findings were associated with impaired cognitive and lower extremity functions [6].

Molecular markers of coagulation activation such as elevated prothrombin fragment, parameters for endothelial cell damage (thrombomodulin and von Willebrand factor), and lipoprotein A were significantly higher in patients with silent lacunes [15]. By contrast, a recently discovered plasma biomarker, SCUBE1, have been shown to associate with different forms of acute cerebral infarction and severity of ischemic stroke based on NIHSS score. The plasma level of SCUBE1 is remarkably

Table 3. Grading of leuko-araiosis (la) in 23 patients with "first-ever" lacunar infarction, high or normal blood pressure, and high or normal fasting plasma glucose.

Grade						Number (%) of Patients with			
Lacunar Syndromes and Positive CT / MRI						Associated Risk Factors			
LA*	PMH (10)	SM (5)	AH(3)	DCH (3)	PS (2)	HBP(18)	NBP(5)	HPG(11)	NPG(12)
0 **	7 / 1	2 / 1	2 / 1	1 / 0	1 / 0	0 **	3	0 **	3
1	1 / 2	1 / 1	1 / 2	1 / 1	0 / 0	5	1	1	5
2	0 / 3	0 / 1	0 / 0	0 / 0	0 / 1	4	1	2	3
3	1 / 3	1 / 1	0 / 0	1 / 2	1 / 1	7	0 **	6	1**
4	1 / 1	1 / 1	0 / 0	0 / 0	0 / 0	2	0 **	2	0 **

*LA: leuko-araiosis; PMH: pure motor hemiparesis; SM: sensorimotor syndrome; AH: ataxic hemiparesis; DCH: dysarthria-clumsy-hand syndrome; PS: pure sensory stroke; HBP: high blood pressure; NBP: normal blood pressure; HPG: high fasting plasma glucose; NPG: normal fasting plasma glucose.
grade 0: absent of LA
grade 1: mild, LA confined to the frontal poles of lateral ventricles ("frontal caps")
grade 2: moderate, LA confined to the frontal and occipital poles of lateral ventricles ("dual caps")
grade 3: severe, LA involved the whole lateral ventricles ("dual caps" plus "rims")
grade 4: advanced, confluent LA lesions surround the centrum semiovale
** $p < 0.05$

elevated in patients with acute infarction due to a trunkor branched-vessel disease, and slightly elevated or normal in patients with acute lacunar infarction [20]. Plasma level of oxidized small dense low-density lipoprotein (LDL-3) is an independent predictor of silent lacunar infarction [21].

In conclusion, evolving-stroke occurs in about one-fifth of patients with "first-ever" acute lacunar infarction within the first 24 hours of stroke onset. Eighty-three percent of Chinese/Taiwanese patients with first-ever lacunar stroke had additional silent brain infarction on MRI. The presence of both hypertension and diabetes mellitus predicts the likelihood of high grade leukoaraiosis and additional SBI. Use of neuroimage studies in conjunction with scrutinized clinical assessments, monitoring stroke risk factors, and measurement of certain molecular or plasma biomarkers may be a future strategy for early diagnosis and treatment of patients at risk for silent brain lacunar infarctions.

5. ACKNOWLEDGEMENTS

The authors thank clinicians and technicians that contributed to the performance of brain CT and MRI examinations. Prof Rossane Harrigan is supported by a grant #1R25RR019321 for Clinical Research Education and Career Development (CRECD) in Minority Institutions; and Prof Linda Chang is supported by a grant #IU54 NS56883-03 for Imaging Studies in Neurotoxicity and Neurodevelopment, University of Hawaii Clinical Specialized Neuroscience Research Program (SNRP).

Conflict of interest: none to declare.

REFERENCES

[1] Fisher, C.M. (1982) Lacunar strokes and infarcts: A review. *Neurology*, **32(8)**, 871.

[2] Ropper, A.H. and Brown, R.H. (2005) Cerebrovascular diseases. In: *Adams and Victor's Principles of Neurology*, 8th Edition, McGraw-Hill Co., Inc., New York, **34**, 682-684.

[3] Tuszynski, M.H., Petito, C.K. and Levy, D.E. (1989) Risk factors and clinical manifestations of pathologically verified lacunar infarctions. *Stroke*, **20(8)**, 990-999.

[4] Vermeer, S.E., Longstretch, W.I. and Kaudstaal, P.J. (2007) Silent brain infarcts: A systemic review. *Lancet Neurology*, **6(7)**, 611-619.

[5] Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H. and Hartung, H.P. (1999) MRI white matter hyperintensities: Three-year follow-up of the Austrian stroke prevention study. *Neurology*, **53(1)**, 132-139.

[6] Longstreth, W.T., Manolio, T.A., Arnold, A., Burke, G.L., Bryan, N., Jungreis, C.A., *et al*. (1996) Clinical correlates of white matter findings on cranial magnetic resonance imaging of 3301 elderly people: The cardiovascular health study. *Stroke*, **27(8)**, 1274-1282.

[7] Vermeer, S.E., Koudstaal, P.J., Oudkerk, M., Hofman, A. and Breteller, M.M.B. (2002) Prevalence and risk factors of silent brain infarcts in the population-based Rotterdam scan study. *Stroke*, **33(1)**, 21-25.

[8] Kwon, H.M., Kim, B.J., Lee, S.H., Choi, S.H., Oh, B.H. and Yoon, B.W. (2006) Metabolic syndrome as an independent risk factor of silent brain infarction in healthy people. *Stroke*, **37(2)**, 466-470.

[9] Kobayashi, S., Okada, K. and Yamashita, K. (1991) Incidence of silent lacunar lesion in normal adults and its relation to cerebral blood flow and risk factors. *Stroke*, **22(4)**, 1379-1383.

[10] Folstein, M.F., Folstein, S.E. and Mc Hugh, P.R. (1975) Mini mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12(3)**, 189-198.

[11] Thajeb, P., Thajeb, T. and Dai, D.F. (2007) Cross-cultural study using modified mini mental test (MMT) for healthy subjects and patients with various forms of vascular dementia. *Journal of Clinical Neuroscience*, **14(1)**, 236-241.

[12] Hachinski, V.C., Lassen, N.H. and Narshall, J. (1974) Multi-infarct dementia, a cause of mental deterioration in the elderly. *Lancet*, **2(7874)**, 207-210.

[13] Poirier, J. and Derouesne, C. (1985) The concept of cerebral lacunae from 1838 to the present (article in French, abstract in English). *Review Neurology* (*Paris*), **141(1)**, 3-17.

[14] Leifer, D., Buonanno. F.S. and Richardson, E.P.Jr. (1990) Clinicopathologic correlations of cranial magnetic resonance imaging of periventricular white matter. *Neurology*, **40(6)**, 911-918.

[15] Kario, K., Matsuo, T., Kobayashi, H., Asada, R. and Matsuo, M. (1996) Silent cerebral infaction is associated with hypercoagulability, endothelial cell damage, and high Lp(a) levels in elderly Japanese. *Arteriosclerosis, Thrombosis, and Vascular Biology*, **16(6)**, 734-741.

[16] Ishikawa, J., Tamura, Y., Hoshide, S., Eguchi, K., Ishikawa, S., Shimada, K. and Kario, K. (2007) Low-grade inflammation is a risk factor for clinical stroke events in addition to silent cerebral infarcts in Japanese older hypertensives: The Jichi medical school ABPM study, wave 1. *Stroke*, **38(3)**, 911-917.

[17] Grau-Olivares, M., Arboix, A., Bartre-Faz, D. and Junque, C. (2007) Neuropsychological abnormalities associated with lacunar infarction. *Journal of Neurological Science*, **257(1)**, 160-165.

[18] Thajeb, P., Thajeb, T. and Dai, D.F. (2007) Clinical analysis of first-ever acute ischemic stroke in the territory of paramedian mesencephalic arteries. *Central European Journal of Medicine*, **2(1)**, 37-46.

[19] Schmidt, R., Enzinger, C., Ropele, S., Schmidt, H. and Fazekas, F. (2003) Progression of cerebral white matter lesions: 6-year results of the Austrian stroke prevention study. *Lancet*, **361(9374)**, 2046-2048.

[20] Dai, D.F., Thajeb, P., Tu, C.F., Chiang, F.T., Chen, C.H., Yang, R.B. and Chen, J.J. (2008) Plasma concentration of SCUBE1, a novel platelet protein, is elevated in patients with acute coronary syndrome and ischemic stroke. *Journal of the American College of Cardiology*, **51(22)**, 2173-2180.

[21] Kato, T., Inoue, T., Yamagishi, S., Morooka, T., Okimoto, T. and Node, K. (2006) Low-density lipoprotein subfractions and the prevalence of silent lacunar infarction in subjects with essential hypertension. *Hypertension Research*, **29(5)**, 303-307.

J. Biomedical Science and Engineering, 2010, 3, 448-453 **JBiSE**
doi:10.4236/jbise.2010.35062 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

The effect of moment arm length on high angled femoral neck fractures (Pauwels' III)

Matthew S. LePine[1], William R. Barfield[1,3], John D. DesJardins[2], Langdon A. Hartsock[1]

[1]Department of Orthopaedic Surgery, Medical University of South Carolina, Charleston, SC;
[2]Department of Bioengineering, Clemson University, Clemson, SC;
[3]Health & Human Performance, College of Charleston, Charleston, SC.
Email: lepine@musc.edu; jdesjar@clemson.edu; hartsock@musc.edu; barfielb@musc.edu

Received 21 January 2010; revised 6 February 2010; accepted 8 February 2010.

ABSTRACT

This study investigated loads among five fixation types (FT) [three cannulated screws (CS), dynamic hip screws with and without derotational screws (DHS-DS and DHS), and dynamic helical hip screws with and without derotational screws (DHHS-DS and DHHS)] across three fracture moment lengths (ML) in Pauwels' Type III fractures. Methods: Seventy-five sawbones were tested (5 FT × 5 trials × 3 ML). The study hypothesis was that significant differences in axial loading to failure would be demonstrated when CS was compared with the other four FT at the three MLs. Each construct was exposed to an axial compressive load to failure. Construct failure was defined as 5 mm of migration at the fracture site or fixation failure. Shapiro-Wilk was used to test for data normality. Subsequently, independent t-tests with Bonferroni correction was used for paired comparisons. Results: At fracture Moments A and B there were no statistical differences between CS and the other FT. At fracture Moment C all four FT yielded significantly higher ($p \leq 0.001$) loads compared with CS. Conclusions: for basicervical fractures CS is a suboptimal form of fixation compared with DHS and DHHS both with and without derotation screws.

Keywords: Pauwels' Fracture, Biomechanical Testing, Fixation

1. INTRODUCTION

Pauwels' Type III fracture is described as a fracture angle of the femoral neck and shaft greater than 50º in the coronal plane, [1-5] with dominating shear stress and varus loading and less compressive loading and stress [2,3]. As a consequence, internal fixation of these unstable fractures surgically has been associated with high rates of non unions [3,6] sparking debate over the most effective internal fixation technique [7-10]. Biomechanical and clinical studies have shown noncomminuted femoral neck fractures in the subcapital and transcervical regions fixed with three cannulated screws yield optimal stabilization results, [11-13] while fractures in the basicervical region show best results when fixed with a sliding hip screw [13,14].

Other variables contribute to the nonunion rates of fixed femoral neck fractures including; femoral neck fracture type, vascular insufficiency, and inaccurate reduction [15]. Stankewich *et al.* expanded this list to include femoral head bone density, percent comminution of the fracture surface, moment arm length, and orientation angle of the fracture surface relative to the femoral shaft [8].

Although biomechanical tests have compared the effectiveness of various fixation devices for different fractures, [16] no study has sought to make a direct comparison based on moment arm length while maintaining a constant fracture angle.

The purpose of this study was to compare the failure load of four different methods of fixation; Dynamic Hip Screws with (DHS-DS) and without derotational screws (DHS), and Dynamic Helical Hip Screw with (DHHS-DS) and without derotational screws (DHHS) (Synthes USA, West Chester, PA) for Pauwels' Type III fracture across three different fracture moment lengths with three cannulated screws (CS), the traditional "gold standard" [17-19]. The study hypothesis was that no statistically significant differences in resistance to axial loading to failure would be demonstrated when the CS was compared with the other four fixation types at each of the three moment lengths.

2. MATERIALS AND METHODS

Seventy-five new 3rd Generation Composite left femoral sawbones (Sawbones Worldwide-Pacific Research Laboratories, Inc. Vashon, WA) of equal density was utilized for this experiment. The 3rd Generation Composite Bone model is designed to simulate natural cortical bone through use of glass fibers and epoxy resin pressure injected around a foam core. Density of the sawbones was

1.7 g/cc with 90 MPa of tensile strength and a modulus of 12.4 GPa. The compressive strength was 120 MPa and the modulus was 7.6 GPa. Other investigators [20,21] have found that 3rd Generation Composite femurs fall within the normal range for bending, torsional and axial loading compared with both fresh-frozen and dried rehydrated cadaver bone. Sawbones also show less variability compared with cadaveric specimens, with the interfemur variability of composites being 20-200 times lower than that for cadaveric specimens, thereby allowing for smaller differences to be characterized with small sample sizes [20,21]. Our apriori power analysis supported use of 5 samples/fixation/moment arm length. Five sawbones for each fracture–fixation and moment arm length were performed, resulting in a total of 75 constructs being tested (5 fixation types × 5 sawbones/fixation type × 3 moment arm lengths).

The three groups of sawbone fractures were created, based on the fracture moment arm length (**Figure 1**; A = 2.0 cm, B = 3.1 cm, C = 4.2 cm). The moment arm distance was measured from the femoral head center to the center of the fracture line on the sawbone specimen using Vernier calipers as shown in **Figure 1**. A custom cutting guide was engineered to ensure fracture line reproducibility.

A femoral neck osteotomy at 70° to the horizontal (**Figure 1**), was created by a single investigator with a bandsaw in each of the sawbones, extending from the superior femoral neck to subcapital, transcervical and basicervical regions. All fractures were repaired by a single investigator using one of five forms of fixation.

The three CSs (7.3 cannulated screw 100 mm, 16 mm thread–Synthes, Paoli, PA, USA Product # 208.900) were placed in an inverted triangle configuration at a 135°. A cannulated screw set jig supplied by Synthes was used for screw insertion to reduce variability.

The four other fixation types included: 1) DHS-DS (4 hole 135° Std plate-Synthes Product # 281.140, 90 mm compression lag screw–Synthes Product # 280.900) with a derotational screw (7.3 cannulated screw 100 mm, 16 mm thread–Synthes Product # 208.900); 2) DHS (4 hole 135° Std plate–Synthes Product # 281.140, 90mm screw–Synthes Product # 280.900) without a derotational screw; 3) DHHS (4 hole 135° Std LCP-Plate–Synthes Product # 282.612, 100 mm helical screw–Synthes Product # 282.239) with a derotational screw (7.3 cannulated screw 100 mm, 16 mm thread–Synthes Product # 208.900); 4) DHHS-DS (4 hole 135° std LCP-Plate–Synthes Product # 282.612, 100 mm helical Screw-Synthes Product # 282.239) without a derotational screw. All hardware was inserted according to the technique guide provide by the manufacturer. The 4-hole plate has traditionally been used for fixation of proximal femur fractures. A PMMA appliance (**Figure 2**) was created which cupped the displaced femoral head and proximal femoral neck to insure reduction accuracy. Implants were properly positioned by direct visualization on the surface of the bone. The hip lag screw guide wire was consistently placed in the center of the femoral head through use of an aiming device.

The angle of implant insertion was in accordance with the technique guide provided by the manufacturer including pre-drilling of all holes prior to fixation (**Figure 3**). A representative group of each fixation type was X-rayed to insure anatomic reduction.

The femurs were uniformly potted in dental stone (Modern Materials) at 25° of adduction in the coronal plane and neutral position in the sagittal plane to simulate the terminal stance phase of gait [22-25]. The Instron Mechanical Testing System (Instron Bi-Axial Servo-Hydraulic Testing Machine Model # 8874-Norwood, MA) incrementally loaded the femoral head with an evenly applied compressive axial load through a polished stainless steel cup attached to the ram. Potted femurs were securely locked in place with the appliance pictured in **Figure 4** for testing. The appliance was free to move in the medio-lateral plane.

The loading protocol utilized was from Chang *et al.*, 1987 [25], designed to simulate single leg stance. In this

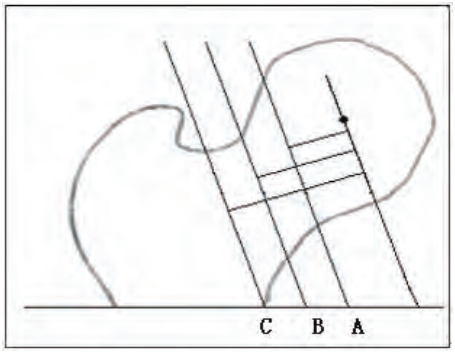


Figure 1. Schematic demonstration of moment arm length (A,B,C). A = Subcapital; B = Transcervical; C = Basicervical.

Figure 2. PMMA appliance used to hold the 2-part Pauwels' fracture in the vise for accurate reduction of fractures.

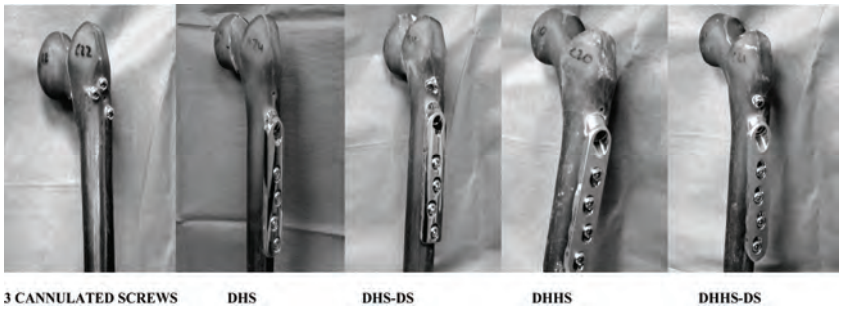


Figure 3. Pictures of the five fixation forms.

Figure 4. Testing jig used for mechanical testing of constructs.

single-leg phase, body weight is transferred ahead of the forefoot, creating loading conditions of nearly five times body weight [26]. Constructs were positioned under a servohydrolic materials testing system producing a vertically-downward force. Displacement was applied at a rate of 0.5 mm/s and terminated when the construct reached failure criteria. Failure was defined as 5 mm of migration at the fracture site through either translation and/or a combination of migration and rotation of the femoral head or catastrophic failure of device or sawbone [13].

Force data were captured using Instron's Max v7.1 computer software at 100 Hz. Statistical analysis was performed using SPSS Version 14 software. Shapiro-Wilk was utilized to test for normality. Unpaired t-tests were subsequently used to compare the four constructs; DHS, DHS-DS, DHHS, and DHHS-DS to the CS constructs at each of the three moment lengths. Due to multiple unpaired comparisons a Bonferroni adjustment was used. Our apriori alpha level was $p \leq 0.05$. With four comparisons the adjusted alpha was lowered to $p \leq 0.013$.

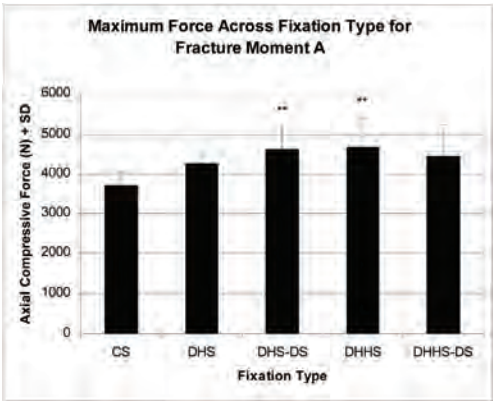
3. RESULTS

Shapiro-Wilk test for normality indicated that the data was normally distributed (> 0.05). At fracture Moment A (subcapital) there was no statistical difference in yield forces ($p = 0.026$) (**Figure 5(a)**). At fracture Moment B (transcervical) there were no statistical differences in yield forces ($p = 0.022$). (**Figure 5(b)**) At fracture Moment C (basicervical) all four fixation forms yielded statistically significant ($p \leq 0.001$) high axial loads when compared with CS as seen in **Figure 5(c)**.

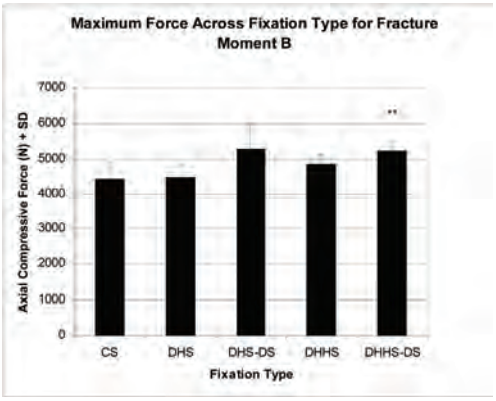
4. DISCUSSION

In the present investigation five fixation types at three different fracture moment lengths (A [subcapital], B [transcervical], C [basicervical]), were axially loaded and biomechanically testing in 75 sawbone models. Our study hypothesis was that no statistically significant differences in resistance to axial loading to failure would be demonstrated when the CS was compared with the other four fixation types at each of the three moment lengths. Results supported our hypotheses at each moment length except the basicervical model where statistically different results were seen

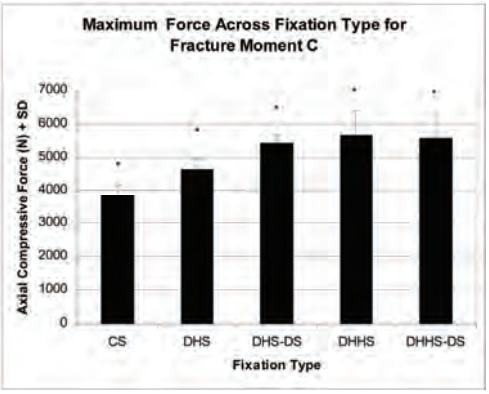
Aminian *et al.* in the *J Orthop Trauma* in 2007, biomechanically tested 32 cadaver specimens utilizing four fixation types. Two of our fixation types were the same as Aminian *et al.* (CS and DHS). Aminian *et al.* specimens were cyclically loaded at 1400 N at 3 Hz for 10,000 cycles followed by a load-to failure test. All CS specimens failed during incremental loading with a failure load of 862 N [16]. Tan *et al.* tested five pairs of fresh frozen cadaver femurs and used cyclical axial loading of 750 N at 0.5 Hz for 200 cycles [27] and found that the load at the yield point was significantly higher in the group with more horizontally oriented screws leading to



(a)



(b)



(c)

Figure 5. Maximum axial force for fracture Moment A** =statistical trend compared with CS.

the suggestion that two horizontal screws in the femoral neck provide better stiffness.

In our study both the CS and DHS construct failed at higher loads ([Moment A-3709.9 N] [Moment B-4423.4 N] [Moment C-3830.8 N]), which we partially attribute to the lack of cyclical fatigue testing, specimen differences (cadaver bone vs. sawbones of uniform density) and different mechanical testing based on our moment arm length differences and different fixation types. The moment arm distance, measured from the center of the femoral head to the fracture line, plays a key role. As the distance increases the torque across the fracture line increases. The authors of the present work evaluated the current English literature through traditional search techniques and could not locate any studies that have investigated the mechanical properties of various fixation techniques in Pauwels' Fracture Type III with moment length considered.

Our testing protocol from Chang *et al*., 2002 [25], best approximated the study goal of assessing maximal axial compressive loading leading to eventual failure among the fixation types, although other protocols exist [27-30]. The authors recognized through earlier unpublished pilot data that there would be minimal flexure of the sawbone, similar to what was observed in Aminian *et al.*

Our fracture moments results supported the null hypothesis are at the subcapital and transcervical region, but not at the basicervical region.

The fracture Moment C model yielded statistically significant differences ($p < 0.001$) between CS and each of the other four fixation types. Our results, based on the use of synthetic femurs, suggest that for a subcapital high angle fracture (Moment A) the DHHS and DHS-DS may be better clinical choices for fixation than the standard CS and at the intermediate fracture Moment B the DHHS-DS would be the preferred form of fixation although statistically significant differences did not exist. At fracture Moment C, a high angle basicervical fracture, the study results suggest that all of the fixed angle fixation forms that we tested are superior to CS in combating axial load. At Moment C, there is a shorter amount of screw length supported by the intact femur and when a force is applied the screw is not supported by adjacent cortical bone due to the width at the base of the femoral neck. Fixation with CS can lead to greater displacement under the applied load.

The literature is inconclusive with respect to DHS fixation. Baitner *et al.* demonstrated that the DHS construct was significantly stronger than the cannulated screw construct [13], however Husby *et al*., found no significant differences between the same two constructs [10].

In a normal gait cycle at heel strike and toe off, the hip experiences approximately 4-5 times body weight [24]. In the instance of a 100 kg (981 N) individual walking, the hip joint can experience between 3924-4905 N. These numbers approximate the values at failure found in our investigation. At fracture Moments A, B, and C the loads for all fixation types are at or exceed the upper range of axial loading compared with the example provided for normal gait. Obviously, an *in vitro* study and what occurs *in vitro* is much different due to soft tissue damping which occurs in humans as we independently ambulate. However, our biomechanical results

suggest that fixation type is a function of the moment arm length in Pauwels; Type III fracture.

The US population is getting older and progressively more obese. A 120 kg (1177 N) patient can experience 4708-5885 N in the hip which surpasses the load bearing capacity of several of the fixation forms tested independent of moment arm length. In the event of a stumble, forces in the hip joint approach eight times body weight [24]. The result of this could be catastrophic based on the results demonstrated by this laboratory study.

The results found in this project are clinically relevant because we have demonstrated that significant differences do exist across moment lengths and fixation types in Pauwels Type III fractures. Even with adequate reductions, failure rates are higher with cannulated screws than with fixed angle devices [7]. Probe and Ward [31] reported that both fixed angle and multiple screws with divergent paths were superior to parallel screws. The current work supports the prior body of knowledge associated with the difficult clinical issue of surgical repair of Pauwels' Type III fractures.

Strengths of our study include: bone uniformity, fracture accuracy, a single investigators placement of screws and other types of fixation, and uniformity of testing and recording of data. The weaknesses of our investigation include uniaxial loading. One perceived weakness among some readers may be that we used sawbones as opposed to fresh cadaver femurs. While this perspective has merit, our position is that by using a material that possesses uniformity, our protocol measured the axial loading of the constructs and was not dependent on human bone quality in terms of density. Varying bone densities can introduce a confounding variable, which we chose to control in our work. This position is supported by studies reported in 1996 and 2001 using sawbones and comparing the mechanical results with human cadaver bone. The findings in these studies support the structural equivalent of composite bones with natural bones in axial, bending and torsional loading [20,21]. Our loading protocol at failure approximated load during single leg stance. Because of hip musculature stabilization our fixation model does not completely replicate *in situ* loading. Different directional force vectors can negatively affect fixation. Static loading in a uniaxial direction does not mimic the complex tensile, compressive and shear loads across the hip *in situ*.

This study is clinically relevant in that for subcapital high angle fractures differences existed with the DHHS and DHS-DS being better fixation choices when compared with the CS construct. For transcervical fractures the best choice was DHHS-DS and for high angle basicervical fractures DHS, DHS-DS, DHHS, and DHHS-DS were all axially superior to CS for fracture fixation. However, as noted in a recent clinical paper that evaluated the efficacy of internal fixation of Pauwels' Type III femoral neck fractures, despite advances clinically and biomechanically the ideal fixation form remains unclear. Clearly patient bone quality, patient age and implant position all affect fixation results in patients as noted in a recent clinical paper [32].

Finally, our model was used purely as a model for comparison of stability of individual types of internal fixation. Our *in vitro* study results suggest that surgical fixation for Pauwels' Type III unstable fractures is a function of the moment arm distance and loading magnitude.

REFERENCES

[1] Nilsson, L., Johansson, A. and Stromqvist, B. (1993) Factors predicting healing complications in femoral neck fractures: 138 patients followed for 2 years. *Acta Orthopaedica Scandinavica*, **64(2)**, 175-177.

[2] Parker, M.J. and Dynan, Y. (1998) Is Pauwels' classification still valid? *Injury*, **29(7)**, 521-523.

[3] Bartonicek, J. (2001) Pauwels' classification of femoral neck fractures: Correct interpretation of the original. *Journal of Orthopaedic Trauma*, **15(5)**, 358-360.

[4] Pauwels, F. (1976) Biomechanics of the normal and diseased hip. Springer, Berlin.

[5] Caviglia, H.A., Osorio, P.Q. and Comando D. (2002) Classification and diagnosis of intracapsular fractures of the proximal femur. *Clinical Orthopaedics and Related Research*, **399(7)**, 17-27.

[6] Marti, R.K., Schuller, H.M. and Raaymakers, E.L. (1989) Intertrochanteric osteotomy for non-union of the femoral neck. *Journal of Bone and Joint Surgery*, **71(5)**, 782-787.

[7] Haidukewych, G.J., Liporace, F. and Gaines, R. (2005) Pauwels' type 3 vertical femoral neck fractures – What is the best fixation device? *Orthopaedic Trauma Association* 21*st Annual Meeting*, **21(8)**, 2005, 130-131.

[8] Stankewich, C.J., Chapman, J., Muthusamy, R., Quaid, G., Schemitsch, E., Tencer, A. and Ching, R.P. (1996) Relationship of mechanical factors to the strength of proximal femur fractures fixed with cancellous screws. *Journal of Orthopaedic Trauma*, **10(4)**, 248-257.

[9] Weinrobe, M., Stankewich, C.J., Mueller, B., Tencer, A.F. (1998) Predicting the mechanical outcome of femoral neck fractures fixed with cancellous screw: An in vivo study. *Journal of Orthopaedic Trauma*, **12(1)**, 27-37.

[10] Husby, T., Alho, A., Hoiseth, A. and Fonstelien, E. (1987) Strength of femoral neck fixation, comparison of six techniques in cadavers. *Acta Orthopaedica Scandinavica*, **58(6)**, 634-637.

[11] Barnes, R., Brown, J.T., Garden, R.S. and Nicoll, E. (1976) Al subcapital fractures of the femur. *Journal of Bone and Joint Surgery*, **58B(1)**, 2-24.

[12] Bout, C.A., Cannegieter, D.M. and Juttman, J.W. (1997) Percutaneous cannulated screw fixation of femoral neck fractures: The three point principle. *Injury*, **28(2)**, 135-139.

[13] Baitner, A.C., Maurer, S.G., Hickey, D.G., Jazrawi, L.M., Kummer, F.J., Goldman, S. and Koval, K.J. (1999) Vertical shear fractures of the femoral neck. *Clinical Orthopaedics & Related Research*, **367(65)**, 300-305.

[14] Deneka, D.A., Simonian, P., Stankewich, E.D., Chapman, J.R. and Tencer, A.F. (1997) Biomechanical comparison

of internal fixation techniques for the treatment of unstable basicervical femoral neck fractures. *Journal of Orthopaedic Trauma*, **11(2)**, 337-343.

[15] LaVelle, D.G. (2003) Delayed union and nonunion of fractures. In: Canale, S.T., Ed., *Campbell's Operative Orthopaedics*, St. Louis, **(10)3**, 3125-3165.

[16] Aminian, A., Gao, F., Fedoriw, W.W., Zhang, L.Q., Kalainov, D.M. and Mark, B.R. (2007) Vertically oriented femoral neck fractures:mechanical analysis of four fixation techniques. *Journal of Orthopaedic Trauma*, **21(8)**, 544-548.

[17] Hernefaulk, L. and Messner, K. (1995) Femoral stiffness after osteosynthesis of a subcapital osteotomy in osteoporotic bone: An in vitro comparison of nine fixation methods. *Journal of Bone and Joint Surgery*, **9(58B)**, 2-24, 464-469.

[18] Kuokkanen, H., Korkala, O., Antti-Poika, I., Tolonen, J., Lehtimaki, M.Y. and Silvennoinen, T. (1991) Three cancellous bone screws versus a screw-angle plate in the treatment of Garden I & II fractures of the femoral neck. *Acta Orthopaedica Belgium*, **57(1)**, 53-57.

[19] Parker, M.J. and Blundell, C. (1998) Choice of implant for internal fixation of femoral neck fractures: Meta-analysis of 25 randomized trials including 4,925 patients. *Acta Orthopaedica Scandinavica*, **69(2)**, 138-143.

[20] Christofolini, L., Vineconti, M., Cappello, A. and Toni, A. (1996) Mechanical validation of whole bone composite femur models. *Journal of Biomechanics*, **29(4)**, 525-535.

[21] Heiner, A.D. and Brown, T.D. (2001) Structural properties of a new design of composite replicate femurs and tibias. *Journal of Biomechanics*, **34(6)**, 773-781.

[22] Choueka, J., Koval, K.J., Kummer, F.J., Crawford, G. and Zuckerman, J.D. (1995) Biomechanical comparison of the sliding hip screw and the dome plunger: Effects of material and fixation design. *Journal of Bone and Joint Surgery*, **77-B(2)**, 277-283.

[23] Goh, J.C.H., Shah, K.M. and Bose, K. (1995) Biomechanical study on femoral neck fracture fixation in relation to bone mineral density. *Clinical Biomechanics*, **10(6)**, 304-308.

[24] Kubiak, E.N., Bong, M., Park, S.S., Kummer, F., Egol, K. and Koval, K.J. (2004) Intramedullary fixation of unstable intertrochanteric hip fractures: One or two lag screws. *Journal of Orthopaedic Trauma*, **18(1)**, 12-17.

[25] Chang, W.S., Zuckerman, J.D., Kummer, F.J. and Frankel, V.H. (1987) Biomechanical evaluation of anatomic reduction versus medial displacement osteotomy in unstable intertrochanteric fractures. *Clinical Orthopaedics & Related Research*, **225(6)**, 141-146.

[26] (1986) Gait analysis. In: McAinsh, T.F., Ed., *Physics in Medicine & Biology Encyclopedia: Medical Physics, Bioengineering and Biophysics*, Pergamon Press Inc, New York, **1(5)**, 356.

[27] Tan, V., Wong, K.L., Born, C.T, Harten, R. and DeLong, W.G. (2007) Two-screw femoral neck fracture fixation: A biomechanical analysis of 2 different confirurations. *American Journal of Orthopedics*, **36(9)**, 481-485.

[28] Chaim, S.H., Mukherjee, D.P., Ogden, A.L., Mayeux, R. H., Sadasivan, K.K. and Albright, J.A. (2002) A biomechanical study of femoral neck fracture fixation with the VHS-Vari-Angle hip fixation system. *American Journal of Orthopedics*, **31(Suppl 1)**, 22-24.

[29] Aboulafia, A.J., Price, M.M., Kennon, R.E. and Hutton, W.C. (1999) A comparison of mechanical strength of the femoral neck following locked intramedullary nailing using oblique versus transverse proximal screws. *Journal of Orthopaedic Trauma*, **13(3)**, 160-163.

[30] Kukla, C., Gaebler, C., Pichl, R.W., Prokesch, R., Heinze, G. and Heinz, T. (2002) Predictive geometric factors in a standardized model of femoral neck fractures. Experimental study of cadaveric human femurs. *Injury*, **33**, 427-433.

[31] Probe, R. and Ward, R. (2006) Internal fixation of femoral neck fractures. *Journal of the Academy of Orthopaedic Surgeons*, **14(9)**, 565-571.

[32] Liporace, F., Gaines, R., Collinge, C. and Haidukewych, G.J. (2008) Results of internal fixation of Pauwels' type-3 vertical femoral neck fractures. *Journal of Bone and Joint Surgery*, **90(A)**, 1654-1659.

J. Biomedical Science and Engineering, 2010, 3, 454-458 **JBiSE**
doi: 10.4236/jbise.2010.35063 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Determination of isotretinoin in pharmaceutical formulations by reversed-phase HPLC

Carla Aiolfi Guimarães[1*], Farid Menaa[2], Bouzid Menaa[3*], Ivo Lebrun[4], Joyce S. Quenca-Guillen[1], Aline Vivian Vatti Auada[4], Lucildes P. Mercuri[5], Paula Ferreira[5], Maria Inês Rocha Miritello Santoro[1]

[1]Department of Pharmacy, Faculty of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil;
[2]Department of Dermatology, School of Medicine, University of Wuerzburg, Wuerzburg, Germany;
[3]Fluorotronics, Inc., 1425 Russ. Bvld, San Diego Technology Incubator, San Diego, USA;
[4]Department of Biochemistry and Biophysics, Butantan Institute, São Paulo, Brazil;
[5]Department of Exact and Earth Sciences, Diadema, Federal University of São Paulo, São Paulo, Brazil.
Email: carlaaiolfi@usp.br; bouzid.menaa@gmail.com

Received 4 February 2010; revised 1 March 2010; accepted 2 March 2010.

ABSTRACT

The development of facile and rapid quantification of biologically active biomolecules such as isotretitoin in therapeutic drugs contained in many generic formulations is necessary for determining their efficiency and their quality to improve the human health care. Isotretritoin finds its applications in the maintenance of epithelial tissues. Different processes to date such as normal phase HPLC, or gas chromatrography among others are able to separate and quantify isotetroin. However, the extraction is quite complex and in the case of HPLC, the analysis requires long retention times. In such context, an isocratic reversed-phase high-performance liquid chromatography (HPLC) technique coupled with an UV-vis detector is described here for easy separation and quantification of 13-cis-retinoic (isotretinoin) from soft gelatin capsule formulations. The isotretinoin was extracted from three different commercial drug samples with tetrahydrofuran (THF) solvent by a procedure that can be completed in less than 10 minutes. Subsequent separation and quantification were accomplished in less than 5 minutes under isocratic reversed-phase conditions on a Lichrospher RP18 column and a mobile phase consisting of 0.01% TFA/acetonitrile (15/85, v/v) at a flow rate of 1.0 mL/min. Isotretoin was detected for the three samples via its UV-vis absorbance at 342 nm. The method was validated and the results showed good linearity, precision and accuracy for sensitive and selective quantitative determination of isotretinoin in the different pharmaceutical formulations. We found that the average isotretinoin content in two of the three commercial products fell outside the 90-110% United States Pharmacopeia specifications. Consequently, the facile extraction and the precise method for the biomolecule quantification open up tremendous possibilities in improving the quality control of drugs which can exist as different generic brands.

Keywords: 13-*Cis* Retinoic Acid; Reversed-Phase Chromatography; Isotretitoin Extraction; Isotretinoin Quantification; Pharmaceuticals Formulation

1. INTRODUCTION

Retinoic acid is a very potent biomolecule in promoting growth and controlling differentiation and maintenance of epithelial tissue for vitamin A deficient animals. Indeed, all-*trans*-retinoic acid (tretinoin) appears to be active form of vitamin A in all tissues except the retina, and it is 10-100-fold more potent than retinol in various systems in vitro. Isomerization of this compound in the body yields to 13-*cis*-retinoic acid (isotretinoin), which is nearly as potent as tretinoin in many of its actions on epithelial tissues. Vitamin A and other retinoids have found wide applications in the treatment of skin disorders and may find important roles in cancer chemoprevention and therapy [1].

One of the most clinically useful vitamin A derivatives is 13-*cis*-retinoic acid (isotretinoin). It was approved in 1982 to treat severe recalcitrant nodular acne and remains the drug of choice for this therapeutic treatment. Isotretinoin is the only treatment that affects all the major factors involved in the pathogenesis of acne. Its exhibits several activities, in particular a capacity to decrease sebaceous gland activity, to correct the keratinisation defect in acne and to reduce the population of the bacterium Propionibacterium acnes [2].

Isotretinoin is a strong lipophilic molecule, almost insoluble in water and only partially soluble in oil. It is labile upon heating, light and air. In such conditions, these physicochemical properties impose considerable

limitations upon formulation options. For instance, the degradation of isotretinoin (in its solid state) in the presence of air is an autocatalytic process that proceeds rapidly and a simple dry powder formulation would not have a sufficient shelf-life. A suspension in oil is more stable [2]. Suspension in oil also minimizes chemical degradation, as lipid preparations, especially those with a low peroxide number, have been shown to enhance stability [3]. Owing to the photolability and sensitivity to heat and oxidation of the retinoids, their quantitative determination in pharmaceutical is particularly important for the quality control of finished products and for stability-indicating assays [1].

Indeed, for all pharmaceutical agents, one of the most important therapeutic determinants is the amount of drug the patient is exposed to. Previous studies comparing the quality of a variety of generic with innovator products have revealed important deficiencies in the amount of active ingredient among generics [4,5]. Isotretinoin and its trans-isomer, tretinoine, are known to be teratogens that can cause fetal malformations during pregnancy for intance [6,7]. It is possible for a patient to be treated with a different isotretinoin-based generic products during a course of therapy because of an assumption that 'generic isotretinoins' are essentially identical and bioequivalent to the innovator product. Under these circumstances, the pharmaceutical quality and consistency among products become of paramount importance [2].

A number of previous methods for determination of isotretinoin have been described in the literature. The analytical method generally employed in the United State Pharmacopoeia (USP) is a normal phase HPLC system [8]. Isotretinoin has also been determined by gas chromatography in soft and hard capsules formulations [9]. A reversed phase HPLC method with fluorescence detection has been described [1] for the determination of retinoids in pharmaceutical dosage forms. Simultaneous determination of tretinoin and isotretinoin in dermatological formulations has been also reported via normal phase HPLC [10]. HPLC [1], Microcalorimetry [11], UV radiation monitor [12], thin layer chromatography (TLC) and preparative layer chromatography (PLC) [13], capillary zone electrophoresis (CZE) and micellar electrokinetic capillary chromatography (ME-KC) [14] have been also employed to study retinoid degradation products in commercial preparation and raw materials.

Although the above methods are able to separate and quantify isotretinoin, the extraction procedure is quite complex and HPLC analysis involves long retention times [10]. In that context, the aim of this study was to develop a rapid method using an isocratic high performance liquid chromatography (HPLC) coupled to UV-vis detector to determine and quantify the isotretinoin in soft capsule formulations from three commercial samples. The validation method with its precision and accuracy was checked and the isotretinoin quantification was carried out for three commercial products for which the content of the biomolecule was compared to the United States Pharmacopeia specifications.

2. MATERIALS AND METHODS

2.1. Chemicals and Reagents

The content of isotretinoin was determined from the concentrations of standard solutions. The standards for 13-*cis*-retinoic acid were from United States Pharmcopeia reference standard (USP:RS). The three commercial products were purchased from two different pharmaceutical companies (their names have been omitted for commercial reasons) named as Branded Original Product A (Product A), Generic Product B (Product B) and Generic Product C (Product C). All three commercial soft-gel capsules contained 20 mg of isotretinoin. Acetonitrile, trifluoroacetic acid (TFA), tetrahydrofuran (THF) were HPLC grade and purchased from Merck (USA).

2.2. Laboratory Precaution

All experiments were performed using amber glass volumetric flasks under yellow light conditions to avoid degradation of isotretinoin.

2.3. HPLC Instrumentation and Conditions

The HPLC system consisting of Agilent model 1100 (Agilent, USA) connected to a UV/Visible absorbance detector (agilent) was used for all separation. Chromatographic separations were performed using reversed-phase chromatographic using a HPLC column [Lichrospher RP18 (5 µm, 125 × 4 mm i.d., Merck, USA)]. Isotretinoin was detected by the UV-vis absorbance at 342 nm [8] with the sensitivity set at 0.1 aborbance unit full scale (Aufs).

The mobile phase used was composed of 0.01% trifluoroacetic acid and acetonitrile (15/85, v/v) at a flow rate of 1.0 mL/min. The injection volume was 20 µL and the column was maintained at 40ºC.

2.4. Standard Preparation Procedures

Stock solutions were prepared by dissolving the appropriate amounts of isotretinoin in acetonitrile. A set of working standard solutions was produced by diluting aliquots of the stock solutions to give the desired concentrations of the analytes. For the method linearity assessment, the concentration range was 5.0 to 30.0 µg/mL (the concentrations were determined using the calibration graphs. The standard solutions were stable for at least 3 days at 4ºC).

2.5. Extraction of Soft-Gel Capsules Formulations

Isotretinoin capsules were opened carefully using a sharp blade and the content of these capsules was ex-

tracted directly with THF solvent. The solution containing isotretinoin was shaked for 5 min. 1 mL of aliquot for the resulting solution was diluted in acetonitrile to obtain the final solution concentration (20 µg/mL) that was analyzed by HPLC in comparison with the appropriate standard solution.

2.6. Quantification of Isotretinoin

A standard curve was constructed by injecting samples containing isotretinoin standard at concentrations ranging from 5.0 to 30.0 µg/mL. The peak area was determined and plotted versus the concentration of isotretinoin. For the recovery studies, known volumes of isotretinoin standard solutions were analyzed, and the absolute recovery was calculated by comparing the peak area obtained from isotretinoin in the commercial capsule with the peak area of samples derived from the standard solutions.

3. RESULTS AND DISCUSSION

The objective of this work was to develop a rapid method using an isocratic high performance liquid chromatography (HPLC) system to determine quantitatively the isotretinoin in hard gelatin capsule formulations. In order to validate an efficient method for analysis of drugs in pharmaceutical formulations, preliminary tests were performed with the objective to select adequate and optimum conditions. Parameters such as detection wavelength, ideal mobile phase and their proportions, concentration of the standard solutions, and flow rate were exhaustively tested and the quantitative determination of isotretoin in the gel capsules was made possible under the ideal conditions described in this paper. The proposed method is simple and do not involve laborious time-consuming sample preparation.

3.1. Chromatography

The conditions for a rapid and simple HPLC separation with UV detection were developed using an isocratic elution with a mobile phase composed of 0.01% trifluoroacetic acid and acetonitrile (15/85, v/v). These conditions gave well resolved sharp peaks (**Figure 1**) for

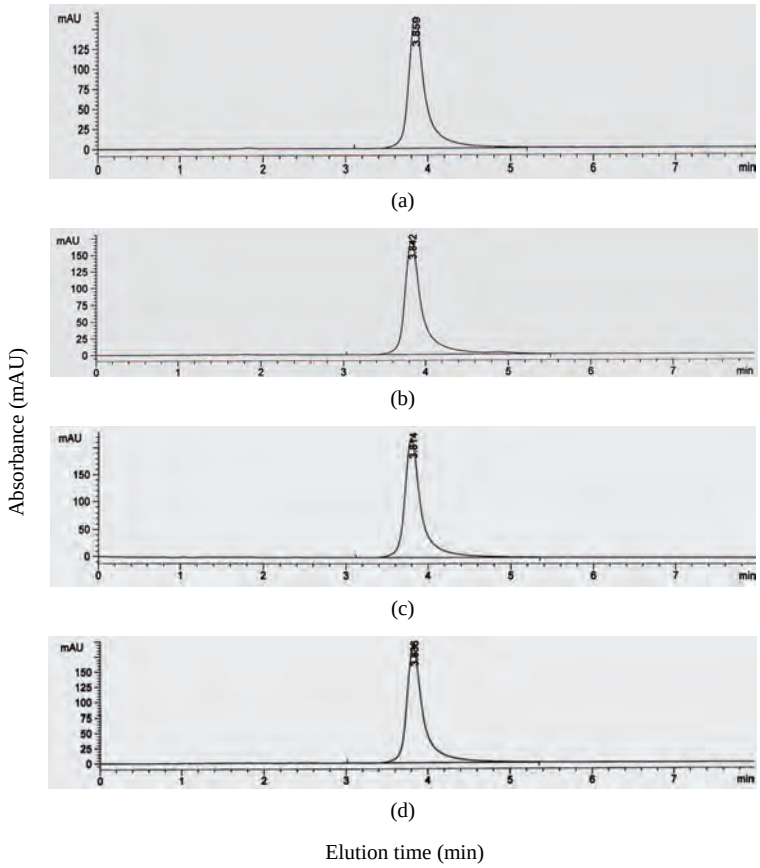


Figure 1. Chromatogram of samples. (a) isotretinoin standard (20 µg/mL); (b) Product A (20 µg/mL); (c) Product B (20 µg/mL); (d) Product C (20 µg/mL). Conditions: Lichrospher RP18 column (5 µm, 125 × 4 mm i.d., Merck), mobile phase 0.01% TFA /acetonitrile (15/85, v/v), flow rate 1.0 mL/min, UV detection at 342 nm and ambient temperature (24 ± 2ºC).

the isotretinoin commercial formulations but also for the isotretinoin standards. The retention time was of approximately 3.8 min. The shorter elution time makes the method especially useful for routine analysis of isotretinoin in pharmaceutical formulations. The validation experiments as described in the following section were completed to determine if the method could achieve the reproducibility and accuracy required for analysis of formulations.

3.2. Method Validation

3.2.1. Method Validation

The quantification of the chromatogram was performed using the peak area of isotretinoin standards for known content of isotretinoin and as references for the determination of the content for the commercial samples to which we determined the peak areas of their respective chromatograms. Five standard solutions ranging from 5.0 to 30.0 µg/mL in concentration (three replicates each), were injected into the HPLC system. The calibration graph was obtained by plotting the peak areas of the standard solutions against the theoretical standard concentrations. The linearity was evaluated by linear least-squares regression analysis. The correlation coefficient obtain with linear regression of curve was 0.997. Statistical analysis indicated excellent linearity as shown in **Table 1**.

3.2.2. Accuracy

To evaluate the accuracy of the proposed method, recovery tests were carried out with all samples. The measurements were performed by adding known amounts of standard solutions to sample followed by analysis using the proposed method. The recovery values obtained were 101.2, 99.4 and 99.3% (why don't you have 100%) for the three soft gelatin capsule formulations, confirming the accuracy of the proposed method. The percentage of recovery results are presented in **Table 1**.

3.2.3. Precision

The precision of proposed method was evaluated through intra-day repeatability of responses after replicate (n = 10) injection of sample solutions (20.0 µg/mL). The precision is expressed as the relative standard deviation (RSD) amongst responses. The standard deviation amongst replicate responses and relative standard deviation values (RSD) were less than 1.0%, indicating precision of the method. The statistical data results obtained in the analysis of commercially available samples are shown in **Table 2**.

According to the results obtained only one formulation (Product A (branded named)), was found to be good agreement with the claimed content of the drugs. Indeed, the average percentage of isotretinoin content (**Table 3**) for the Product A was evaluated to 105 ± 0.05% which fits with United States Pharmacopeia (USP) specifications. In contrast, the content for the generic products formulations Products (Product B and Product C), fell outside the United States Pharmacopeia (USP) specifications. The content percentage of isotretinoin was found to be 128.5 ± 0.08 and 115.5 ± 0.04 in Product B and Product C, respectively. The USP XXX [8] monograph for isotretinoin capsules states that the capsules must contain not less than 90.0% and not more than 110% of the labeled amounts of isotretinoin. The results and the accuracy of the method shows that it is necessary to control the content of the active biomolecule for drug safety, but also that the precision of the method and rapid acquisition of the data can be a plus for pharmaceutical industries for the quality control of their products.

Table 1. Statistical results of linear regression analysis in the determination of isotretinoin by proposed method.

Statistical parametrs	Results
Intercept of curve (b)	–328.09
Slope of curve (a)	138.86
Linear correlation coefficient (r)	0.997

Table 2. Recovery data of standard solutions added to the samples analyzed by using the proposed HPLC method.

Sample	Amount added (µg/mL)	Amount found* (µg/mL)	Recovery(%)	Mean recovery(%)
Product A	25.0	26.0	100.0	101.2
	30.0	31.3	103.0	
	35.0	36.1	100.6	
Product B	25.0	30.6	98.0	99.4
	30.0	35.8	101.0	
	35.0	40.6	99.3	
Product C	25.0	28.2	102.0	99.3
	30.0	32.9	98.0	
	35.0	37.8	98.0	

*Avarage of three determination

Table 3. Statistical data obtained in the analysis of samples using the proposed method.

Sample	Declared theoretical concentration (µg/mL)	Found experimental concentration (µg/mL)	Content (%)	RSD (*) (%)	Content (%) Confidence interval (**)
Product A	20.0	21.0	105.0	0.43	105.0 ± 0.05
Product B	20.0	25.7	128.5	0.70	128.5 ± 0.08
Product C	20.0	23.1	115.5	0.30	115.5 ± 0.04

(*) Average of ten determination (**) 95.0% confidence level.

4. CONCLUSIONS

The proposed HPLC methods for quantitative determination showed good linearity, precision and accuracy for sensitive and selective quantitative determination of isotretinoin in pharmaceutical formulations. This method is not time-consuming and is easy to perform in any laboratory. We ultimately show, according the United States Pharmacopeia specifications that the average amount of isotretinoin in 2 of the 3 commercial products ranged outside the 90-110%. Safe use of isotretinoin, particularly with respect to teratogenicity, remains a central issue for therapeutics development. Therefore, it is important, that future research determines a comparative study with pharmaceutical quality of isotretinoin products on patient exposure to isotretinoin by comparative bioequivalence tests.

5. ACKNOWLEDGEMENTS

The authors acknowledge the CNPq (National Counsel of Technological and Scientific Development) CAPES (Foundation to Higher Level People Improvement) and FAPESP (Research Support Foundation in the State of São Paulo) for the financial support.

REFERENCES

[1] Gatti, R., Gioia, M.G. and Cavrini, V. (2000) Analysis and stability study of retinoids in pharmaceuticals by fluorescence. *Journal of Pharmaceutical and Biomedical Analysis*, **23(1)**, 147-159.

[2] Taylor, P.W. and Keenan, M.H.J. (2006) Pharmaceutical quality of generic isotretinoin products, compared with Roaccutane. *Current Medical Research and Opinion*, **22(3)**, 603-615.

[3] Ioele, G., Cione, E., Risoli, A., Genchi G. and Ragno, G. (2005) Accelerated photostability study of tretinoin and isotretinoin in liposomes formulations. *International Journal of Pharmaceutics*, **60(1-2)**, 251-260.

[4] Wetterich, U. and Mutschler, E. (1995) Quality of cefotaxine sodium preparations. *Arzneimittel-Forschung Drug Research*, **45(1)**, 74-80.

[5] Lambert, P.A. and Conway, B.R. (2003) Pharmaceutical quality of ceftriaxone generic products compared with Rocephin. *Journal Chemotherapy*, **15(4)**, 357-368.

[6] Lammer, E.J., Chen, D.T., Hoar, R.M., Agnish, N.D., Benke, P.J. and Braun, J.T. (1985) Retinoic acid embryopathy. *New England Journal of Medicine*, **313(14)**, 837-841.

[7] Willhite, C.C., Wier, P.J. and Berry, D.L. (1989) Dose response and structure-activity considerations in retinoid-induced dysmor phogênesis. *Critical Reviews in Toxicology*, **20(2)**, 689-695.

[8] (2007) United States pharmacopeia. XXX, *Pharmacopeia Convention*, Maryland, 2428-2429.

[9] Lima, E.M., Diniz, D.G. and Antoniosi-Filho, N.R. (2005) Development of a gas chromatography method for the determination of isotretinoin and its degradation products in pharmaceuticals. *Journal of Pharmaceutical of Biomedical Analysis*, **38(4)**, 678-685.

[10] Tashtoush, M.B. Jacobson, E.L. and Jacobson, M.K. (2007) A rapid HPLC method for simultaneous determination of tretinoin and isotretinoin in dermatological formulations. *Journal of Pharmaceutical and Biomedical Analysis*, **43(3)**, 859-864.

[11] Tan, X., Meltzer, N. and Lindenbaum, S. (1992) Solid-state stability studies of 13-cis-retinoic acid and all-transretinoic acid using microcalorimetry and HPLC analysis. *Pharmaceutical Research*, **9(9)**, 1203-1208.

[12] Allwood, M.C. and Plane, J.H. (1984) The degradation of vitamin A exposed to ultraviolet radiation. *International Journal of Pharmaceutics*, **19(2)**, 207-213.

[13] Crank, G. and Pardijanto, M.S. (1995) Photooxidations and photosensitized oxidations of vitamin A and its palmitate ester. *Journal of Photochemistry and Photobiology A*: *Chemistry*, **85(1-2)**, 93-100.

[14] Bempong, D.K., Honigberg, I.L. and Meltzer, N.M. (1993) Separation of 13-*cis* and all-trans retinoic acid and their photodegradable products using capillary zone electrophoresis and micellar electrokinetic chromatography (MEC). *Journal of Pharmaceutical and Biomedical Analysis*, **11(9)**, 829-833.

J. Biomedical Science and Engineering, 2010, 3, 459-469 **JBiSE**
doi: 10.4236/jbise.2010.35064 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

An analysis of quantitative PCR reliability through replicates using the C_t method

Chris C. Stowers[1], Frederick R. Haselton[2], Erik M. Boczko[3*]

[1]Bioprocess R&D Division Dow AgroSciences LLC Indianapolis, Indiana, USA;
[2]Department of Biomedical Engineering Vanderbilt University Nashville, Nashville, USA;
[3]Department of Biomedical Informatics Vanderbilt University Medical Center Nashville, Nashville, USA.
Email: erik.boczko@vanderbilt.edu

Received 6 February 2010; revised 1 March 2010; accepted 2 March 2010.

ABSTRACT

There is considerable interest in quantitatively measuring nucleic acids from single cells to small populations. The most commonly employed laboratory method is the real-time polymerase chain reaction (PCR) analyzed with the crossing point or crossing threshold (C_t) method. Utilizing a multiwell plate reader we have performed hundreds of replicate reactions each at a set of initial conditions whose initial number of copies span a concentration range of ten orders of magnitude. The resultant C_t value distributions are analyzed with standard and novel statistical techniques to assess the variability/reliability of the PCR process. Our analysis supports the following conclusions. Given sufficient replicates, the mean and/or median C_t values are statistically distinguishable and can be rank ordered across ten orders of magnitude in initial template concentration. As expected, the variances in the C_t distributions grow as the number of initial copies declines to 1. We demonstrate that these variances are large enough to confound quantitative classification of the initial condition at low template concentrations. The data indicate that a misclassification transition is centered around 3000 initial copies of template DNA and that the transition region correlates with independent data on the thermal wear of the TAQ polymerase enzyme. We provide data that indicate that an alternative endpoint detection strategy based on the theory of well mixing and plate filling statistics is accurate below the misclassification transition where the real time method becomes unreliable.

Keywords: Misclassification Transition; Single Molecule Counting; Rank Ordering Running Title; PCR; Replicates and Reliability

1. INTRODUCTION

Real time Polymerase Chain Reaction (PCR) is widely used for quantitative analysis [1-3] in a variety of clinical and research areas including the study of genetically modified foods, vaccine efficacy, and in systems biology [4-6]. In a real time PCR reaction the DNA amplification process is recorded. The goal of a quantitative analysis is to use the amplification time series data, $y(n) = AX(n)$, to solve the inverse problem of determining a reasonable proxy for the amount of initial template, $X(0)$. A satisfactory solution of this inverse problem has been hampered by the amplification of error, dilution error, the multivariate nature of the enzyme system and the lack of a model that accounts for the variability [1,7-11].

In its stead, the ad-hoc crossing point, C_t, method has emerged [12]. The heuristic behind the selection of the crossing threshold is predicated on the observation that amplification curves from identically prepared initial conditions diverge dramatically with iteration. In accordance with this observation a threshold value is chosen close to the detection threshold to limit variability from amplification error. But variability remains. It is a well known result from the statistics of Subsection 4.8 [13] that averaging over replicates reduces variance in ratio to the number of replicates. In practice, from 1 to 5 replicates appears typical. Since there is a balance between cost, time, effort, and accuracy it is of practical interest to understand the requirement for replicates and its dependence on initial template concentration.

The goal of this study was to examine the variability in the distribution of C_t values generated from a large number of identically prepared replicates as a function of the initial template concentration to determine the following:

1) Are measures of central tendency informative of initial template concentration over 10 orders of magnitude?

2) How many replicates are required to discriminate between different initial template concentrations?

3) How does the number of replicates depend on the initial template concentration?

Using a multiwell plate format we measured hundreds of replicates to produce C_t value distributions. Using standard and novel statistical techniques we analyze the C_t value distributions and demonstrate that the sample mean and/or median C_t values are statistically significantly distinguishable over ten orders of magnitude. Furthermore, we show that the sample mean C_t values are reliably ordered according to the initial concentration of template. In other words, if x and y are initial template concentrations with $x < y$ and μ_x and μ_y are the corresponding sample mean C_t values then $\mu_x > \mu_y$. The order reverses because less initial template requires more cycles of PCR to amplify. We utilize ordering as a convenient and natural device to quantify the role of replicates on reliability. We ask and answer the following question: Given an unlabeled dilution series how many replicates are required to reliably order the tubes? We find that the answer depends on the range of initial template.

A focus of this work was to cover as broad a range of initial conditions as possible with the same experimental format. We observed that the mean and/or median C_t values had the smallest variance above 10^4 initial copies. Most published standard curves focus on this range [2]. Few studies have analyzed issues of variability and robustness below this range. We show that below 10^4 initial copies the probability of misclassification of the initial template concentration given a C_t value grows rapidly and saturates near a half. The dispersion in the C_t value distributions and the rise in misclassification correlate with an independent measure of the thermal wear of the TAQ polymerase enzyme.

Driven by the observed broadening of the C_t value distributions below a thousand initial copies, and inspired by elegant methods that sidestep the issues created by the dynamics of exponential growth [14-18], we examined a format for single molecule detection utilizing an endpoint analysis and the statistical properties of well mixing and plate filling. We present data that such an assay is accurate where the real time method becomes unreliable.

2. MATERIALS AND METHODS

2.1. PCR

Rt-PCR results were generated using linearized double stranded EC3 plasmid DNA containing the ybdO gene. The plasmid was linearized by digestion with the restriction enzyme BamH1 prior to PCR. The following primer sequences were used.

1) Forward: 5'-AAT TAT TCT AAA ACC AGC GTG TC-3'

2) Reverse: 5'-TTT GGG ATT GAA TCA CTG TTT C-3'

The PCR supermix was prepared as described in [19], with the exception that we used Qiagen HotStarTaq Cat # 203203, Roche dNTPs Cat# 13583000, DMSO Sigma # D8418 at 2%, and Sybr Green (Sigma # 86205) at 5-times the recommended concentration. Primers were used at a concentration of 1 μM. All samples were run on the 384 well plate platform using an Applied Biosystems 7900HT thermocycler and the SDS 2.3 software. The C_t value threshold was set at 5.0 RFU (Relative Fluorescence Units) for all samples. The DNA concentrations of concentrated stocks were measured using a Nano-Drop 100 spectrophotometer prior to use. Subsequent dilutions were performed using sterile, nuclease free water from Ambion # AM9937. The following thermo-cycling program was used.

1) 2 min at 50°C Initial Warmup Phase;
2) 15 min at 95°C Initial TAQ Activation Step;
3) 1 min at 95°C DNA Denaturation;
4) 1 min at 50°C Primer Annealing;
5) 1 min at 72°C DNA Extension;
6) 0.25 min at 80°C Fluorescence Measurement;
7) Repeat Steps 3-6 forty times.

2.2. Preparation of Identical Replicates

To ensure uniformity in the face of pipetting error the PCR supermix was prepared in well-mixed batches in a 14-mL conical tube. Each sample consisted of 184 replicates and 8 negative controls, requiring exactly half of a 384 well plate. All of the components except for DNA were loaded into a 14 mL conical tube in the following order: 800 μL PCR buffer, 5.6 mL of nuclease free water, 160 μL DMSO, 320 μL MgCl2 (Qiagen Cat # 124113012), 160 μL of a primer mix (a 1:1 mix of the forward and reverse primer stored at a concentration of 50 M each), 160 μL Sybr Green (100X stored in DMSO), 160 μL of dNTPs, and lastly 80 μL of Taq polymerase. We have noticed that the order at which these are added affects the reproducibility of the assay. The mixture was vortexed at high speed for 1 minute. 335 μL of supermix was then removed to be used as a negative control and placed into a 1 mL eppendorf tube and 25 μL of water was added. This mixture was then briefly vortexed to ensure well mixing. The remaining 7.105 mL of supermix was then split equally four ways into 2 mL cryostat tubes, and 134 μL of water plus the amount of desired DNA was added to each cryostat tube. Each tube was then briefly vortexed. For each reaction contained within a single well of the plate, 10 μL of the respective reaction mix was loaded into a well of the 384 well plates.

2.3. TAQ Polymerase Pre-Wear Assay

The PCR supermix was prepared as described above, but without template DNA. Steps 1 and 2 of the PCR process were executed following which samples were pre-worn by thermocycling the supermix as described in steps 3 through 6 above. Samples were pre-worn from 5 to 40 cycles. 10^8 copies of initial template DNA were

added to the preworn enzyme with subsequent resumption of cycling. An efficiency was calculated by averaging the derivative over the resultant amplification curve.

2.4. Statistical Analysis of C_t Distributions

The sample mean C_t values for each initial template concentration were compared pairwise using a permutation test that is asymptotically valid and robust in situations where the distributions are not necessarily normal and/or the ratio of the variances is unknown, indicating that a t-test is not supported [20,21]. The test statistic T [20], measures the difference in mean rank of the samples within their union, scaled by a consistent estimator of their variance. Because the C_t value distributions may be skewed by outliers, we also considered the median as a measure of central tendency. The median C_t values were compared pair-wise using a bootstrap test that has been shown to outperform all reasonable alternative methods [22].

Given a linear regression, $y = mx + b$, of the mean/median C_t values against $x = \log(n)$, the log of the number of initial copies of template, a relative error was calculated from the quantiles of the C_t value distributions as follows. Allow Q_h^x and Q_l^x to be high and low quantile values chosen from the C_t value distribution generated by initial log template x. Since the C_t value generally increases with decreasing amount of initial template the slope m of the regression line(s) is negative. Thus, the difference in the predicted amount of initial template DNA from the distributions divided by the input amount is given as:

$$\frac{\Delta n}{n} = \frac{10^{(Q_l^x - b)/-m} - 10^{(Q_h^x - b)/-m}}{10^x} \quad (1)$$

Let U represent the universe of possible C_t values, and let T stand for the collection of possible initial template concentrations. The initial template concentrations are thought of as the class labels. We consider the probability of misclassifying an observed C_t value given a known class label. Suppose that we draw a C_t value from a given class, how likely is it to find that value in any of the other classes? The mean misclassification probability is estimated from the C_t value distributions corresponding to different initial template concentrations according to the following formula.

$$P(x) = \sum_{i \in U} P(i \mid x) P(i \mid T \setminus x) \quad (2)$$

where $P(i \mid x)$ is the conditional probability of finding the C_t value i, given the initial template concentration x, and the $P(i \mid T \setminus x)$ is the conditional probability of observing that same value given any initial template concentration other than x. The later conditional probability is interpreted as the probability of misclassification. The conditional probabilities are estimated from the measured C_t value frequency distributions.

2.5. Plate Filling with Microbeads

Experiments were performed using 20μm latex beads from Beckman Coulter (#PN6602798) using flat bottom 96 well plates from Becton Dickinson Labware. 96 well plates were used in place of 384 well plates for ease of microscopic analysis. Various dilutions of beads were prepared using a Beckman Multisizer Coulter Counter 2. 25 μL of each dilution was loaded into each well of the 96 well plate. The number of beads in each well was counted with a Nikon TE-2000 microscope.

2.6. Plate Filling Simulations

The statistics of the plate filling stochastic process were modeled using Monte-Carlo simulation. For instance, the expected number of empty wells in a 96 well plate was estimated by simulation using the following function:

```
Table[Mean[
  Table[Length[
    Complement[Range[96],
      RandomInteger[Range[96], m]]], {10000}]], {m,
1, 600}]
```

Here m is the number of molecules being plated from a well mixed solution. The mean is estimated from 10,000 realizations. The standard deviation is computed by replacing the function Mean by Standard Deviation. A graph of these functions is shown in **Figure 9**. All simulations and analysis were carried out in Mathematica 6.03 (Wolfram Research), and the notebooks are available upon request.

3. RESULTS

The C_t value data are summarized in **Figure 1**. The figure shows that above 10^4 copies the data are distributed about the median with smaller variance than those below. Outliers exist across all the data, mostly trending upward of the median, indicative of reactions lagging behind the pack. The data show the distributions as collected across ten orders of magnitude in initial template.

3.1. Mean and Median C_t values

While the C_t value distributions below 10^4 copies of initial template DNA are broad and noisy, the sample means and medians form a monotone increasing series when stratified according to initial template. These data are shown in **Table 1**. The means and medians are very nearly equal in all cases and the data distributions appear unimodal.

At initial template concentrations larger than 10^4 initial copies the data distributions seen in **Figure 1** appear

Table 1. Sample mean and median C_t values. The corresponding distributions are shown in **Figure 1**. Observe that the sample means and medians form an ordered sequence stratified by initial template.

Distribution Number	1	2	3	4	5	6	7	8	9	10	11
Log Copies Template	9.0	8.0	7.3	6.3	5.7	4.1	3.1	2.1	1.9	1.1	0.3
Replicates	184	182	183	183	184	183	184	179	730	175	178
Mean C_t Value	6.6	10.6	12.8	15.6	19.5	25.5	27.6	28.8	30.7	35.5	36.4
Median C_t Value	6.6	10.4	12.8	15.7	19.3	25.5	28.1	28.7	30.7	35.0	35.7

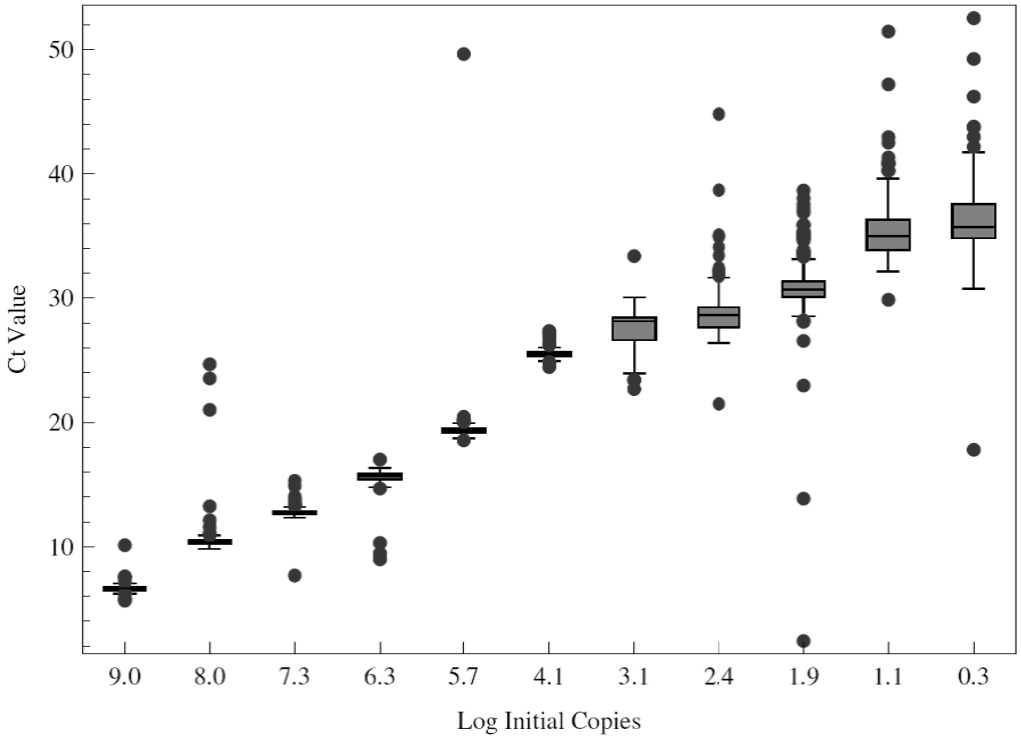


Figure 1. Summary of the C_t value data, stratified as a function of the log of the initial number of copies of DNA amplified. As described in the materials and methods section, a minimum of 175 replicates were run at each initial condition. The box covers two quartiles about the median with outliers shown as the red dots. Outliers are defined as points beyond 3/2 the discrete interquantile range from the edge of the box.

to identify distinguishable mean and median values and no hypothesis test appears necessary. However at lower template concentrations it is not clear that the means are not within a fraction of a standard deviation of one another. When all of the data points are taken into account the means and medians are statistically different at a level of significance greater than 1/20000 with the exception of the smallest two medians that were statistically different at the 3/1000 level of significance. The result of the significance test for the means is displayed in **Table 2**.

The C_t value distributions may not be normal and because we do not a priori know the ratio of the variances, standard hypothesis tests to compare means are not justified. Several non-parametric methods have been proposed for this common situation [20,21] and have shown to have power. The last row of **Table 2** contains the T-statistic comparing template-adjacent distributions. The values of the T-statistic are generally declining as template decreases and come closer to the upper boundary of the null distribution. Because the fourth and fifth distributions are non-overlapping the variance estimator takes its minimum value and this explains the large T value. All of the other pairs of distributions contain some overlap and hence have much larger variances. The data indicate that with approximately 184 replicates the C_t method can provide discrimination over nine orders of magnitude. One goal of

our work, to be described below, is to provide some analysis of how few replicates are required to reliably make the same claim.

3.2. Standard Curves

A linear regression of the mean C_t values against log initial copies of DNA template is shown in **Figure 2**. The regression line captures the data reasonably well over the entire range of initial template concentration. In contrast, **Figure 1** gives the impression that the data are described by a function that is initially linear and then contains at least two sigmoid like transitions. This can also be seen in the oscillation of the data about the regression line in **Figure 2**.

In an independent experiment with the same enzyme and supermix, we examined the wear that the enzyme experiences as a function of thermocycling alone. The middle of the transition region corresponds to a C_t value of approximately 25. As seen in **Figure 1**, the thermal wear transition corresponds to the region near 10^4 copies where the distributions begin to broaden. Because of this independent observation, we split the data in two at this transition point and considered a piecewise linear regression. The data are summarized in **Figure 4**.

Given a regression model of the data we computed a measure of the relative error of the inversion process, from the spread in the C_t value distributions, using the equation for n, as described in Subsection 2.4. The first and third quartiles were used to produce the data shown in **Figure 5**. The data show that the relative error is approximately 20% above 10^4 initial copies, and rises sharply below. The piecewise linear regression, shown as the red line, produces a larger relative error in the transition region, and agrees elsewhere.

Table 2. Results of testing the null hypothesis of stochastic equality. A hypothesis test was applied pairwise to C_t value distributions adjacent in initial template concentration. The extreme values of the distribution of the T-statistic under the null hypothesis are shown along with the T-statistic for adjacent distributions. The distributions, e.g. "2v1", are labeled as in **Table 1**. In each test, the null distribution was simulated using 20,000 random permutations applied to the pooled data as described in [20]. In each test the calculated T-statistic fell outside of the range indicating that the p-value is less than 1/20000. In each case the null hypothesis is rejected with confidence.

Property	2v1	3v2	4v3	5v4	6v5	7v6	8v7	9v8	10v9	11v10
Min	-5.02	-3.85	-4.63	-4.25	-4.06	-4.10	-3.94	-4.62	-4.50	-3.57
Max	3.93	4.17	3.87	3.74	4.13	5.03	4.77	4.87	3.73	3.74
T	684.33	39.61	51.45	23809.7	91.0	24.27	7.29	18.37	81.7	4.22

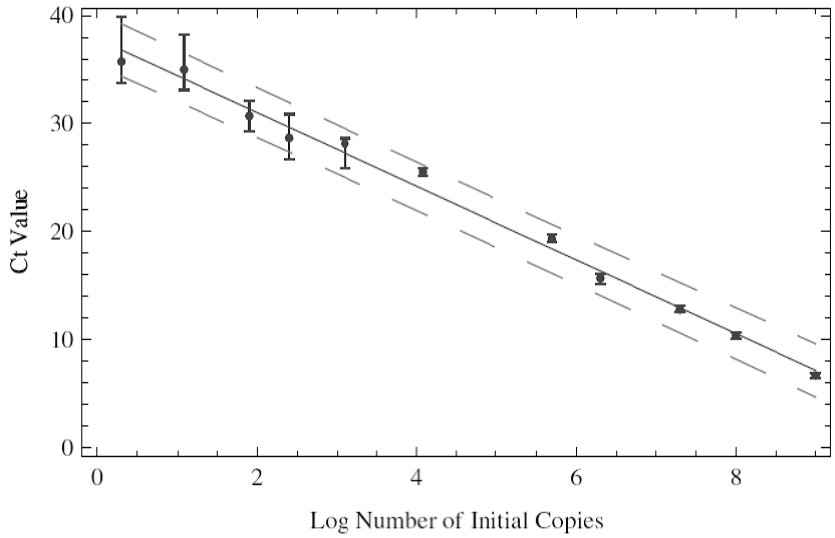


Figure 2. Linear regression of the mean C_t values with log initial copy number. The regression line is shown in red along with a 95% confidence interval in dashed line. The error bars on the individual data points reflect one standard deviation computed from the C_t value distributions.

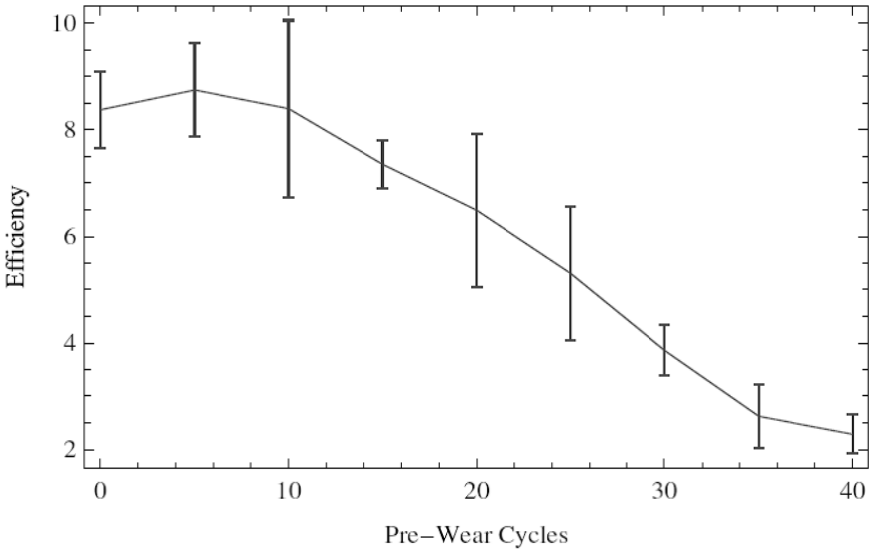


Figure 3. TAQ efficiency as a function of thermo-cycling pre-wear. An efficiency is computed as the average derivative of relative fluorescence over the amplification curves. Error bars represent a standard deviation over three independent experiments.

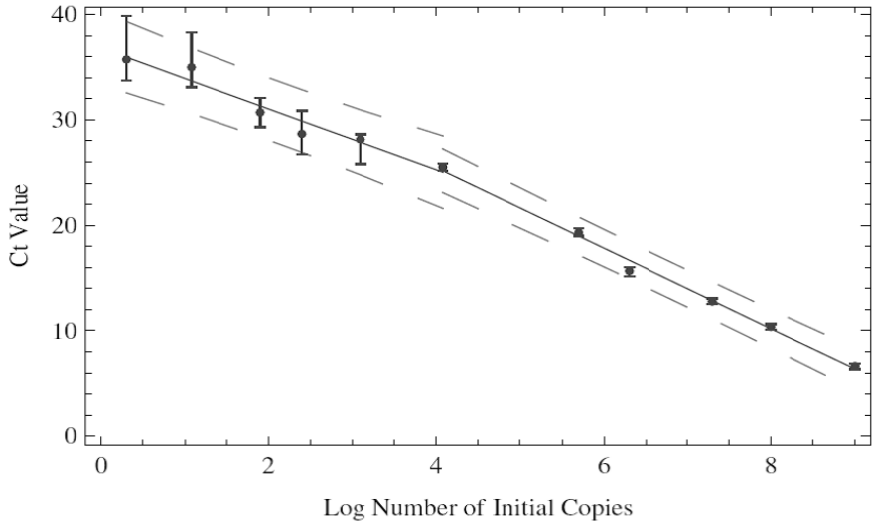


Figure 4. Piecewise linear regression of the mean C_t values. The data were split according to the trends observed in **Figures 1** and **3**.

3.3. Misclassification

The graph of the relative error of the linear regression, shown in **Figure 5** is an attempt to quantify the effect that the broadening of the C_t value distributions has on the inverse problem of assigning an initial template concentration from a measured C_t value. In an alternative attempt to summarize the impact of the variance of the C_t value distributions on the process of categorizing the initial template concentration based on a measured C_t value we considered statistics such as that summarized in the equation for $P(x)$ described in Subsection 2.4.

The construction of a standard curve, as described in the previous section, is one way to assign a template concentration to a measured C_t value. Now consider another. Consider each distribution of C_t values as representing a class. The class label is the initial template concentration. The heuristic is this: Given a measured C_t value coming from one of these classes, how likely is it to misassociate this value with the wrong class label? The larger the overlap between the meat of the distributions, the more likely it is that an unknown value will be misclassified. The formula for $P(x)$ given quantifies this notion. The results, conditioned on our data set, are shown in **Figure 6**.

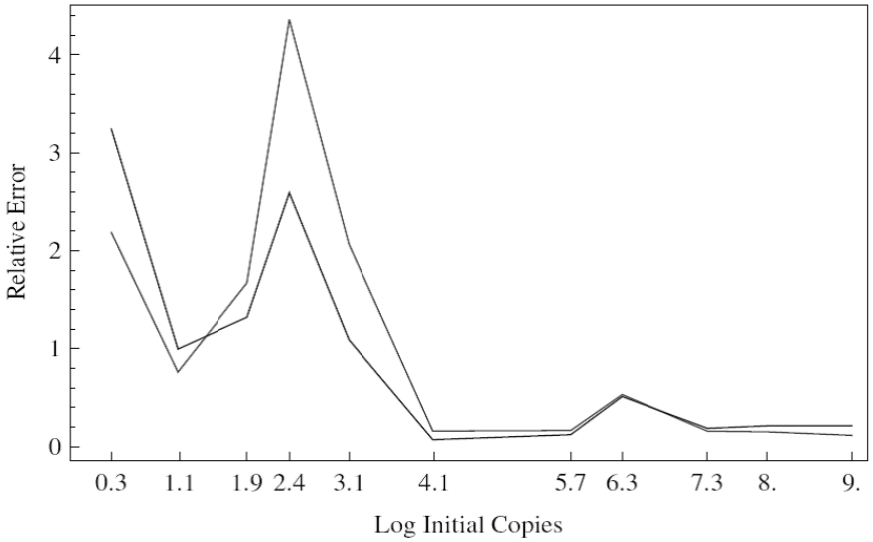


Figure 5. Relative error of the PCR process as calculated from the C_t value distributions and the standard curves shown in **Figures 2** and **4**, Blue and Red respectively. The C_t values corresponding to the first and third quartile were used to calculate a ΔDNA value whose limits were calculated using the regression line(s), according to the definition of n given in Subsection 2.4.

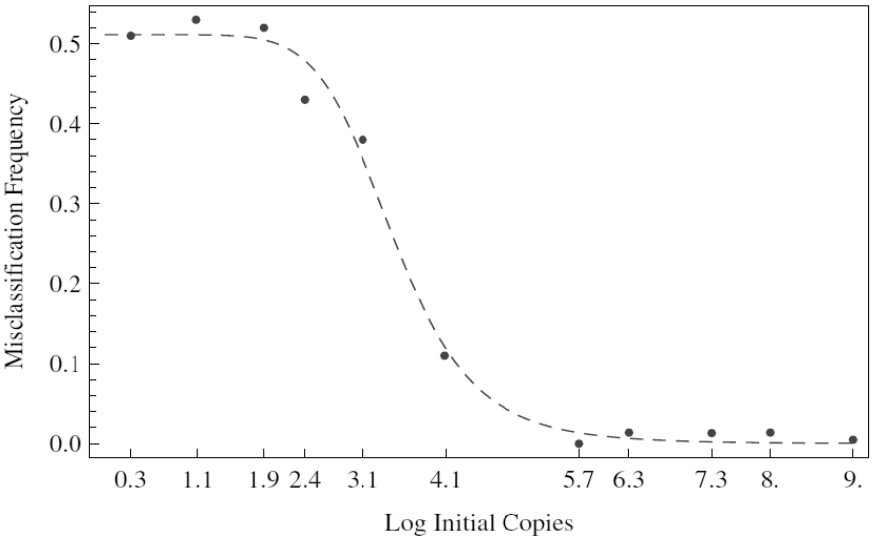


Figure 6. Misclassification frequency calculated according to the definition of $P(x)$ given in Subsection 2.4 and conditioned upon the C_t value distributions shown in **Figure 1**. A best fit Hill's function is shown as the dashed line. The midpoint of the transition occurs at approximately 2950 copies of initial template DNA.

The analysis indicates, as is clearly correlated with the results shown in **Figure 1**, that above 10^5 copies the probability of misclassification are small, less than 0.1. Over a narrow range spanning the next two decades of concentration the probability rises to one half. Subsequently, at lower concentrations, the probability more or less saturates at the value of one half and equates classification to coin flipping. The two sided nature of the flip is interpreted as the chance of being in the right class or out, in a one versus all sense. From a best-fit Hill's function, shown as the dotted curve, it is deduced that the midpoint of the misclassification transition occurs at $10^{3.47} \approx 2950$ initial copies of template DNA. As can be seen in **Figure 2**, the ordered pair (3.47,25) lies

close to the regression line and well within the error stripe, indicating the concordance between the relative error, misclassification and pre-wear data.

3.4. Rank Ordering with Replicates

In the first section we summarized the findings that the sample mean and median C_t values computed from all the data are statistically distinguishable and rank ordered. This was the case using all 184 replicates per initial template concentration. Doubtless it is of interest to understand how few data are required to make this same claim.

In this section we imagine the following experiment. Suppose that an investigator has produced a serial dilution in a set of tubes, with no error other than that coming from a correctly calibrated pipetteman. However, the tubes containing the dilution series are unlabeled and become scrambled as to order, but not contents. It is unequivocal that the amount of DNA within the tubes form a monotone series. The question is to determine how many replicate PCR runs are required such that the resultant sample mean C_t values correctly order, and hence label, the tubes with a given level of confidence? We imagine that rank ordering is one level of quantitation removed from inversion, but a closely related pre-requisite.

Conditioned on our data we can answer this question by Monte-Carlo simulation. From each C_t value distribution we draw k replicates uniformly at random and compute their sample mean μ_k^x. This mimics an experiment with k-replicates. The resulting set are sorted according to their numerical values and if this order agrees with that of their class labels, x, we score this trial positive. We compute the fraction of positive draws that resulted in correctly ordered sample means, from a total of 20,000 draws at each value of k. The results are shown in **Figures 7** and **8**.

Figure 7 shows that individual data points can rank order the highest concentrations from 10^4 to 10^9 initial copies with greater than 90% accuracy conditioned on our data. The use of 8 or more replicates guarantees 99% accuracy. The inset to **Figure 8** shows the performance when all the data are considered together. These data indicate that 35 or more replicates are required to exceed 90% accuracy over the entire concentration range. **Figure 8** shows that the larger variance of the distributions with smaller initial template is responsible for the behavior of the sample over the entire range.

3.5. End Point Detection

In the previous sections we have detailed an analysis of variance that shows that below 10^4 and certainly below 100 initial copies of template DNA, quantitation via standard curves or classification via a C_t value is confounded by error. As an alternative we describe a format for single molecule detection utilizing an endpoint analysis. We have separated the analysis into two parts. In the first part we consider the process of plate filling. We show that theory and experiment are in good agreement. In the second part we show that the process of amplification by PCR simply reveals the pattern of plate filling.

Suppose, as before, that we pipette identical, well mixed, aliquots of a DNA solution into the wells of a multiwell plate and perform 40 cycles of PCR and count the total number of wells that amplified above the detection limitation of the machine. Our data demonstrate that the expected number of unamplified wells is an informative statistic that has the property that its error declines as the total number of molecules of DNA declines [23].

Figure 9 shows the excellent congruence between

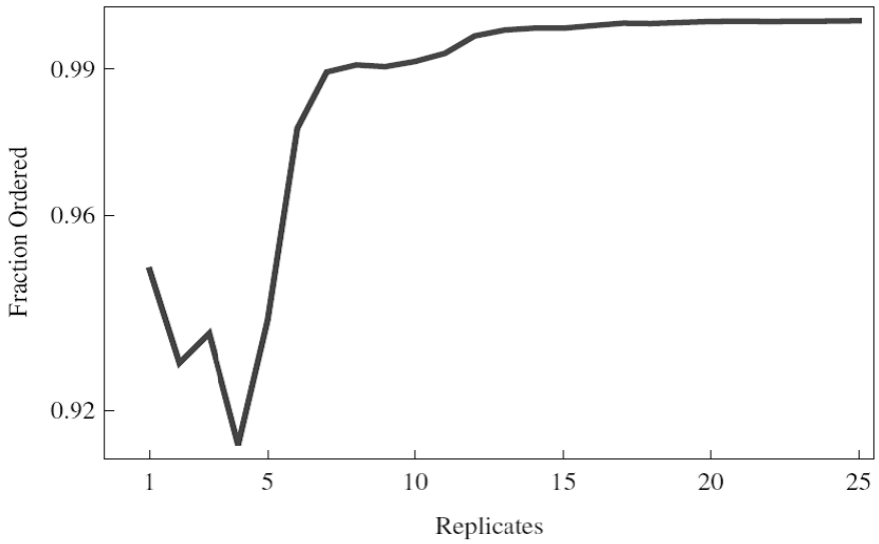


Figure 7. Reliability of replicates given the task of rank ordering a dilution series. The results for rank ordering the initial template concentrations with $x \geq 10^4$.

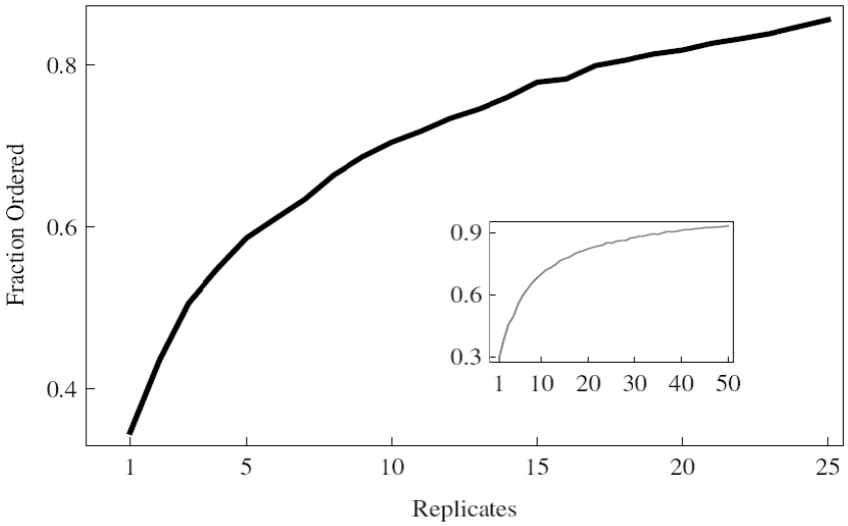


Figure 8. Reliability of replicates given the task of rank ordering a dilution series. The results for rank ordering the initial template distributions with $x < 10^4$. Inset shows the results over the entire range of initial template DNA, see **Table 1**.

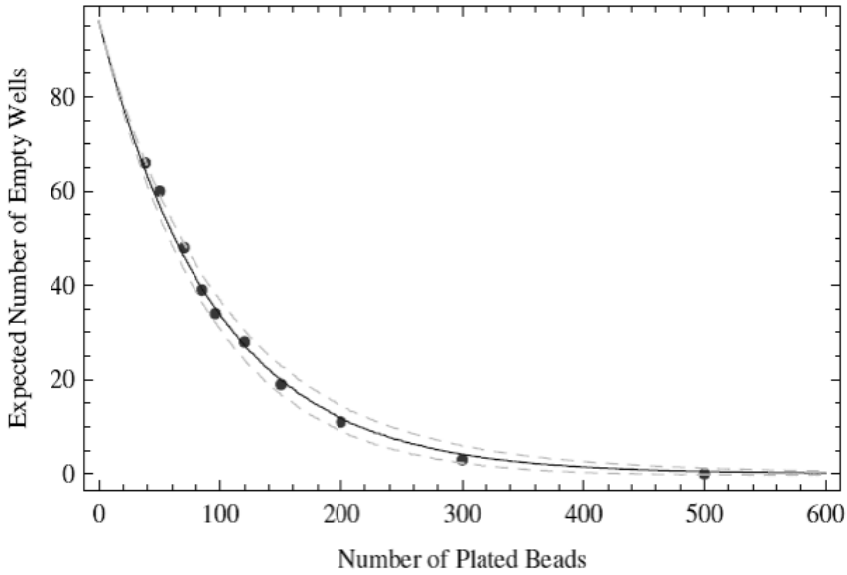


Figure 9. Agreement between experimental and theoretical plate filling statistics. The expected number of empty wells is shown as a function of the total number of beads distributed among the wells of a 96-well plate. The expected number of unfilled wells calculated from theory is shown as the solid blue line, while one standard deviation is shown as the dotted green line. Experimental data are shown in red.

theory and experiment for filling of 96-well plates with solutions of 20 micron latex beads as described in the methods. Utilizing PCR to discriminate wells filled with DNA from those that are not, is more complex than optical bead counting. **Figure 10** describes the results of an endpoint analysis. Each data point represents an independent experiment. The concentrations of DNA plated for PCR were determined through dilution as described in methods. When these putative DNA concentrations are uniformly scaled by a factor of 2.4 the data agree remarkably well with the theory over the entire length of the plate filling curve.

4. CONCLUSIONS

With a sensitive PCR plate reader we have examined the variability of real time PCR process. We have shown

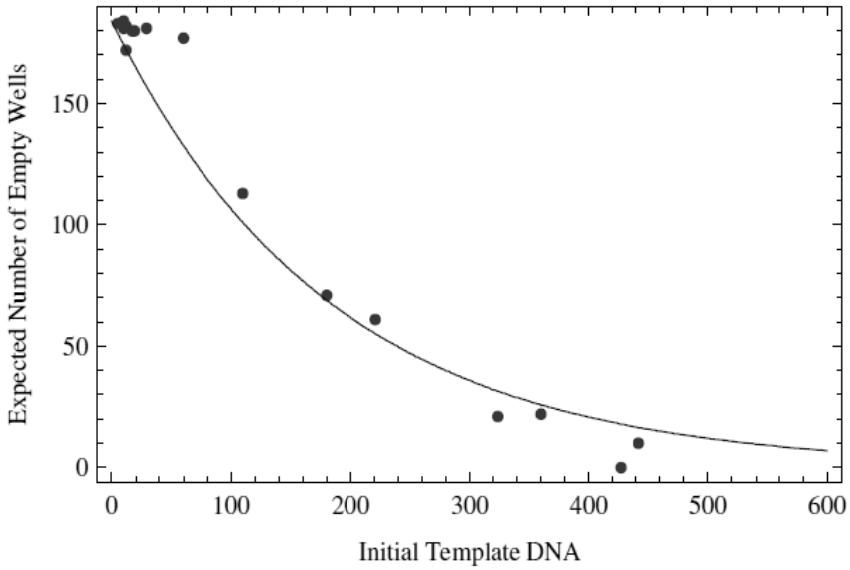


Figure 10. The number of unamplified wells as a function of the number of DNA molecules spread over a 384 well plate. The expected number of empty wells calculated from theory in blue. The red dots represent experimental PCR endpoint data.

that using less than a few hundred replicates per initial template concentration ensures that the mean C_t values are statistically distinguishable and rank ordered corresponding with the initial amount of template DNA. We have shown that the mean/median values can be regressed over ten orders of magnitude.

The C_t value distributions appear noisy below 10^4 copies of initial template DNA and the results of two independent statistical techniques confirm this observation. Independent data on TAQ-wear indicate that the enzyme experiences a transition of decreasing efficiency in the corresponding region.

Given a standard curve, the C_t value distributions were used to quantify the impact of the variability of the C_t values on the process of predicting the initial template concentration. The relative error varies from 7% to 50% over the highest initial template concentrations averaging near 20%. Both the relative error and the misclassification analysis capture the transition from low variance to high. The misclassification frequency is smoother and perhaps more intuitive to interpret while the relative error analysis perhaps more quantitative. The misclassification analysis suggests an alternative clasification approach to solving the inverse problem. Instead of using a standard curve to convert a C_t value into an initial template concentration we consider probabilistic classification into one of a discrete number of template classes. We are currently exploring this idea.

A question central to the analysis of the C_t value distribution data concerns the role of replicates. The value of replicates stems from the statistical fact that the variance of the distribution of sample means is smaller than the variance of the data distributions. What lessons can the data distributions teach about how many replicates are required in practice for resolution and reproducibility? Rank ordering of the sample means is a convenient and meaningful device for exploring this question. Rank ordering simulations with our data suggest that the number of replicates required depends on the range of initial template. Below the transition region individual data points provide better than 90% rank accuracy, while 35 or more replicates are suggested below the transition region for the same degree of rank accuracy. But this is a disheartening result for the following reason: More replicates are required precisely where the sample may be scarce. For samples with small initial template concentrations it may be more accurate to consider an endpoint method using the expected number of (un) amplified wells than to consider replicates.

These data, and the work of other groups [24] have demonstrated that the use of the C_t method in conjunction with statistical replicates renders the process of real time PCR capable of quantitative analysis for initial DNA samples ranging from upwards of 10^4 molecules. However the data presented above show that for tens to hundreds of initial copies real time PCR is unreliable for quantitation. We and other groups have been exploring alternative methods for single molecule detection and counting. In this regard, a process involving endpoint analysis shows significant promise.

We have described a decomposition of single molecule counting into plate filling and amplification. We have demonstrated through simulation and experiment that the expected number of (un)amplified wells is a

robust statistic on which to base an inverse problem or a rigorous hypothesis test to count small numbers of single molecules. The observed linear scaling between expected and perceived DNA concentration indicates that amplification by PCR is directly related to plate filling.

It remains an open problem to determine the conditional probability with which PCR can amplify above threshold in 40 cycles from a single strand of DNA. While it is currently impossible to enumerate individual molecules or particles smaller than a nanometer, it is straightforward to count macroscopic objects such as latex beads or single yeast cells with a Coulter counter. Haploid yeast cells provide a convenient and verifiable means to deliver single copies of Bacillus subtilis genes such as ybdO into the wells of a multiwell PCR plate.

In this way, we are currently exploring the relationship between amplification and DNA copy number.

5. ACKNOWLEDGEMENTS

This work was partially supported through NSF-DMS 0443855, NSF ECS 0601528, NIH EB009235, and the short-lived W. M. Keck Foundation Grant#062014.

REFERENCES

[1] Freeman, W., Walker, S. and Vrana, K. (1999) Quantitative rt-PCR: Pitfalls and potential. *Biotechniques*, **26**, 124-125.

[2] Larionov, A., Krause, A. and Miller, W. (2005) A standard curve based method for relative real time PCR data processing. *BMC Bioinformatics*, **6**, 62.

[3] Yuan, J., Reed, A., Chen, F. and Stewart, C. (2006) Statistical analysis of real-time PCR data. *BMC Bioinformatics*, **7**, 85.

[4] Abramov, D., Trofimov, D. and Rebrikov, D. (2006) Accuracy of a real-time polymerase chain-reaction assay for quantitative estimation of genetically modified food sources in food products. *Applied Biochemistry and Microbiology*, 4**2**, 485-488.

[5] Diehl, F., Li, M., Dressman, D., Yiping, H., Shen, D., Szabo, S., Diaz, L., Goodman, S., David, K., Juhl, H., Kinzler, K. and Vogelstein, B. (2005) Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proceedings of the National Academy of Sciences*, **102**, 16368-16373.

[6] Roussel, Y., Harris, A., Lee, M. and Wilks, M. (2007) Novel methods of quantitative real-time PCR data analysis in murine Helicoba*C*ter pylori vaccine model. *Vaccine*, **25**, 2919-2929.

[7] Jagers, P. and Klebaner, K. (2003) Random variation and concentration effects in PCR. *Journal of Theoretical Biology*, **224**, 299-304.

[8] Lalam, N., Jacob, C. and Jagers, P. (2004) Modeling the PCR amplification process by a size-dependent branching process and estimation of the efficiency. *Advanced Applied Probability*, **36**, 602-615.

[9] Liu, W. and Saint, D. (2002) A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Analytical Biochemistry*, **302**, 52-59.

[10] Nedelman, J., Haegerty, P. and Lawrence, C. (1992) Quantitative PCR: Procedures and precision. *Bulletin of Mathematical Biology*, **54**, 477-502.

[11] Vaerman, J., Saussoy, P. and Inargiola, I. (2004) Evaluation of real-time PCR data. *Journal of Biological Regulators and Homeostatic Agents.* 18, 212-214.

[12] Pfaffl, M. (2001) A new mathematical model for relative quantification in real-time PCR. *Nucleic Acid Research*, **29**, e00.

[13] DeGroot, M.H. (1986) Probability and statistics. 2nd Edition, Addison Wesley Reading MA, Massachusetts.

[14] Cady, N., Stelick, S., Kunnavakkam, M., Lui, Y. and Batt, C. (2004) A microchip-based DNA purification and real-time PCR biosensor for bacterial detection. *Proceedings of IEEE Sensors*, Vienna, 1191-1194.

[15] Matsubara, Y., Kerman, K., Kobayashi, M., Yamamura, S., Morita, Y. and Tamiya, E. (2004) Microchamber assay based DNA quantification and specific sequence detection from a single copy via PCR in nanoliter volumes. *Journal of Biosensors and Electronics*, **20**, 1482-1490.

[16] Mitra, R. and Church, G. (1999) In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Research*, **27**, e34.

[17] Samatov, T., Chetverina, H. and Chetverin, A. (2006) Real-time monitoring of DNA colonies growing in a polyacrylamide gel. *Analytical Biochemistry*, **356**, 300-302.

[18] Williams, R., Peisajovish, S., Miller, O., Magdassi, S., Tawfik, D. and Griths, A. (2006) Amplification of complex gene libraries by emulsion PCR. *Nature Methods*, **3**, 545-550.

[19] Karsai, A., Muller, S., Platz, S. and Hauser, T. (2002) Evaluation of a home-made SYBR Green 1 reaction mixture for real-time PCR quantification of gene expression. *Biotechniques-Short Technical Reports*, **32**, 790-796.

[20] Neubert, K. and Brunner, E. (2007) A studentized permutation test for the non-parametric Behrens-Fisher problem. *Computational Statistics & Data Analysis*, **51**, 5192-5204.

[21] Reiczigel, J., Zakarias, I. and Rozsa, L. (2005) A bootstrap test of stochastic equality of two populations. *American Statistical Association*, **59**, 156-161.

[22] Wilcox, R.R. (2006) Comparing medians. *Computational Statistics & Data Analysis*, **51**, 1934-1943.

[23] Stowers, C. and Boczko, E.M. (2008) Platescale: The birthday problem applied to single molecule PCR. *Biocomp*'08, Las Vegas. http://www.ucmss.com/cr/main/papersNew/papersAll/BIC9135.pdf

[24] Cook, P., Fu, C., Hickey, M., Han, E. and Miller, K. (2004) SAS programs for real-time RT-PCR having multiple independent samples. *Biotechniques*, **37**, 990-995.

J. Biomedical Science and Engineering, 2010, 3, 470-475 **JBiSE**
doi: 10.4236/jbise.2010.35065 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

A digital cmos sequential circuit model for bio-cellular adaptive immune response pathway using phagolysosomic digestion: a digital phagocytosis engine

Sayed Mohammad Rezaul Hasan

Center for Research in Analog and VLSI Microsystem Design, School of Engineering and Advanced Technology, Massey University, Albany, New Zealand.
Email: hasanmic@massey.ac.nz

Received 26 February 2010; revised 11 March 2010; accepted 16 March 2010.

ABSTRACT

Living systems have to constantly counter micro-organisms which seek parasitic existence by extracting nutrition (amino acids) from the host. Phagocytosis is the ingestion of micro-creatures by certain cells of living systems for counter nutrition (breakdown of the micro-creature into basic components) as part of cellular adaptive immune response. These particular cells are called phagocytes, all of which are different types of white blood cells or their derivatives. Phagocytes are activated by certain components of the micro-creatures which act as an antigen, generating antibody secretion by the phagocyte. This paper develops a digital CMOS circuit model of phagocytosis: the immune response biochemical pathway of a phagocyte. A micro-sequenced model has been developed where the different stages in phagocytosis are modeled as different states clocked by circadian time intervals. The model converts the bio-chemical immune system digestive pathway into a cascade of CMOS multi-step logical transformations from micro-creature ingestion to the secretion of indigestible residuals. This modeling technique leads to the understanding of cellular immune deficiency diseases of living systems in the form of logical (electrical) faults in a circuit.

Keywords: Systems Biology; CMOS Circuit; Phagocytosis; Silicon Mimetic

1. INTRODUCTION

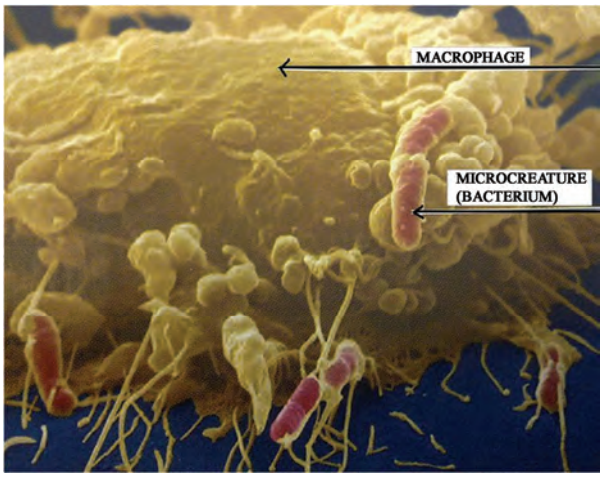
Looking at bio-chemical pathways of biological systems from a micro-electronic circuit diagram perspective is an interesting analytical paradigm in developing the convergence of biosciences, biotechnology and electrical engineering. In an effort towards understanding the biochemical functions of living systems at molecular interaction level researchers working in the developing field of systems biology [1,2] have endeavored to provide models that integrate the molecular interactions between proteins [3], enzymes [4,5], DNA [6], RNA [4], ions etc. within individual bio-cells of different kinds and their derivatives.

Immune response is one of the most fundamental biochemical activity of living systems. Micro-creatures constantly harboring towards living systems for parasitic existence are subjected to counter nutrition mechanisms of certain type of immune system cells, e.g., white blood cells or their derivatives usually known as phagocytes. This process of immune response is know as phagocytosis. It is a cascade of molecular interactions within the phagocyte which result in the breakdown of the micro-creature into its basic components, parts of which are recycled (digested) for nutrition and the rest eliminated from the system as residual waste. Circuit theoretic techniques of modeling this biochemical immune system pathway within a phagocyte compared to other mathematical modeling techniques [7,8] is the central theme of this paper. In recently published papers [9-11] this author described how the states of DNA protein interactions within the bio-cellular operations can be modeled using integrated circuits (*silicon mimetic* model [12,13]). In addition, models for mitochondrial respiration related electron transfer pathway [14], and cell signaling pathway using G-Protein and Phosphorylation cascade [15] has been reported by this author. These models based on integrated circuits are amenable to the large variety of mature simulation tools [16] available in the Very Large Scale Integration (VLSI) Computer aided design (CAD) arena. This type of modeling thus enhances the knowledge required for the realization of the "dream" of designed biochemical pathways and gene circuits [12] with revolutionary outcomes for gene and cell therapy ("*nano-medicine*"). In this work a digital CMOS [17] micro-sequenced model for phagocytosis has been developed where the sequence of biochemical processing stages are clocked by circadian time (hours)

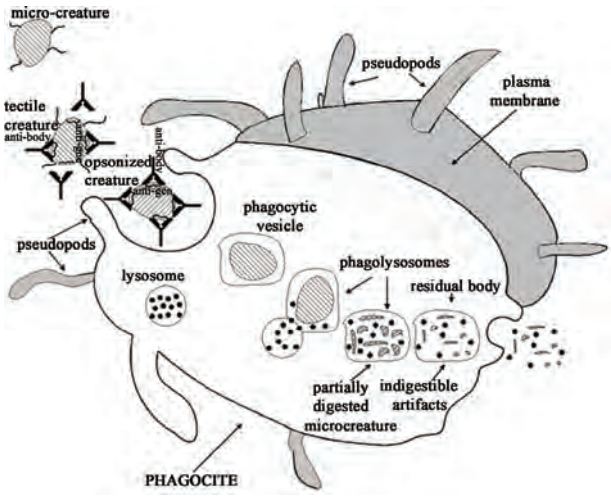
[7] intervals.

2. PHAGOLYSOSOMIC IMMUNE RESPONSE BIO-CHEMICAL PATHWAY

Infection of living systems by micro-creatures results in various types of white blood cells becoming phagocytes and moving into infected tissue. These phagocytes enlarge and develop into a macrophage engulfing and devouring micro-creatures in a counter nutrition effort. **Figure 1(a)** shows the microphotograph of such a macrophage digesting micro-creatures (red bacterium). **Figure 1(b)** shows the mechanism of phagocytosis in terms of the bio-chemical pathway in a phagocyte. The bio-chemical pathway of Phagocytosis [4,5] can be divided into several states such as chemotaxis (or phototaxis), adherence, opsonization, ingestion, digestion and residue ejection. Chemotaxis (or, phototaxis) by chemical (or optical) stimuli is the mechanism of approach through "run" and "tumble" of micro-creatures towards favorable host living systems. Phagocytes are attracted by the micro-creatures through their antigenic components, resulting in the release of antibody serum proteins that opsonize the microcreatures. Opsonized microcreatures then easily adhere to the plasma membrane of the phagocyte which is referred to as a state of adherence. After adherence, there is ingestion, whereby, the phagocyte extend its flagella type "arms" called pseudopods which surround and then engulf the microcreature. Once ingested, the surrounding pseudopods fuse and enclose the micro-creature in a fluid sack called "phagosome" or phagocytic vesicle. The membrane of this vesicle has enzymes which pump hydrogen ions (H^+) into the vesicle, thereby reducing the pH to around 4, so that hydrolytic enzymes can be activated for the breakdown of the micro-creature. The next stage of phagocytosis is the digestive pathway which begins with the phagosome "pinching off" from the plasma membrane and entering the cytoplasm. Inside the cytoplasm the phagosomes attach with lysosome sacks containing digestive enzymes and microbcidal substances. On attachment, the membranes of lysosome and phagosme merge forming a larger composite sack called phagolysosome. The degenerative components inside the phagolysosome take only sub-circadian time interval (less than an hour) to completely breakdown ("kill") micro-creatures inside it. This digestive pathway contains several main and subprocesses. Lysosomic enzyme lysozyme directly attacks the cell-wall of the micro-creature and initiates the hydrolysis of the cell-wall. A host of other lysosomic enzymes are also active at the same time, so that the microcreatures macro-components are disintegrated, such as its lipids by the enzyme lipase, its proteins by protease, its ribonucleic acids by ribonuclease, and, its deoxyribonucleic acids by deoxyribonuclease. Lysosomes also contain certain enzymes which can initiate a process known as oxidative burst which result in toxic oxygen products such as super oxide radical (O_2^-), hydrogen per oxide (H_2O_2), hydroxyl radical ($OH\bullet$) and singlet oxygen ($^1O_2^-$). Other lysosomic enzymes combine with these toxic oxygen products to produce secondary microbcides which hydrolyze and breakdown the microcreatures. For instance, the enzyme myeloperoxidase (MY) converts chlorine ions (Cl^-) and hydrogen per oxide (H_2O_2) into highly toxic hypochlorous acid ($HOCl$) which contain the well-known anti-microbial



(a)



(b)

Figure 1. (a) Photomicrograph [5] of phagocytic digestion (microphage) of red micro-creatures (bacterium); (b) a schematic illustration of the mechanism of phagocytosis (bio-chemical pathway) in a phagocyte.

agent hypochlorous ions.

3. A MICRO-SEQUENCED MODEL OF PHAGOCYTIC BIOCHEMICAL PATHWAY

A *silicon mimetic* [11] model of the bio-chemical pathway of phagocytosis has been developed and is shown in **Figure 2.** The discussion in the previous section indicates the existence of many intermediate steps in the biochemical digestive pathway of phagocytosis whose proper co-ordination is an important factor in the functional phagocyte of living systems. Similar to the concept introduced in [15], the intermediate steps in the bio-chemical path of phagocytic digestion of micro-creatures can be modeled as states in Delay flip-flops (Registers) which indicate the states of molecular interactions (binding, hydrolysis and breakdown) that constitute the behavior of the phagocyte. Using this modeling technique malfunction in the immune system (immune deficiency syndrome: cause of numerous diseases) can be conveniently modeled as circuit faults, such as, outputs of logic gates or flip-flops (registers) are stuck at logic "0" or at logic "1", thus unifying the approach to solving faults in electronic or biological circuits (in this case the biological circuit of the immune response sys tem). At any time-instant n, presence of Gram positive (GP) or Gram negative (GN) micro-creatures in close proximity to the phagocyte is stored as u1(n) in the D-FF_1. The presence of chemo-tectile micro-creatures due to the presence of chemical signals is stored as u2(n) in the D-FF_2. On the other hand, the presence of phototectile micro-creature due to the presence of optical signals is stored as u3(n) in the D-FF_3. The presence of chemo-tectile or phototectile micro-creature opsonized by anti-body serum protein (SP) is stored as u4(n) in the D-FF_4. The presence of ingested micro-creatures engulfed by pseudopods (PP) is stored as u5(n) in the D-FF_5. The state of the ingested micro-creatures within phagosomes with H^+ ions pumped inside the phagocytic vesicle is indicated by u6(n) in the D-FF_6. The "pinching off" of phagosome from the plasma membrane and entering the cytoplasm is registered as u7(n) in D-FF_7. The attachment and merger of lysosome with the phagosome in the phagolysosome, and the resulting state of the digestive degradation of the microcreatures is represented by,

$$u14(n) = \sum_{m=10}^{13} um(n-1) \tag{1}$$

The hydrolysis of micro-creature inside the phagolysosome due to hypochloric acid (a byproduct of the oxidative burst) is indicated by the state u8(n) stored in the D-FF_8. The breakdown of the micro-creature's cell-wall by the enzyme lysozyme (LYE) is indicated by the state u9(n) in the D-FF_9. The breakdown of lipids inside the degenerated micro-creatures by the lysosomic enzyme lipase is indicated by the state u10(n) stored in D-FF_10. Next, the breakdown of the bacterial proteins inside the degenerated microbe by the lysosomic enzyme protease is indicated by the state u11(n) stored in D-FF_11. The breakdown of the bacterial ribonucleic acids inside the degenerated microbe by the lysosomic enzyme ribonuclease (RNE) is indicated by the state u12(n) stored in D-FF_12. And, finally the breakdown of the bacterial DNA by the lysosomic enzyme deoxyribonuclease (DNE) is indicated by the state u13(n) stored in D-FF_13. Based on the micro-step sequenced model of **Figure 2**, the following discrete-time state equations can be written for the phagocytic digestive pathway:

$$u1(n) = GP + GN \tag{2}$$

$$u2(n) = CS \times ul(n-1) \tag{3}$$

$$u3(n) = OS \times ul(n-1) \tag{4}$$

$$u4(n) = [u3(n-1) + u2(n-1)] \times SP \tag{5}$$

$$u5(n) = u4(n-1) \times pp \tag{6}$$

$$u6(n) = u5(n-1) \times H^+ \tag{7}$$

$$u7(n) = u6(n-1) \tag{8}$$

$$u8(n) = [u7(n-1) \times \{MY \times Cl^- \times (LE1 \times H_2O)\}] \tag{9}$$

$$u9(n) = [u7(n-1) \times LYE] \tag{10}$$

$$u10(n) = [u8(n-1) + u9(n-1)] \times LPE \tag{11}$$

$$u11(n) = [u8(n-1) + u9(n-1)] \times PTE \tag{12}$$

$$u12(n) = [u8(n-1) + u9(n-1)] \times RNE \tag{13}$$

$$u13(n) = [u8(n-1) + u9(n-1)] \times DNE \tag{14}$$

$$u14(n) = [u10(n-1) + u11(n-1) + u12(n-1) + u13(n-1)] \tag{15}$$

Modeling immune response deficiency as electrical faults: Using these circuit model immune deficiency diseases can be modeled as electrical faults. For example, if the serum protein (anti-body) is not produced, the invading micro-creatures cannot be engulfed, and, the state at u(4) will be stuck at "0", resulting in all the down-stream digestive degradation pathway states being stuck at "0", *i.e.*, becomes non-functional. The solutions can then be designed in terms of CMOS circuit functionality and converted into the equivalent biochemical solution ("*nano-medicine*").

4. MICRO-SEQUENCED PHAGOCYTE PATHWAY STATES AND SILICON MIMETIC SIGNAL TRANSDUCTION RESULTS

Figure 3 displays how the immune system response

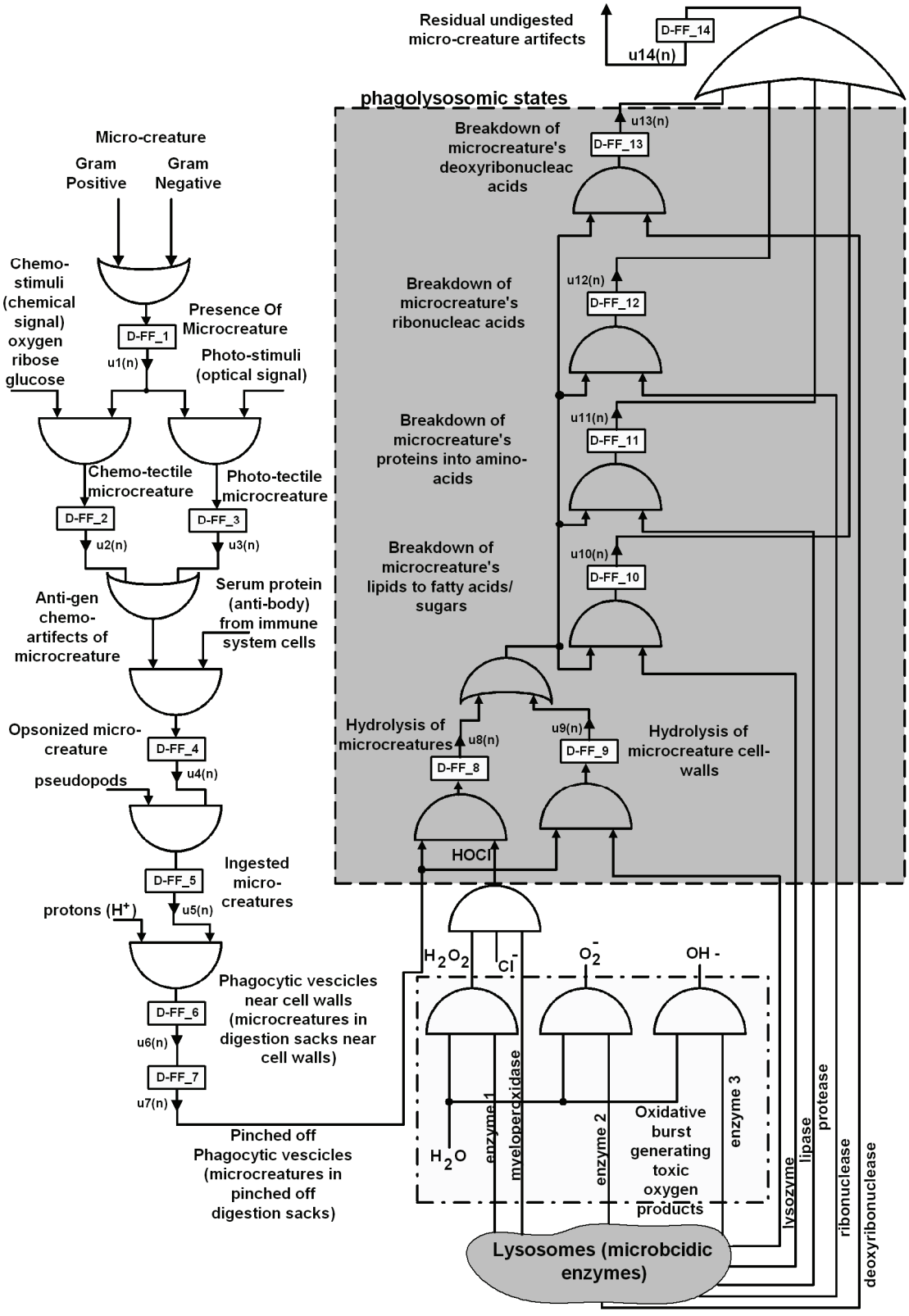


Figure 2. A micro-sequenced digital CMOS model of the bio-chemical pathway of phagocytosis using logic gates and D-flip-flops.

bio-chemical events form a pipeline through the *silicon mimetic* model of phagocytosis in a digital CMOS micro-sequenced logic simulation process. It also shows the sequence of intermediate states (D flip-flop states) corresponding to whether a certain bio-chemical state is activated. The sequence of molecular interactions is considered to be taking place in the circadian time-scale (hours or few minutes). The D flip-flops are thus driven by a clock with circadian time period. **Figure 3** shows molecular events occurring over 24 circadian time peri ods during which micro-creatures (bacteria or microbes) arrive towards phagocytes until the period n = 8. This

A micro-step sequence for phagocytosis

Path-way stste	Circadian Time-instant											
	n = 0	n = 1	n = 2	n = 3	n = 4	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10	n = 11
D-FF_1	RESET	microb	microb	microb	microb	microb	microb	microb	microb	0	0	0
D-FF_2	RESET	0	1	1	1	1	1	1	1	1	0	0
D-FF_3	RESET	0	1	1	1	1	1	1	1	1	0	0
D-FF_4	RESET	0	0	1	1	1	1	1	1	1	1	0
D-FF_5	RESET	0	0	0	1	1	1	1	1	1	1	1
D-FF_6	RESET	0	0	0	0	1	1	1	1	1	1	1
D-FF_7	RESET	0	0	0	0	0	1	1	1	1	1	1
D-FF_8	RESET	0	0	0	0	0	0	1	1	1	1	1
D-FF_9	RESET	0	0	0	0	0	0	1	0	1	1	1
D-FF_10	RESET	0	0	0	0	0	0	0	0	1	1	1
D-FF_11	RESET	0	0	0	0	0	0	0	0	1	1	1
D-FF_12	RESET	0	0	0	0	0	0	0	0	1	1	1
D-FF_13	RESET	0	0	0	0	0	0	0	0	1	1	1
D-FF_14	RESET	0	0	0	0	0	0	0	0	1	1	1

(a)

A micro-step sequence for phagocytosis(contd.)

Path-way stste	Circadian Time-instant											
	n = 12	n = 13	n = 14	n = 15	n = 16	n = 17	n = 18	n = 19	n = 20	n = 21	n = 22	n = 23
D-FF_1	0	microb	microb	microb	microb	0	0	0	0	0	0	0
D-FF_2	0	0	1	1	1	1	0	0	0	0	0	0
D-FF_3	0	0	1	1	1	1	0	0	0	0	0	0
D-FF_4	0	0	0	1	1	1	1	0	0	0	0	0
D-FF_5	0	0	0	0	1	1	1	1	0	0	0	0
D-FF_6	1	0	0	0	0	1	1	1	1	0	0	0
D-FF_7	1	1	0	0	0	0	1	1	1	1	0	0
D-FF_8	1	1	1	0	0	0	0	1	1	1	1	0
D-FF_9	1	1	1	0	0	0	0	1	1	1	1	0
D-FF_10	1	1	1	1	0	0	0	0	1	1	1	1
D-FF_11	1	1	1	1	0	0	0	0	1	1	1	1
D-FF_12	1	1	1	1	0	0	0	0	1	1	1	1
D-FF_13	1	1	1	1	0	0	0	0	1	1	1	1
D-FF_14	1	1	1	1	1	0	0	0	0	1	1	1

(b)

Figure 3. Sequence of intermediate steps (D flip-flop states) in the phagocytic bio-chemical pathway, (a) for circadian time-instants n = 0 to n = 11; and (b) for circadian time-instants n = 12 to n = 23.

causes the immune response pathway to activate. This results in the presence of either chemo-tectile (D-FF_2) or photo-tectile (D-FF_3) micro-creatures during the circadian interval <2,9>. Also, opsonized microcreatures are present during the interval <3,10>, microcreatures are ingested during the interval <4,11>, digestive vesicles are formed near the phagocyte membrane during the interval <5,12>, vesicles are detached and moves into the cytoplasm during the interval <6,13>, phagolysosome is formed and the microcreature is hydrolyzed (state in D-FF_8) and microcreature cell-wall is hydrolyzed (state in D-FF_9) during the interval <7,14>. The breakdown of lipids (state in D-FF_10), proteins (state in D-FF_11), ribonucleic acid (state in D-FF-12) and deoxyribonucleic acid (state in D-FF_13) take place simultaneously in the interval <8,15>. The undigested micro-creature artifacts (state in D-FF_14) are removed during the interval <9,16>. As the micro-creatures are is not present during the interval <9,12>, there is a short pause in the digestive bio-chemical process which is evident from the diagonal zero states in **Figure 3(b)**. Also, beginning the interval <13,16> the presence of a short burst of microcreatures results in a short pipeline of digestive pathway activities (reactions and interactions) that is evident through the diagonal array of "1" as shown in **Figure 3(b)**.

5. CONCLUSIONS

A digital circuit based model for the phagocytic biochemical digestive pathway in living systems has been developed and discussed in detail. The model corresponds quite well with the immune response phenomenon depicting striking resemblance of CMOS logic circuit (with states in D flip-flops) to states in bio-cellular phenomenon. Compared to mathematical modeling, model derived from analogies with integrated circuit allows VLSI CAD circuit and logic simulators to be conveniently used as a *biological simulation program*. Hence this work provides an alternative route for further systems biological investigation into more comprehensive integrated circuit models for more extensive biochemical pathways in living systems. This investigation will thus contribute to the desired manipulation of biological processes at the cellular level leading to electrical circuit modeling of diseases and "*nano-medicine*".

REFERENCES

[1] Dhar, P.K., Zhu, H. and Mishra, S.K. (2004) Computational approach to systems biology: From fraction to integration and beyond. *IEEE Transactions on Nanobioscience*, **3(3)**, 144-152.

[2] Khammash, M. and El-Samad, H. (2004) Systems boilogy: From physiology to gene regulation. *IEEE Control Systems Magazine*, **24(4)**, 62-76.

[3] Vaidyanathan, P.P. (2004) Genomics and proteomics: A signal processor's tour. *IEEE Circuits and Systems Magazine*, **4(4)**, 6-29.

[4] Campbell, N.A. and Reece, J.B. (2005) *Biology.* Pearson, San Francisco.

[5] Tortora, G.J., Funke, B.R. and Case, C.L. (2007) Microbiology: An introduction. Pearson Benjamin Cummings, San Francisco.

[6] Watson, J.D. and Crick, F.H.C. (1953) A structure for DNA. *Nature*, **171(4356)**, 737-738.

[7] Goldbeter, A. (1995) A model for circadian oscillations in the drosophila period protein. *Proceedings of Biological Sciences*, **261(1362)**, 319-324.

[8] Chen, L. and Wang, R. (2006) Designing gene regulatory networks with specified functions. *IEEE Transactions on Circuits and Systems*-I: *Regular Papers*, **53(11)**, 2444-2450.

[9] Hasan, S.M.R. (2005) A novel CMOS integrated circuit model for cellular DNA-protein regulatory mRNA transcription process. *Proceedings of* 12*th International Conference on Biomedical Engineering, International Federation for Medical and Biological Engineering CD proceedings*, Singapore.

[10] Rezaul Hasan, S.M. (2005) A novel integrated circuit model for mRNA transcription in bio-cellular processes," *Proceedings of* 12*th Electronics New Zealand Conference*, Manukau, 7-12.

[11] Hasan, S.M.R. (2008) A novel mixed-signal integrated circuit model for DNA-protein regulatory genetic circuits and genetic state machines. *IEEE Transactions on Circuits and Systems*-I: *Regular Papers*, **55(5)**, 1185-1196.

[12] Hasty, J., McMillen, D. and Collins, J.J. (2002) Engineered gene circuits. *Nature*, **420(14)**, 224-230.

[13] Simpson, M.L., Cox, C.D., Peterson, G.D. and Sayler, G.S. (2004) Engineering in the biological substrate: Information processing in genetic circuits. *Proceedings of the IEEE*, **92(5)**, 848-863.

[14] Hasan, S.M.R. and Ula, N. (2008) Analog CMOS charge model for molecular redox electron-transfer reactions and bio-chemical pathways. *Proceedings of International Symposium on Circuits and Systems*, Geneva.

[15] Rezaul Hasan, S.M. (2008) A Micro-sequenced CMOS model for cell signaling pathway using G-protein and phosphorylation cascade. *International Journal of Intelligent Systems Technologies and Applications*, **19(12)** 57-62.

[16] Schiek, R.L. and May, E.E. (2003) Development of a massively-parallel, biological circuit simulator. *Proceedings of the Computational Systems Bioinformatics*, Cambridge, 620-622.

[17] K. Martin, (2000) Digital integrated circuit design. Oxford University Press, Oxford.

J. Biomedical Science and Engineering, 2010, 3, 476-483 **JBiSE**
doi: 10.4236/jbise.2010.35066 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Inhibition of mammalian target of rapamycin induces phenotypic reversion in three-dimensional cultures of malignant breast epithelial cells

Ross Booth, Soonjo Kwon*, Eric Monson

Department of Biological Engineering, Utah State University, Logan, USA.
Email: soonjo.kwon@usu.edu

Received 27 February 2010; revised 11 March 2010; accepted 17 March 2010.

ABSTRACT

Inhibition of mammalian target of rapamycin (mTOR) is a potential method for cancer treatment. Effects of rapamycin (RAP) on the reversion of malignant breast epithelial cells were investigated on three-dimensional (3D) basement membrane extract (BME) cultures. Through continuous exposure to 20 nM of RAP, cell colony size was significantly reduced in 3D BME cultures of malignant breast epithelial cells, while normal cell colony size appeared unaffected. In unfixed 3D BME cultures of normal and RAP-treated malignant breast epithelial cells, the presence of luminal cell death was confirmed by ethidium bromide and propidium iodide labeling. Increased structural organization was observed by im- munofluorescence staining of F-actin and β-catenin in RAP-treated malignant breast epithelial cells. In monolayer cultures of normal and malignant breast epithelial cells, continuous exposure to 20 nM of RAP increased caspase 3/7 activity and decreased proliferation. Reverse transcriptase polymerase chain reaction (RT-PCR) array analysis indicated a fold increase in the expression of a number of proteins related to polarity, cell-cell adhesion, and cell-matrix adhesion in the presence of RAP. Our data showed that phenotypic reversion of malignancy can be achieved through RAP exposure on 3D BME cultures. This 3D BME culture system will provide correct microenvironments for observing the effects of other mTOR inhibitors on phenotypic reversion of malignant breast epithelial cells.

Keywords: Rapamycin; Three-Dimensional Culture; Breast Cancer Reversion; Basement Membrane Extract; mTOR Inhibitors

1. INTRODUCTION

The use of traditional two-dimensional (2D) cultures for breast cancer studies provides abundant fundamental information on the regulation of breast cancer, but its results often fail to properly translate into clinical applications [1]. The need to effectively address this problem has led to a demand for culture systems which more accurately represent the tissue environment found *in vivo*. In recent years, three-dimensional (3D) culture systems have been developed which use extra-cellular matrix (ECM) proteins [2,3]. 3D cultures provide epithelial cells with spatial freedom to allow the formation of 3D structures. Furthermore, the presence of reconstituted basement membrane extract (BME) promotes polarization of breast epithelial cells into acinar structures with luminal cell death through communication with membrane-bound integrins [3-7]. BME products are derived from the Engelbreth-Holm Swarm (EHS) tumor and are rich in laminin-1, with other ECM components including procollagen IV, heparin sulfate proteoglycan, and entactin [8]. However, most breast cancer cell lines fail to exhibit a normal phenotype in BME cultures unless accompanied with one or more signal molecules [9]. Several signal molecules have previously been shown to affect phenotypic reversion through limited cell proliferation, increased multi-cellular organization, basolateral polarization, and programmed luminal cell death. These reversion factors include inhibitors of β1-integrin, epidermal growth factor receptor (EGFR), and the MAPK & PI3K kinases [10-12]. Such reversion studies have provided valuable insight for potential methods of cancer treatment.

Rapamycin (RAP), the primary inhibitor for mammalian target of rapamycin (mTOR), has been used for its immunosuppressive properties under the drug name Sirolimus [13]. However, the mTOR signaling pathway has been shown to be a component of several signaling pathways related to breast cancer, including Akt [14], PTEN [15], and PI3K [16]. The web of signaling pathways related to mTOR is complex, and its malignant effects include increased cell growth, proliferation, and

resistance to apoptosis [17]. Presently, RAP and its analogues are under investigation as potential cancer treatments [15]. In breast epithelial cells transformed to exhibit hyperactivity of the Akt kinase, RAP exposure inhibits morphological disruptions caused by Akt overactivity [18]. However, the effects of mTOR inhibition on breast cancer cell lines cultured in 3D BME are not well characterized.

To provide further insight into the value of mTOR inhibition as a potential cancer treatment, we observed and analyzed the phenotypic reversion of malignant breast epithelial cells in 3D BME following exposure to RAP. Two malignant breast epithelial cells (BT-483 and MDA-MB-468) were exposed to RAP and compared to normal breast epithelial cells (184-B5). Changes in multi-cellular morphology, cell proliferation, polarity, and gene expression were investigated in the presence of RAP on 3D BME cultures.

2. MATERIALS AND METHODS

2.1. Cell Culture

Malignant human breast epithelial cells (BT-483 and MDA-MB-468) and normal human breast epithelial cells (184-B5) were purchased from ATCC (VA). For monolayer cultures: BT-483 was cultured using RPMI-1640 (Sigma-Aldrich, MO) with 20% of fetal bovine serum (FBS) (Thermo Fisher Scientific, UT) and 0.01 mg/mL of bovine insulin, MDA-MB-468 was cultured using Leibovitz's L-15 (Sigma-Aldrich) with 10% of FBS in air with no added CO_2, and 184-B5 was cultured using MEGM (Lonza, Switzerland) with 1 ng/mL of cholera toxin. To achieve luminal morphology, cells were cultured using 3D on-top method as previously described [2,3]. 3D culture matrix BME (phenol red free), with reduced growth factors, was obtained from Trevigen (MD). The 3D culture media were modified to contain 2-10% of BME, with 2% of FBS for malignant cells with or without EGF (epidermal growth factor) supplements. All media were supplemented with 1% of Penicillin/Streptomycin (Invitrogen, CA) and 0.1% of Fungizone (Hyclone, UT). Cells were seeded at 0.25×10^5 (184-B5), 0.22×10^5 (BT-483), and 0.18×10^5 (MDA-MB-468) cells/cm^2. Media were changed every two days subsequent to seeding. RAP (Sigma-Aldrich, MO) was reconstituted in DMSO and stored at -20°C at a stock concentration of 1 μM.

2.2. Morphological Observation

Morphology was observed using a TE2000-S Eclipse Microscope (Nikon, Japan) or a TS100 Eclipse Microscope (Nikon, Japan). All images were taken with a Digitial Sight DSQi1Mc Camera (Nikon, Japan). Fluorescence images were detected with an Intensilight C-HGFI lamp (Nikon, Japan). For the detection of non-viable (apoptotic) cells, non-fixed cultures were incubated for 15-30 minutes with 1 μM of ethidium bromide or 500 nM of propidium iodide and observed for fluorescence. Cell colonies were extracted using PBS-EDTA solution, fixed using 4% of paraformaldehyde, and blocked with 10% of goat serum and 1% of goat F (ab') IgG (Sigma-Aldrich, MO). Monoclonal anti-actin (1A4) and monoclonal anti-β-catenin (15B8) were obtained from Sigma-Aldrich (St. Louis, MO) for use as primary antibodies. Secondary antibody for FITC-conjugated goat anti-mouse was obtained from Santa Cruz Biotechnology (CA). Image analysis was conducted using NIS-Elements v3.0 (Nikon). Areas of cell colonies (in the x-y plane) in all images were analyzed for the calculation of mean areas under different conditions.

2.3. Caspase 3/7 Assay

Monolayers of BT-483 and 184-B5 cells were prepared in a 96-well plate to analyze the effects of RAP on caspase activity. Caspase activity was measured at 24 and 48 hours following exposure to RAP at 0, 5, and 20 nM. One hundred μL of Caspase-Glo 3/7 Reagent (Promega, WI) was added to each sample, including two blank wells containing 100 μL of growth media, and incubated for 2 hours. Luminescence was then measured using a Synergy 4 Plate Reader (Biotek, VT).

2.4. Cell Viability Assay

CellTiter-Blue Cell Viability Assay reagent was obtained from Promega (WI). MDA-MB-468 and 184-B5 cells were seeded in 96-well plates, beginning on day 2, and were treated continuously with 0 nM and 20 nM of RAP. On day three and day seven, cell viability assays were conducted and measured for fluorescence using a Synergy 4 Plate Reader (Biotek, VT).

2.5. Real-Time RT-PCR Arrays

Real time RT-PCR (reverse transcriptase polymerase chain reaction) Arrays for human extracellular matrix and adhesion molecules were obtained from SABiosciences (MD). MDA-MB-468 cells cultured in 3D BME were treated with 0 nM and 20 nM of RAP for 6 days. Cell colonies were extracted using PBS-EDTA, mRNA was extracted using RNeasy RNA isolation kit (QIAGEN, CA), and cDNA synthesis was conducted using the Reverse Transcriptase First Strand Kit (SABiosciences, MD). Real time RT-PCR arrays were run using a Bio-Rad Opticon 2 thermocycler system (MJ Research, Canada) and analyzed using the Opticon 2 software provided. C_t values were acquired graphically.

2.6. Statistical Analysis

With the SYSTAT 12 software, statistical analyses on all quantitative data were carried out using ANOVA (ana-

lyses of variance) to verify statistical significance ($p < 0.05$).

3. RESULTS

3.1. Effects of EGF on Phenotype of Normal Breast Epithelial Cells

Correct signaling from growth factors (EGF) and ECM proteins (BME) contribute to the development and maintenance of a hollow luminal structure within normal human breast epithelial acini, which is schematically described in **Figure 1**. Cells seeded on BME proliferated to form acini. Several events contributing to the formation of a luminal structure preceded the suppression of proliferation within the normal breast epithelial acini. Two populations of cells within each acinus became evident from day 6 to 8:1) An outer layer of cells that was in direct contact with the matrix and 2) an inner subset of cells that underwent luminal death (apoptosis) (**Figure 1(c)**) and that lacked contact with the matrix. Cells in contact with BME were apparently protected from a global death signal. Normal (184-B5) breast epithelial cells were initially cultured in a MEGM-based medium containing epidermal growth factor (EGF) additive. Resulting cultures contained large, morphologically disrupted structures with no apparent limitations to growth. Culturing the cells in EGF-reduced media resulted in the formation of significantly smaller structures with uniformity in shape (**Figure 2(a)**). The mean area for cell colonies in EGF-reduced media at day 7 was significantly lower than those in EGF-enhanced media at day 7 (**Figure 2(b)**).

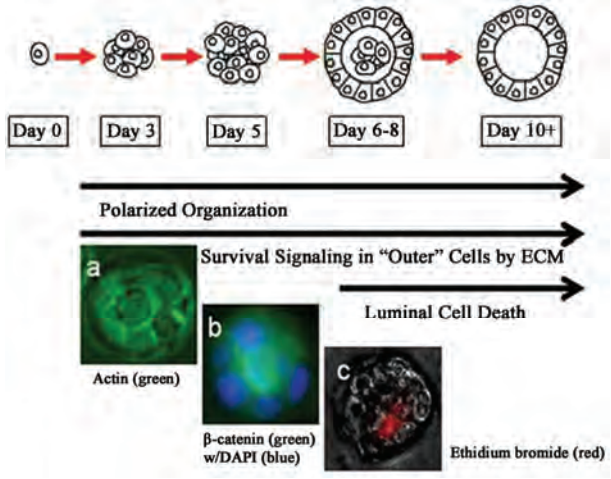
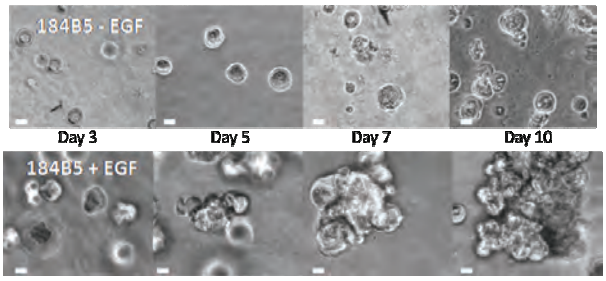
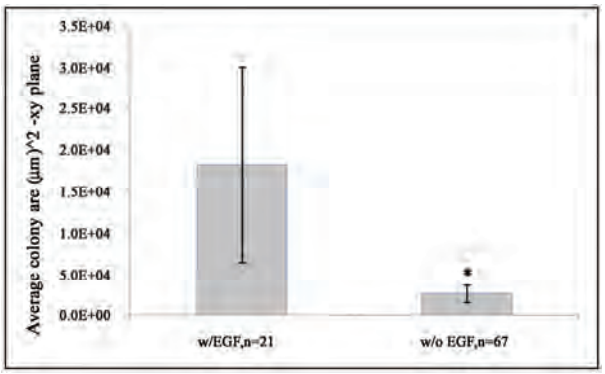


Figure 1. Schematic of biological events during normal human breast epithelial cells' acinar morphogenesis. Normal human breast epithelial cells were cultured using the 3D on-top method. Cell colonies were immunostained against. (a) actin (green); (b) β-catenin (green); (c) ethidium bromide (red). Nuclei were counterstained with DAPI (blue) in (b). Experimental procedures were adapted from [2,3].



(a)



(b)

Figure 2. (a) Effects of EGF-reduction on morphology of normal human breast epithelial cells (184-B5) in 3D BME cultures with and without EGF. Top: EGF-reduced media induced a normal phenotype, including smaller, round morphology with maximum size at day 5 of culture. Bottom: EGF-enhanced media led to increased proliferation and disrupted morphology. (Scale bars = 25 μm). (b) Effects of EGF on average cell colony size: Average cell colony areas were measured from normal human breast epithelial cells (184-B5) in 3D BME cultures with and without EGF on day 7 of culture. Colonies in EGF-reduced media showed a significantly lower average area than the colonies in EGF-enhanced media. (Error bars ± st. dev, * significantly lower than control, $p < 0.05$).

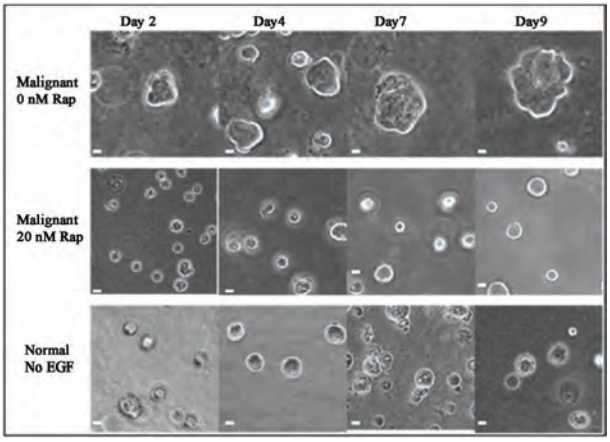
3.2. Effects of RAP Exposure on Malignant Cell Morphology

Morphologies of normal and malignant breast epithelial cells (BT-483 and MDA-MB-468) were compared in 3D BME cultures. Malignant breast epithelial cells did not form organized structures, had no uniformity in shape, and continued to proliferate. Malignant cell colonies exposed to 20 nM of RAP, starting on day 2, formed significantly smaller colonies with a maximum size reached at 6-8 days (**Figure 3(a)**). The mean area for RAP-treated malignant cell colonies was significantly reduced than the untreated malignant cell colonies on day 7 (**Figure 3(b)**).

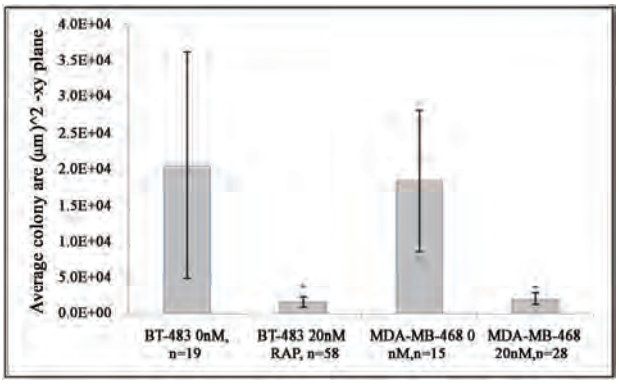
An inner subset of cells undergoing luminal cell death (apoptosis) was observed with ethidium bromide and propidium iodide staining. Apoptotic colonies were commonly found at the center of colonies in normal

(184-B5) breast epithelial cell cultures on day 8, which indicated luminal formation. This was also observed in RAP-treated malignant (BT-483 & MDA-MB-468) cell colonies, providing evidence that the treated malignant breast epithelial cells were reverting to normal phenoltypes (**Figure 4(a)**).

Structural organization was also observed using anti-actin and anti-β-catenin immunofluorescence staining (**Figure 4(b)**). RAP-treated malignant colonies exhibited cells arranged around the center, showing similarities to the structure of normal breast epithelial cells in 3D BME

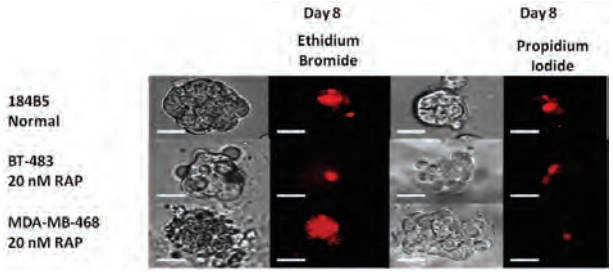


(a)

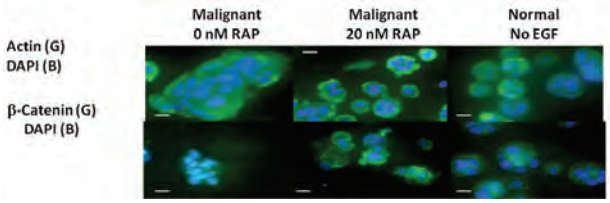


(b)

Figure 3. (a) Comparison of normal, malignant, and RAP-treated malignant human breast epithelial cells in 3D BME culture. RAP was exposed to BT-483 malignant cells beginning on the second day of culture at 0 and 20 nM, and compared to 184-B5 normal cells. Normal cells (bottom) were compared in colony shape and size to RAP-treated malignant cells (middle**)**. Malignant cells (top**)** exhibited severe morphological disruption when cultured without treatment. (Scale bars = 25 μm); (b) Effects of RAP exposure on average cell colony size: Average cell colony areas were measured from BT-483 and MDA-MB-468 malignant cells treated with 0 and 20 nM of RAP on day 7 of 3D BME culture. No significant difference was found between the two cell lines, but RAP treatment significantly reduced cell colony size of malignant cells. (Error bars ± st. dev, * significantly lower than control, $p < 0.05$).



(a)



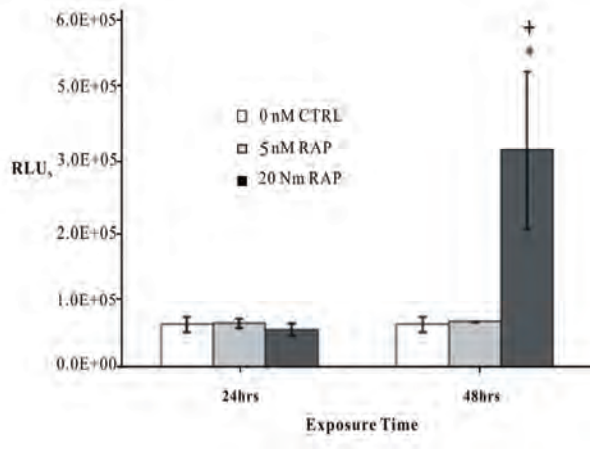
(b)

Figure 4. (a) Selective staining of non-viable cells. Top: Staining of day 8 3D cultures of normal (184-B5) human breast epithelial cells with ethidium bromide or propidium idodie resulted in detection of cell death at the center of cell colonies. Bottom: Similar occurences of centrally-localized apoptosis were detected in both malignant cell lines when treated continuously with 20 nM of RAP; (b) Immunofluorscence staining of fixed cell colonies from 3D culture. Malignant (BT-483) cell colonies from day 7 of culture, after continuous treatment of 20 nM RAP from day 2, exhibited multi-cellular organization similar to day 7 cultures of normal (184-B5) human breast epithelial cells. Top: Anti-Actin (green) counterstained with DAPI (blue). Bottom: Anti-β-Catenin (green) counterstained with DAPI (blue). (Scale bars = 25 μm)

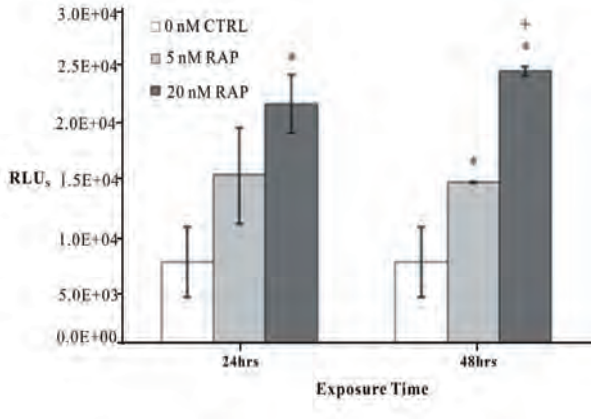
cultures. Exposure to up to 20 nM of RAP had no apparent effect on the morphology of either EGF- reduced or EGF-enhanced cultures of normal (184-B5) cells (Data not shown).

3.3. Effects of RAP Exposure on Caspase Activity and Proliferation

Confluent monolayers of normal (184-B5) and malignant (BT-483) breast epithelial cells were exposed to RAP at 0, 5, and 20 nM. Caspase activity was measured after 24 and 48 hours of exposure (**Figure 5**). For normal breast epithelial cells, the only condition which showed significant deviation from the control was 20 nM of RAP after 48 hours of exposure (**Figure 5(a)**). For malignant breast epithelial cells, 20 nM of RAP exposure showed a significantly higher effect on caspase activity than 5 nM, but no significant difference was observed with respect to exposure time (**Figure 5(b)**). An increase in caspase activity indicates the presence of programmed cell death, which is essential for the formation of the luminal space in 3D BME cultures.



(a)



(b)

Figure 5. Effect of RAP exposure on caspase activity. (a) Normal human breast epithelial cell monolayers: Caspase activity was measured on samples of 184-B5 exposed to RAP at 5 and 20 nM for 0, 24, and 48 hours following confluency. The only condition which yielded a significant difference from the control group was 20 nM of RAP at 48 hours of exposure. (n = 2, * significantly higher than control, + significantly higher than 5 nM, $p < 0.05$); (b) Malignant human breast epithelial cell monolayers: Caspase activity was measured on samples of BT-483 cells exposed to RAP at 5 and 20 nM for 0, 24, and 48 hours following confluency. While there was no significant difference between 24 and 48 hours exposure, all treated samples were significantly higher than the control (n = 2, * significantly higher than control, + significantly higher than 5 nM, $p < 0.05$).

RAP's effect on cell proliferation was analyzed in monolayers of normal (184-B5) and malignant (MDA-MB-468) breast epithelial cells (**Figure 6**) after 2 and 6 days of RAP exposure. Both cell lines showed significantly decreased total cell viability when exposed to 20 nM of RAP, with a significantly higher reduction seen on day 7 (not significant at 5 nM of RAP; data not shown).

3.4. Effects of RAP Exposure on Gene Expression of Extracellular Matrix and Adhesion Molecules

RT-PCR arrays were prepared from malignant breast epithelial cells (MDA-MB-468) in 3D BME cultures containing 0 and 20 nM of RAP on day 6 of culture. Quality controls included with the arrays verified successful RT-PCR and the absence of genomic DNA contamination. C_t values were cut off at > 35. Genes were grouped into six major classifications, and reported with expression fold change (DDCT method) of untreated controls compared to RAP treated cells (**Table 1**). In RAP-treated samples, 25 genes related to ECM and adhesion molecules showed a significant (at least 5-fold) increase in expression, while no genes showed a significant decrease in expression (**Table 1**).

4. DISCUSSION

In this study, the effects of RAP exposure on the phenotypic reversion of malignant breast epithelial cells were investigated in 3D BME culture. Microscopic observation of malignant breast epithelial cells treated with RAP showed greater similarities in morphology to normal cells than in untreated malignant breast epithelial cells in 3D BME cultures. Comparison of the average horizontal area confirmed that RAP-treated malignant cell colonies are significantly smaller in size. Increased structural organization was observed by immunofluorescence staining of anti-Actin and anti-β-Catenin. On day 8 of 3D BME culture, RAP-treated malignant breast epithelial

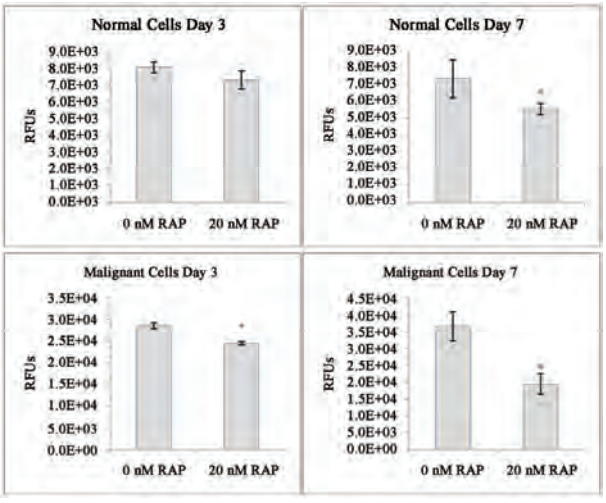


Figure 6. Effect of RAP exposure on total cell viability of normal and malignant human breast epithelial cell monolayers. Cultures of MDA-MB-468 and 184-B5 were exposed to RAP at 0 and 20 nM beginning on the second day of culture. Total cell viability was measured on day 3 and day 7. Both cell lines showed a significant decrease in proliferation at day 7 (n = 3, * significantly lower than control, $p < 0.05$).

Table 1. RAP's effects on gene expression of human extracellular matrix and adhesion molecules. Fold change was calculated using the DDCT method, with GADPH as the normalizing gene.

Gene Category	Symbol	Description	Fold Change
Basement Membrane Components	LAMA2	Laminin, alpha 2	13.77
	LAMA3	Laminin, alpha 3	19.05
	LAMB3	Laminin, beta 3	12.57
	SPARC	Secreted protein, acidic, cysteine-rich	13.64
Cell-Cell Adhesion	CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	6.40
	CTNND1	Catenin (cadherin-associated protein), delta 1	7.13
	CD44	CD44 molecule (Indian blood group)	6.95
	ITGA2	Integrin, alpha 2	9.84
	ITGA6	Integrin, alpha 6	5.26
	ITGAV	Integrin, alpha V (vitronectin receptor)	8.99
	ITGB1	Integrin, beta 1 (fibronectin receptor)	14.30
	ITGB2	Integrin, beta 2	9.27
	ITGB5	Integrin, beta 5	5.48
	SGCE	Sarcoglycan, epsilon	5.25
	COL1A1	Collagen, type I, alpha 1	5.79
	COL4A2	Collagen, type IV, alpha 2	9.77
	FN1	Fibronectin 1	16.51
Other ECM Molecules	THBS1	Thrombospondin 1	16.02
	COL12A1	Collagen, type XII, alpha 1	17.50
	CTNNA1	Catenin (cadherin-associated protein), alpha 1	7.82
	CTNNB1	Catenin (cadherin-associated protein), beta 1	11.40
	LAMC1	Laminin, gamma 1 (formerly LAMB2)	6.36
Other ECM Molecules	TGFBI	Transforming growth factor, beta-induced	10.04
	VCAN	Versican	17.70
Trans-membrane Molecules	ITGA4	Integrin, alpha 4	5.87

cells stained with ethidium bromide and propidium iodide showed similar structural organization with luminal cell death to that observed in normal breast epithelial cells, indicating re-polarization of the epithelial cell layers. Supporting data from monolayer cultures indicated that levels of caspase activity were increased in the presence of RAP, and that proliferation over time was decreased. RT-PCR array analysis of RAP-treated malignant breast epithelial cells in 3D BME culture indicated an increase in the expression of several ECM and adhesion molecules important to polarity and structural organization. This collection of data supports our hypothesis that RAP exposure can revert malignant breast epithelial cells to a normal phenotype.

In order to understand why breast cancer develops, as well as to predict the outcome of pharmacological treatments, we need to model the structure and function of organs or tissues in culture so that our experimental manipulations occur under physiological contexts. 3D cultures provide a model to analyze the spatial and temporal aspects of key biological processes (e.g., proliferation and apoptosis) and signal transduction events during morphogenesis. Many of these studies are impractical or impossible to perform in native breast tissue. Cultured normal human breast epithelial cell monolayers in the absence of BME fail to assemble organized structures, and arrest growth when they reach confluence (**Figure 7(a)**). We recently found that signaling from epidermal growth factor (EGF) and ECM proteins contributes to the development and maintenance of a hollow luminal

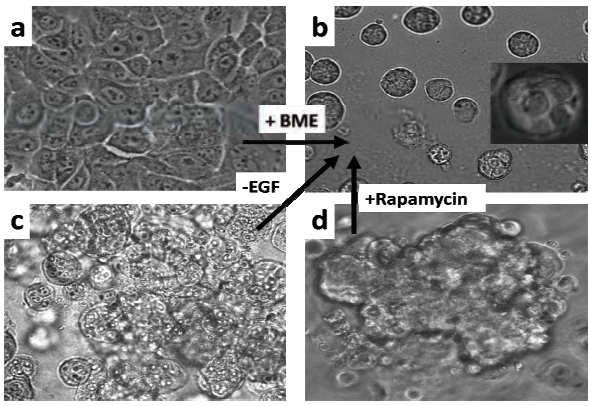


Figure 7. Comparison of normal and malignant human breast epithelial cells grown in 2D monolayer and 3D culture matrix. Normal human breast epithelial cells grown in the 3D BME culture form a luminal structure with normal phenotype (images not to scale). (a) Normal human breast epithelial cells in monolayer without BME (2D); (b) Normal human breast epithelial cells in 3D BME with reduced EGF; (c) Normal human breast epithelial cells in 3D BME with EGF; (d) Malignant human breast epithelial cells in 3D BME with reduced EGF. Phenotypic reversion of malignancy can be achieved through RAP exposure on EGF-reduced 3D BME cultures.

structure within normal breast epithelial acini. In our 3D culture system, non-malignant breast epithelial cells seeded at low density developed into polarized acinar units when EGF is withdrawn. Two populations of cells within each acinus became evident from day 6 to 8: An outer layer of cells that was in direct contact with the ECM, and an inner set of cells not in contact with the matrix that underwent apoptosis (**Figures 1** and **7(b)**). This emulated the ductal structures of breast epithelial cells found in vivo. In the presence of EGF, normal breast epithelial cells exhibited severe morphological deformities, including colony overgrowth, luminal filling, and resistance to apoptosis (**Figure 7(c)**). It was observed that EGF disrupted the formation of luminal structures of normal breast epithelial cells either in monolayer or 3D cultures with BME (**Figures 7(a)** and **7(c)**). We also found evidence that rapamycin and its derivative can revert cells' structure from disorganized to organized (**Figures 3**, **4**, and **7**).

Our criteria for phenotypic reversion are: increased structural organization, limited colony size/proliferation, and the presence of basolateral polarity. Immunofluorescence staining of F-actin and -catenin showed a much higher occurrence of morphologically spherical colonies, indicating structural reorganization. This was also supported by the increase in expression of cell-cell adhesion molecules, transmembrane molecules, and ECM proteases shown by the RT-PCR array data (**Table 1**). Smaller mean area of RAP-treated cell colonies provided strong evidence of limited proliferation. A decreased number of total cell viability over time and increased level of caspase activity confirmed that RAP had a growth-limiting effect on cancer cells. The presence of centrally located cell death indicated the process of lumen formation, directed by caspase-induced programmed apoptosis, and was a result of polarization of the epithelial layers. Up-modulation of caspase, in addition to increased gene expression of basement membrane components and cell-matrix adhesion molecules such as integrins, gave evidence that RAP exposure promoted development of basolateral polarity (**Figure 8**).

In previous studies, RAP has been seen to prevent morphological disruption and inhibit runaway proliferation normally caused by Akt activation in breast epithelial cells (MCF-10A) [18], and to significantly inhibit growth of human lung cancer cells [14]. The value of inhibiting mTOR as a potential cancer treatment was further validated by this study. However, due to its poor water solubility and limited stability, analogues of RAP have been developed with improved pharmaceutical properties for parenteral delivery, including CCI-799 (tensirolimus), RAD-001 (everolimus), and AP-23573 (ARIAD) [15,19,20], which also inhibit mTOR. The effects of these analogues on the phenotypic reversion of malignant breast epithelial cells cultured in 3D BME

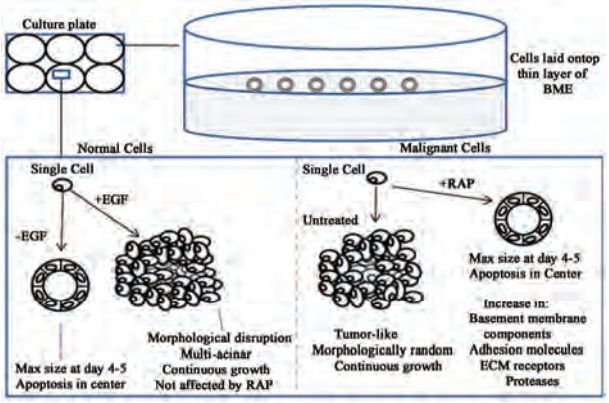


Figure 8. Summary of experimental results. With normal human breast epithelial cells, normal phenotype was achieved by removing EGF additive from the media, leading to limited proliferation, limited cell colony size, and luminal apoptosis. RAP was found to limit proliferation of malignant human breast epithelial cells and induce luminal apoptosis. RT-PCR array data indicated an increase in expression of cell-cell adhesion molecules, transmembrane molecules, and ECM proteases.

should also be studied. Synergistic or additive effects of RAP with other cancer drugs have been studied, including paclitaxel, carboplatin and vinorelbine [21]; mellerian inhibiting substance [22]; doxorubicin [23]; and gemcitabine [24]. Synergistic or additive effects of RAP with other drugs on the phenotypic reversion of malignant cells in 3D BME culture should also be investigated. This study can be used as an experimental template for such future investigation.

In conclusion, by showing that RAP can revert malignant breast epithelial cells to a normal phenotype, our study has provided supporting data for the value of mTOR inhibition as a potential anti-tumor strategy. To provide further insight, two further directions of research should be taken: Further studies on the effects of RAP's analogues on phenotypic reversion, and studies on the synergistic effects of mTOR inhibition with other cancer targets.

5. ACKNOWLEDGEMENTS

Funding for this project was provided by NIH Grant No. 1 R21 CA 131798-01A1. Special thanks to Dr. Timothy Doyle and Dr. David Britt for their contributions.

REFERENCES

[1] Xu, F. and Burg, K.J. (2007) Three-dimensional polymeric systems for cancer cell studies. *Cytotechnology*, **54(3)**, 135-143.

[2] Lee, G.Y., Kenny, P.A., Lee, E.H. and Bissell, M.J. (2007) Three-dimensional culture models of normal and malignant breast epithelial cells. *Nature Methods*, **4(4)**, 359-365.

[3] Debnath, J., Muthuswamy, S.K. and Brugge, J.S. (2003)

Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods*, **30(3)**, 256-268.

[4] Roskelley, C.D., Desprez, P.Y. and Bissell, M.J. (1994) Extracellular matrix-dependent tissue-specific gene expression in mammary epithelial cells requires both physical and biochemical signal transduction. *Proceedings of the National Academy of Sciences*, **91(26)**, 12378-12382.

[5] Bello-DeOcampo, D., Kleinman, H.K., Deocampo, N.D. and Webber, M.M. (2001) Laminin-1 and alpha6 beta1 integrin regulate acinar morphogenesis of normal and malignant human prostate epithelial cells. *Prostate*, **46(2)**, 142-153.

[6] Reginato, M.J. and Muthuswamy, S.K. (2006) Illuminating the center: Mechanisms regulating lumen formation and maintenance in mammary morphogenesis. *Journal of Mammary Gland Biology and Neoplasia*, **11(3-4)**, 205-211.

[7] Barcellos-Hoff, M.H., Aggeler, J., Ram, T.G. and Bissell, M.J. (1989) Functional differentiation and alveolar morphogenesis of primary mammary cultures on reconstituted basement membrane. *Development*, **105(2)**, 223-235.

[8] Kleinman, H.K., McGarvey, M.L., Liotta, L.A., Robey, P.G., Tryggvason, K. and Martin, G.R. (1982) Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma. *Biochemistry*, **21(24)**, 6188-6193.

[9] Nelson, C.M. and Bissell, M.J. (2005) Modeling dynamic reciprocity: Engineering three-dimensional culture models of breast architecture, function, and neoplastic transformation. *Seminars in Cancer Biology*, **15(5)**, 342-352.

[10] Liu, H., Radisky, D.C., Wang, F. and Bissell, M.J. (2004) Polarity and proliferation are controlled by distinct signaling pathways downstream of PI3-kinase in breast epithelial tumor cells. *Journal of Cell Biology*, **164**, 603-612.

[11] Wang, F., Hansen, R.K., Radisky, D., Yoneda, T., Barcellos-Hoff, M.H., Petersen, O.W., Turley, E.A. and Bissell, M.J. (2002) Phenotypic reversion or death of cancer cells by altering signaling pathways in three-dimensional contexts. *Journal of the National Cancer Institute*, **94(19)**, 1494-1503.

[12] Weaver, V.M., Petersen, O.W., Wang, F., Larabell, C.A., Briand, P., Damsky, C. and Bissell, M.J. (1997) Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies. *Journal of Cell Biology*, **137(1)**, 231-245.

[13] Sabers, C.J., Martin, M.M., Brunn, G.J., Williams, J.M., Dumont, F.J., Wiederrecht, G. and Abraham, R.T. (1995) Isolation of a protein target of the FKBP12-rapamycin complex in mammalian cells. *Journal of Biological Chemistry*, **270(2)**, 815-822.

[14] Sun, S.Y., Rosenberg, L.M., Wang, X., Zhou, Z., Yue, P., Fu, H. and Khuri, F.R. (2005) Activation of Akt and eIF4E survival pathways by rapamycin-mediated mammalian target of rapamycin inhibition. *Cancer Research*, **65(16)**, 7052-7058.

[15] Vignot, S., Faivre, S., Aguirre, D. and Raymond, E. (2005) mTOR-targeted therapy of cancer with rapamycin derivatives. *Annals of Oncology*, **16(4)**, 525-537.

[16] Heinonen, H., Nieminen, A., Saarela, M., Kallioniemi, A., Klefstrom, J., Hautaniemi, S. and Monni, O. (2008) Deciphering downstream gene targets of PI3K/mTOR/p70S6K pathway in breast cancer. *BMC Genomics*, **9(4)**, 348-359.

[17] Bjornsti, M.A. and Houghton, P.J. (2004) The TOR pathway: A target for cancer therapy. *Nature Reviews Cancer*, **4(5)**, 335-348.

[18] Debnath, J., Walker, S.J. and Brugge, J.S. (2003) Akt activation disrupts mammary acinar architecture and enhances proliferation in an mTOR-dependent manner. *Journal of Cell Biology*, **163(2)**, 315-326.

[19] Gadducci, A., Cosio, S. and Genazzani, A.R. (2006) Old and new perspectives in the pharmacological treatment of advanced or recurrent endometrial cancer: Hormonal therapy, chemotherapy and molecularly targeted therapies. *Reviews in Oncology/Hematology*, **58(2)**, 242-256.

[20] Figlin, R.A., Brown, E., Armstrong, A.J., Akerley, W. and Benson, A.B., Burstein, H.J., Ettinger, D.S., Febbo, P.G., Fury, M.G., Hudes, G.R., Kies, M.S., Kwak, E.L., Morgan, R.J., Mortimer, J.J., Reckamp, K., Venook, A.P., Worden, F. and Yen, Y. (2008) NCCN task force report: mTOR inhibition in solid tumors. *Journal of the National Comprehensive Cancer Network*, **6(Suppl 5)**, **S1-S20**, **S21-S22.**

[21] Steelman, L.S., Stadelman, K.M., Chappell, W.H., Horn, S., Basecke, J., Cervello, M., Nicoletti, F., Libra, M., Stivala, F., Martelli, A.M. and McCubrey, J.A. (2008) Akt as a therapeutic target in cancer. *Expert Opinion on Therapeutic Targets*, **12(9)**, 1139-1165.

[22] Pieretti-Vanmarcke, R., Donahoe, P.K., Pearsall, L.A., Dinulescu, D.M., Connolly, D.C., Halpern, E.F., Seiden, M.V. and MacLaughlin, D.T. (2006) Mullerian inhibiting substance enhances subclinical doses of chemotherapeutic agents to inhibit human and mouse ovarian cancer. *Proceedings of the National Academy of Sciences of the USA*, **103(46)**, 17426-17431.

[23] Avellino, R., Romano, S., Parasole, R., Bisogni, R., Lamberti, A., Poggi, V., Venuta, S. and Romano, M.F. (2005) Rapamycin stimulates apoptosis of childhood acute lymphoblastic leukemia cells. *Blood*, **106(4)**, 1400-1406.

[24] Mondesire, W.H., Jian, W., Zhang, H., Ensor, J., Hung, M.C., Mills, G.B. and Meric-Bernstam, F. (2004) Targeting mammalian target of rapamycin synergistically enhances chemotherapy-induced cytotoxicity in breast cancer cells. *Clinical Cancer Research*, **10(20)**, 7031-7042.

J. Biomedical Science and Engineering, 2010, 3, 484-487 **JBiSE**
doi:10.4236/jbise.2010.35067 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

The determination of acidity of the dilute solutions of weak multibasic organic acids

Elene Kvaratskhelia, Ramaz Kvaratskhelia

R. Agladze Institute of Inorganic Chemistry and Electrochemistry, Tbilisi, Georgia.
Email: elicko@mail.ru; ekvarats@yahoo.com

Received 12 February 2010; revised 1 March 2010, accepted 5 March 2010.

ABSTRACT

The new theoretical method for the accurate determination of acidity of dilute solutions of weak multibasic organic acids (which are widely used in medicine, pharmacology, various branches of industry and participate in important biological processes in living organisms) is suggested. The concepts of the contributions of the separate dissociation steps to the [H^+] value, x_m, are used for an analysis of complex equilibria of the processes of dissociation of these acids. The cases of weak dibasic and tribasic organic acids with the "overlapping" dissociation equilibria and a general case of weak multibasic acids, H_nA, are considered. From the conditions of equality of the concentrations of various ionized and non-ionized forms in the dilute solutions of weak multibasic organic acids the areas of dominance of these forms in connection with the corresponding x_m values are formulated.

Keywords: Dibasic Acids; Tribasic Acids; Hydrogen Ions Concentration; Equations

1. INTRODUCTION

Weak multibasic organic acids are widely used in medicine, pharmacology, chemical, food and cosmetic industries. Some of these acids participate in a series of important biological processes occurring in living organisms (for example, in the Krebs cycle). The majority of drugs are weak acids and/or bases. Their biopharmaceutical properties are directly connected with the dissociation constants and degrees of these compounds, consequently, with acidity of their solutions. The latter is the very factor which affects in physiological systems the rate at which the compound is able to diffuse across membranes and various obstacles, determines the acid-base homeostasis and enzyme kinetics in the cell and in the body. It is possible to say that an acidity of the weak multibasic organic acids determines, as a rule, all their useful (or harmful) properties.

Many weak multibasic organic acids have comparatively close values of the dissociation constants for the various steps; this fact causes their simultaneous participation in the determining the hydrogen ion concentration in solutions of these acids (*i.e.*, "overlapping" dissociation equilibria). In this paper a new theoretical method for determination of acidity of the dilute (0.0001-0.1 mol·dm^{-3}) solutions of such acids is suggested.

2. RESULTS AND DISCUSSION

2.1. Dibasic Acids

Dibasic acids form the most numerous group of weak multibasic organic acids with the "overlapping" equilibria effect. In dilute aqueous solutions the primary and secondary steps of dissociation are

$$H_2A \leftrightarrow H^+ + HA^- \quad (1)$$

$$HA^- \leftrightarrow H^+ + A^{2-} \quad (2)$$

In our previous communications [1,2] we have used the concepts of the contributions to the total hydrogen ion concentration, [H^+], being assigned to the primary and secondary dissociation steps, x_1 and x_2, such that $x_1 + x_2 = [H^+]$. The corresponding mass-action equations for both steps dissociation constants are

$$K_1 = \frac{[H^+](x_1 - x_2)}{c - x_1} F_1 = \frac{x_1^2 - x_2^2}{c - x_1} F_1 \quad (3)$$

$$K_2 = \frac{[H^+]x_2}{x_1 - x_2} F_2 = \frac{x_2(x_1 + x_2)}{x_1 - x_2} F_2 \quad (4)$$

where K_1 and K_2 are the thermodynamic dissociation constants, c is the total (analytical) concentration of acid, F_1 and F_2 are the quotients of the activity coefficients:

$$F_1 = \frac{f_{H^+} f_{HA^-}}{f_{H_2A}} \quad (5)$$

$$F_2 = \frac{f_{H^+} f_{A^{2-}}}{f_{HA^-}} \quad (6)$$

The values of the activity coefficients may be approximated by the Debye-Huckel equation:

$$\log_{10} f_i = -\frac{z_i^2 A\sqrt{I}}{1+a_i B\sqrt{I}} \quad (7)$$

where a_i is the cation-anion distance of closest approach, A and B are constants depending on the properties of water at given temperature, z_i is the charge of ion. The ionic strength is given by $I = x_1 + 2x_2$. The activity coefficient of undissociated acid is assumed to be unity.

According to the (3) and (4) the x_1 and x_2 values (and then their sum – the $[H^+]$ value) can be calculated successively by an iterative solution of two quadratic equations:

$$x_1 = \frac{1}{2}\left[-\frac{K_1}{F_1} + \sqrt{\left(\frac{K_1}{F_1}\right)^2 + 4\left(x_2^2 + \frac{K_1 c}{F_1}\right)}\right] \quad (8)$$

$$x_2 = \frac{1}{2}\left[-\left(\frac{K_2}{F_2} + x_1\right) + \sqrt{\left(\frac{K_2}{F_2} + x_1\right)^2 + \frac{4K_2 x_1}{F_2}}\right] \quad (9)$$

We suggest also the empirical equation for the fast approximate determination of the pH values of dilute (0.0001-0.01 mol·dm^{-3}) solutions of weak dibasic (and tribasic with the low K_3 values) organic acids:

$$pH = -1.489 + 0.8\,pK_1 - (1.185 - 0.14\,pK_1)\lg c \quad (10)$$

The maximum value of the relative error for this equation for a series of weak dibasic and tribasic organic acids with the pK_1 values in the interval: 2.5-5 does not exceed 5% (the relative error is the ratio of the difference between the approximate pH value and corresponding accurate value, divided by the approximate pH value, and converted to percent).

2.2. Tribasic Acids

In case of weak tribasic organic acids with the "overlapping" dissociation equilibria, the mass-action equations may be expressed as follows:

$$K_1 = \frac{[H^+](x_1 - x_2)}{c - x_1} F_1 = \frac{(x_1 + x_2 + x_3)(x_1 - x_2)}{c - x_1} F_1 \quad (11)$$

$$K_2 = \frac{[H^+](x_2 - x_3)}{x_1 - x_2} F_2 = \frac{(x_1 + x_2 + x_3)(x_2 - x_3)}{x_1 - x_2} F_2 \quad (12)$$

$$K_3 = \frac{[H^+]x_3}{x_2 - x_3} F_3 = \frac{(x_1 + x_2 + x_3)x_3}{x_2 - x_3} F_3 \quad (13)$$

where

$$F_1 = \frac{f_{H^+} f_{H_2A^-}}{f_{H_3A}} \quad (14)$$

$$F_2 = \frac{f_{H^+} f_{HA^{2-}}}{f_{H_2A^-}} \quad (15)$$

$$F_3 = \frac{f_{H^+} f_{A^{3-}}}{f_{HA^{2-}}} \quad (16)$$

and $I = x_1 + 2x_2 + 3x_3$

The x_1, x_2 and x_3 values (and then their sum – the $[H^+]$ value) can be calculated successively by an iterative solution of three quadratic equations:

$$x_1 = \frac{1}{2}\left[-\left(\frac{K_1}{F_1} + x_3\right) + \sqrt{\left(\frac{K_1}{F_1} + x_3\right)^2 + 4\left(x_2^2 + x_2 x_3 + \frac{K_1 c}{F_1}\right)}\right] \quad (17)$$

$$x_2 = \frac{1}{2}\left[-\left(\frac{K_2}{F_2} + x_1\right) + \sqrt{\left(\frac{K_2}{F_2} + x_1\right)^2 + 4\left(x_3^2 + x_1 x_3 + \frac{K_2 x_1}{F_2}\right)}\right] \quad (18)$$

$$x_3 = \frac{1}{2}\left[-\left(\frac{K_3}{F_3} + x_1 + x_2\right) + \sqrt{\left(\frac{K_3}{F_3} + x_1 + x_2\right)^2 + \frac{4K_3 x_2}{F_3}}\right] \quad (19)$$

2.3. Acids with the Higher Basicity

It is necessary at first to consider the general case of the weak multibasic organic acid H_nA with the "overlapping" dissociation equilibria. For this case we may write the equations connecting the values of x_1, x_2, x_3, … x_{n-1}, x_n with the concentrations of various anions:

$$x_1 = [H_{n-1}A^-] + [H_{n-2}A^{2-}] + [H_{n-3}A^{3-}] + \ldots + [HA^{(n-1)-}] + [A^{n-}] \quad (20)$$

$$x_2 = [H_{n-2}A^{2-}] + [H_{n-3}A^{3-}] + \ldots + [HA^{(n-1)-}] + [A^{n-}] \quad (21)$$

$$x_3 = [H_{n-3}A^{3-}] + \ldots + [HA^{(n-1)-}] + [A^{n-}] \quad (22)$$

$$x_{n-1} = [HA^{(n-1)-}] + [A^{n-}] \quad (23)$$

$$x_n = [A^{n-}] \quad (24)$$

In a general form for the m dissociation step we may write:

$$x_m = [H_{n-m}A^{m-}] + x_{m+1} \quad (25)$$

The total hydrogen ion concentration may be expressed as follows:

$$[\mathrm{H}^+] = \sum_{m=1}^{n} m[\mathrm{H}_{n-m}A^{m-}] = \sum_{m=1}^{n} x_m \quad (26)$$

The mass-action equation for the m dissociation step may be expressed by the following equations:

$$K_m = \frac{[H^+](x_m - x_{m+1})}{x_{m-1} - x_m} F_m = \frac{(x_m - x_{m+1})\sum_{m=1}^{n} x_m}{x_{m-1} - x_m} F_m \quad (27)$$

where

$$F_m = \frac{f_{H^+} f_{H_{n-m}A^{m-}}}{f_{H_{n-(m-1)}A^{(m-1)-}}} \quad (28)$$

The equation for an ionic strength may be written as follows:

$$I = \sum_{m=1}^{n} m x_m \quad (29)$$

In case of weak organic acids with the high (more than tribasic) basicity the conversion of (27) to the forms of (8-9, 17-19) leads to the very complicated expressions. That is why we suggest to solve the complicated problem of determining of acidity of the high-basic acids solutions by more simple method. Let us consider this method for the most difficult case of hexabasic mellitic (benzenehexacarboxylic) acid. First, assume that this acid can be treated as a tribasic acid (taking into account that the main contribution to the $[H^+]$ value is made by first three dissociation steps). Then the x_1, x_2 and x_3 values are determined successively by an iterative solution of (17-19), where the values of F_1, F_2 and F_3 were assumed to be unity. The obtained x_1, x_2 and x_3 values are then used for the determination of the initial estimate of $[H^+]$. Then, with the aid of $[H^+]$ value and the iterative solution of the following equations (which are obtained from (27) for the corresponding dissociation steps):

$$x_4 = \frac{K_4 x_3 + [H^+]F_4 x_5}{K_4 + [H^+]F_4} \quad (30)$$

$$x_5 = \frac{K_5 x_4 + [H^+]F_5 x_6}{K_5 + [H^+]F_5} \quad (31)$$

$$x_6 = \frac{K_6 x_5}{K_6 + [H^+]F_6} \quad (32)$$

the initial values of x_4, x_5 and x_6 are determined (assuming that F_4, F_5 and F_6 values to be unity). These values are used for a correction to the $[H^+]$ value:

$$[H^+] = \sum_{m=1}^{6} x_m$$

and then obtaining the final (for this stage) x_4, x_5 and x_6 values. Then with the aid of the following equations (where F_4, F_5 and F_6 values are assumed to be unity):

$$x_1 = \frac{K_1 c + [H^+]F_1 x_2}{K_1 + [H^+]F_1} \quad (33)$$

$$x_2 = \frac{K_2 x_1 + [H^+]F_2 x_3}{K_2 + [H^+]F_2} \quad (34)$$

$$x_3 = \frac{K_3 x_2 + [H^+]F_3 x_4}{K_3 + [H^+]F_3} \quad (35)$$

improved x_1, x_2 and x_3 values are obtained. At the following stage, with the aid of the obtained six x_m values, the ionic strength I is calculated with the aid of (29). The values of the activity coefficients of H^+ and all anions are approximated by the Debye-Huckel (7). With the aid of the activity coefficient values and (28) the F_1, F_2, F_3, F_4, F_5 and F_6 values are calculated. Using these values in (30) to (35), corrected values of all six x_m values are obtained. With the aid of the latters, the final $[H^+]$ value is determined.

2.4. The Use of the X_m Concept for a Determination of the Concentrations of the Ionized and Non-Ionized Forms and their Distribution in the Dilute Solutions of Weak Multibasic Acids

The determination of the x_m values gives us also the opportunity to calculate the important dissociation parameters: the concentrations of all ionized and non-ionized forms of weak multibasic organic acids in their dilute solutions. For this goal in general case of H_nA acid can be used (25) and (26). Taking into account also the following equation:

$$[H_nA] = c - x_1 \quad (36)$$

We can formulate the conditions of an equality of the concentrations of ionized and non-ionized forms:

$$[H^+] = [H_nA]{:}\, c = 2x_1 + \sum_{m=2}^{n} x_m \quad (37)$$

$$[H_{n-1}A^-] = [H_nA]{:}\, c = 2x_1 - x_2 \quad (38)$$

$$[H_{n-2}A^{2-}] = [H_nA]{:}\, c = x_1 + x_2 - x_3 \quad (39)$$

$$[H_{n-m}A^{m-}] = [H_nA]{:}\, c = x_1 + x_m - x_{m+1} \quad (40)$$

$$[A^{n-}] = [H_nA]{:}\, c = x_1 + x_n \quad (41)$$

Taking into account these conditions, we may formulate the areas of dominance of various ionized and non-ionized forms of acid:

$$[H^+] > [H_nA]{:}\, c < 2x_1 + \sum_{m=2}^{n} x_m \quad (42)$$

(and vice versa)

$$[H_{n-1}A^-] > [H_nA]{:}\, c < 2x_1 - x_2 \quad (43)$$

(and vice versa)

$$[H_{n-2}A^{2-}] > [H_nA]{:}\, c < x_1 + x_2 - x_3 \quad (44)$$

(and vice versa)

$$[H_{n-m}A^{m-}] > [H_nA]{:}\, c < x_1 + x_m - x_{m+1} \quad (45)$$

(and vice versa)

$$[A^{n-}] > [H_nA]{:}\, c < x_1 + x_n \quad (46)$$

(and vice versa)

3. CONCLUSIONS

Many weak multibasic organic acids have the comparatively close values of the dissociation constants of the different steps. This fact causes the participation of all steps in formation the hydrogen ions total concentration in the solutions of such acids. We suggest the new theoretical method for a determination of acidity of these solutions using the concepts of the contributions of the separate dissociation steps to the $[H^+]$ value, x_m. The equations for the accurate calculation of the $[H^+]$ values in cases of dibasic and tribasic acids and also in general case of weak multibasic acid, H_nA, are suggested. The comparatively simple method for a determination of acidity of the dilute solutions of high-basic (more than tribasic) acids is also described. With the aid of the formulated by us conditions of an equality of the concentrations of various ionized and non-ionized forms in the dilute solutions of weak multibasic organic acids the areas of dominance of these forms in connection with the corresponding x_m values are formulated.

REFERENCES

[1] Kvaratskhelia, E. and Kvaratskhelia, R. (2007) The degrees of dissociation of weak multibasic organic acids. *Joural of Solution Chemistry*, **36(6)**, 787-792.

[2] Kvaratskhelia, R. and Kvaratskhelia, E. (2000) Voltammetry of dicarboxylic acids on solid electrodes in aqueous and mixed media. *Russian Journal of Electrochemisry*, **36(3)**, 330-333.

J. Biomedical Science and Engineering, 2010, 3, 488-495 **JBiSE**
doi:10.4236/jbise.2010.35068 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

A consistency contribution based bayesian network model for medical diagnosis

Yan-Ping Yang

School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China.
Email: Yangyanping07@gmail.com

Received 8 February 2010; revised 1 March 2010; accepted 5 March 2010.

ABSTRACT

This paper presents an effective Bayesian network model for medical diagnosis. The proposed approach consists of two stages. In the first stage, a novel feature selection algorithm with consideration of feature interaction is used to get an undirected network to construct the skeleton of BN as small as possible. In the second stage for greedy search, several methods are integrated together to enhance searching performance by either pruning search space or overcoming the optima of search algorithm. In the experiments, six disease datasets from UCI machine learning database were chosen and six off-the-shelf classification algorithms were used for comparison. The result showed that the proposed approach has better classification accuracy and AUC. The proposed method was also applied in a real world case for hypertension prediction. And it presented good capability of finding high risk factors for hypertension, which is useful for the prevention and treatment of hypertension. Compared with other methods, the proposed method has the better performance.

Keywords: Bayesian Network; Medical Diagnosis; Feature Selection; Greedy Search

1. INTRODUCTION

Machine learning classification techniques have been applied to medical diagnosis tasks [1-4]. With many advantages over other methods, Bayesian network has become a promising tool in medical diagnosis [5-9]. It can reason under uncertainty by providing a concise representation of a joint probability distribution over a set of random variables [6,10]. Due to its graphical representation, a Bayesian network is easily understandable [11]. In addition, the diagnostic performance with Bayesian networks is often surprisingly insensitive to imprecision in numerical probabilities [12].

Bayesian network based diagnosis requires the learning of the Bayesian network structure, which is an optimization problem in the search space of the DAGs (directed acyclic graph). In general, researchers treat this problem in two distinctly different ways: the constraint-based method [13] and the search-and-scoring method [14,15]. The constraint-based methods are computationally effective. However, the search-and- scoring methods may produce more accurate results than the constraint-based methods [15]. In addition, constraint-based methods may fail to provide orientations for all edges in a network [13]. In this work we focus on structure learning and consider search-and-scoring methods.

When learning the structure of a Bayesian network model, however, search-and-scoring methods still face great challenges [16], because the true structure may not be identified by a finite number of observations, and it may be computationally intractable when the number of possible network structures grows super exponentially. To prune the search space, the existing methods depend on either setting a uniform bound on the number of parents [17], or conditional dependency test [18], which may degrade the performance of classification. As for the computation efficiency, hill-climbing search algorithm is the most commonly used because of its good trade-off between accuracy of the obtained model and ease of implementation. However, the well-known problem, myopia, may hinder the algorithm from finding the global optima.

In this paper, a new Consistency-Contribution based Bayesian Network algorithm, named CCBN, is proposed for medical diagnosis. To build the skeleton of Bayesian network, a novel feature selection method with the consideration of feature interaction is used to efficiently construct an undirected network. Further more, to enhance the performance of search algorithm, “AND” or “OR” strategy [19] is used to prune search space, while random start, beam-search and look-ahead help to overcome the optima of hill-climbing algorithm. To evaluate the effectiveness of the proposed algorithm in various medical problems, six disease datasets from the UCI machine learning repository were selected, and six off-the-shelf classifiers were used for comparison. Moreover,

a real world application on hypertension dataset was chosen to present its capability of finding high risk factors for hypertension. Its classification performance was further evaluated against other algorithms.

The rest of this paper is organized as follows. Section 2 introduces the necessary preliminaries about Bayesian network. In Section 3, the proposed classifier is described in detail. Then experiments on six disease datasets from the UCI machine learning repository are presented in Section 4. Later in Section 5, the proposed method is applied in a real world case for hypertension prediction to further evaluate its performance. Consequently in Section 6, the main contribution of this paper is summarized.

2. PRELIMINARIES

2.1. Bayesian Network

A Bayesian network $B = (G, \Theta)$ consists of two parts [7]. The first part is a network structure, which is a directed acyclic graph (DAG) G with each variable X_i represented as a node. The edges in the DAG represent statistical dependencies between the variables. The second part, Θ, is a set of parameters describing the probability density.

The problem of learning a BN can be stated as follows: given a finite data set $D = \{X_1, X_2, \cdots X_n\}$ of instances of variable X, find a minima set Φ to construct a graph structure G with proper parameters Θ that best matches D. In the commonly used search-and-scoring methods, a scoring function is usually used to evaluate how well the DAG G explains the data and then to search for the best DAG that optimizes the scoring function. In our proposed algorithm, Bayesian scoring criterion [14] is chosen. It computes the posteriori probability of a network given the data and prior knowledge [20].

2.2. Feature Selection

To get the skeleton of a BN, Φ, is a solution to feature selection [21]. In real-world classification problems, feature selection is indispensable to identify the most useful relevant features from data, and remove the irrelevant and redundant features [22]. Here, we define the feature relevant as that in [14]: in a feature set F, a feature F_i is relevant to a class C if and only if there exists a subset $S = F - \{F_i\}$, for which $P(C|F_i, S) \neq P(C|S)$; otherwise F_i is irelevant.

To find the relevant features, many effective algorithms have been developed [23,24], but they most often assume that features are independent. Due to attribute interaction [25], however, there are many features that can be considered as irrelevant when evaluated independently; but when they are combined with other features, their relevancy may be significant. Thus, unintentional removal of these features could lead to poor classification performance because of the loss of useful information.

Jakulin and Bratko [26,27] attempted to address this issue, but their methods can only deal with low order interaction. Later, Zhen Zhao [28] presented a more efficient algorithm by searching for interacting features. In this paper, we will use consistency measure to construct the skeleton of BN with consideration of feature interaction.

3. METHOD

3.1. Overview

In this section, we present our consistency measure based Bayesian Network algorithm, named CCBN. The algorithm is a two-step process: Firstly, an undirected network is constructed as the prototype of a BN's skeleton Φ based on symmetrical uncertainty and consistency hypothesis. This process finds a set that is as small as possible, but can deal with feature interactions in variable selection, and contain all parents, children and co-parents (attribute interaction) of each variable. In the second step, several methods are integrated together to enhance the performance of search algorithm by restricting the search space and overcoming the myopia of Hill-climbing search.

3.2. Constructing the Skeleton of BN

To construct the skeleton of a BN, an undirected network will be discovered in two phases. In phase-one, all variables (features) are ranked with regard to their SU (symmetrical uncertainty) and stored in the candidate set PC in descending order. In the second phase, the backward phase, features with less consistency-contribution will be removed from the candidate set PC. This elimination process is based on the consistency measure and takes into account feature interaction when removing both irrelevant and redundant features. In the end, the undirected network is obtained with candidate set PC. It will be used as the skeleton of BN.

As a fast correlation measure [24], SU is based on mutual information to measure the common information shared by two variables. It attempts to increase the chance for a strongly relevant feature to remain in the selected subset. The SU between the class label C and a feature F_i is formulated as:

$$SU(F_i, C) = 2\left[\frac{M(F_i, C)}{H(F_i) + H(C)}\right] \quad (1)$$

where $H(F_i)$ and $H(C)$ are the marginal entropies, and $M(F_i, C)$ is the joint entropy.

This formula compensates for information gain's bias toward features with more values and normalizes its values to the range [0,1] with the value 1 indicating that

knowledge of the value of the feature F_i can completely predict the value of class label C and the value 0 indicating that the feature F_i and class label C is independent.

After all features are ranked in descending order with regard to their *SU* values, a backward elimination method will be performed. However, in our method, *SU* will not be the criterion for elimination, because it can't guarantee that the interacting features are ranked highly. Instead, a novel consistency measure with the consideration of feature interaction is used to remove both irrelevant and redundant features.

In a dataset $D = \{d_1, d_2, \cdots d_m\}$, there is inconsistency if at least two instances have the same value for a feature space $F = \{F_1, F_2, \cdots F_n\}$, but they are categorized into different classes. Those inconsistent instances will be grouped together into a subset $\mathscr{D}_i$, named as inconsistent instance set. Every $\mathscr{D}_i$ has its inconsistency count $IC_F(\mathscr{D}_i)$. Let $\mathscr{D}_1, \mathscr{D}_2, \cdots \mathscr{D}_p$ represents all inconsistent instance sets of *D*, the inconsistency rate of *D* is:

$$ICR_F(D) = \frac{\sum_{1 \le i \le p} IC_F(\mathscr{D}_i)}{m} \quad (2)$$

Here, *m* represents the number of instances in *D*.

If F_i is irrelevant or redundant, removing it from *F* will have no effect on the number of inconsistent instantces. Consequently, there will be $ICR_F(D) = ICR_{F-\{Fi\}}(D)$. If F_i is a relevant feature, no matter whether it has feature interaction or not, its removal will bring more inconsistent instances, $ICR_F(D) < ICR_{F-\{Fi\}}(D)$. This unique character, named as consistency contribution, makes it a good metric for feature selection by taking the feature interaction into consideration.

In our method, a feature F_i's consistency contribution is formalized as:

$$CC(F_i, F) = ICR_{F-\{F_i\}}(D) - ICR_F(D) \quad (3)$$

From **Eq.3**, we can conclude that: higher consistency contribution value means higher relevancy; and the zero value indicates that it is either irrelevant or redundant feature. Using certain threshold value δ those weak relevant feature [16] can also be removed from PC if $CC(F_i) < \delta$. It will help to restrict the search space.

Based on the consistency-contribution, the elimination operation will start from the end of the ranked feature list. If the consistency-contribution of a feature is less than a predefined threshold δ ($0 < \delta < 1$), the feature is considered immaterial and can be removed; otherwise it is selected. The process is repeated until all features in the list are checked. A large δ is associated with a high probability of removing relevant features. δ is assigned 0.0001 in this paper if not otherwise mentioned.

3.3. DAG Search

After an undirected network is constructed as the prototype of Φ, a local hill climbing algorithm will be used to search for the best DAG that optimizes the scoring function in the restricted space of DAGs.

To make the greedy hill climbing algorithm more efficient, our method takes into account several aspects, like preprocessing, starting solution, how to acquire neighbors and score metrics for candidate neighbors.

Before performing the hill-climbing algorithm, the search space of DAGs is initially restricted by using "OR" or "AND" strategy [19], because of the likely asymmetric dependencies in the finite sample scenario (*i.e.*, the edge a-b may be found on a but not on b). In "OR" strategy, an arc is excluded only if it was not found in either direction, while "AND" strategy requires edges to be found in both directions.

In addition, our DAG search will start with a random start DAG network rather than an empty one (*i.e.*, the one with no edges). Based on the restricted search space for DAGs, a random number of edges will be chosen randomly. The edge will be added to the original DAG only if it is discovered in the process of constructing the undirected network.

To obtain the neighbor at each search step, the legal operators used are arc addition, arc deletion, and arc reversal. Of course, except in arc deletion, we have to make sure that there is no directed cycle introduced in the graph. For fast evaluation of candidate sub-graphs (neighbors), the Bayesian scoring criterion is used in our work.

To overcome the myopia of hill climbing algorithm, the 2-step look ahead in good directions (LAGD) hill climbing is used. Further more, 5 best evaluation operations will be considered at each search step.

Following is the procedure of our DAG searching:

1) Restrict search space of DAG with "OR" or "AND" strategy;

2) Form an original DAG *Mc* with random starting and get the Bayesian scoring metric for *Mc*;

3) Obtain all neighbors of current candidate DAG *Mc* by getting all possible directional orientations for the edges in *Mc* and compute the Bayesian scoring metric for each neighbor;

4) By ranking Bayesian scoring metric for all neighbors in descending order, store the 5 best valued network structures in a specific set B_1;

5) For each candidate stored in set B_1, get its neighbors in the same way as 3) and get its Bayesian scoring metric;

6) Choose the neighbor with highest Bayesian scoring metric in step 5), and mark it as new candidate DAG *Mnew*;

7) Greedy operation: if new candidate *Mnew* has Bayesian scoring metric higher than that of current Mc, replace Mc with new candidate *Mnew*, and then go back to step 3) for further improvement; otherwise, it means that there is no better solution available and the search ends up with the final graph structure *Mc*;

8) Deliver the out-coming graph structure *Mc*.

4. EXPERIMENTS

To evaluate our method's effectiveness for medical diagnosis, six disease datasets from the UCI machine learning repository [29] is used in experiments. In this section, we will discuss the experiments in detail.

4.1. Datasets Description

Among the selected datasets, datasets *sick* and *hypothyroid* are chosen for the diagnosis of thyroid diseases. Dataset *spectf_test* is selected for diagnosis of heart disease. In addition, the datasets for *mammographic*, *diabetes* and horse's *colic* are also used in experiments.

Table 1 tabulates the detailed information of the experimental datasets. They are further analyzed, respectively, to verify the effectiveness of the proposed CCBN algorithm on medical diagnosis tasks.

4.2. Implementation Results

For comparison, three representative classification algorithms are chosen. They are NaiveBayes [1], C4.5 [4], and IB1 [30]. All are available in the WEKA environment [31]. Our CC-BN method is also implemented in the WEKA's framework. Moreover, Bayesian Networks with different local search algorithms [32], like hill-climbing, LAGD hill-climbing and simulated annealing are also chosen for comparison. They are relatively represented by HC-BN, LAGD-BN and SA-BN. To guarantee the impartiality of experiments, for HC-BN, LAGD-BN and SA-BN, the maximum allowed size for the candidate parents' set is the same as the size of features in each dataset. For the evaluation of performance, error rate and ROC [33] were used in experiments with10-fold cross validation.

Table 2 reflects the error rate generated by each method based on 10-fold CV. On average, CCBN (AND) has the best performance with error rate 12.2% and CCBN (OR) is the second best with 12.8%. We can see that both CCBN methods have distinctly better performance than others, and the difference between two CCBN methods is slight. Among the six datasets, at least one of CCBN methods has the lowest error rate in four of the datasets. For the other two datasets, CCBN methods still have the second lowest error rate. It proves that CCBN has a better consistent performance than the other methods used in the experiments.

Impressively, on datasets *sick* and *hypothyroid*, which have a large set of features and instances, all Bayesian Network methods have very low error rates, which are less than 60% of error rate of IB1 and not more than 25% of that of Naïve Bayes or C4.5. This shows the capability of Bayesian Network in dealing with high dimension data. But on other datasets, unlike CCBN, HC-BN, LAGD-BN and SA-BN have no great advantages over other conventional methods. This proves that our proposed, novel, method brings great enhancement to Bayesian Networks for coping with various problems.

Among BN methods, SA-BN has apparent advantage over HC-BN and LAGD-BN. However, simulated annealing search costs much more time than hill climbing methods. For example, on *colic* dataset, it takes less than 10 seconds for HC-BN and LAGD-BN to finish 10-fold CV testing, while it costs 2250 seconds for SA-BN. It proves that hill climbing algorithm has a good trade off between accuracy and efficiency. Further when compared with CCBN, SA-BN only wins on *colic*. It may be because CCBN gets benefit from our proposed method that brings improvement to HC and LAGD methods.

To study the impact of threshold δ on CCBN's performance, more experiments were performed with various thresholds δ. For each dataset, the error rate of LAGD-BN in **Table 2** is used as reference. In **Figure 1**, the ratio of error rate between CCBN(AND) and LAGD-

Table 1. Summary of experimental disease datasets from UCI.

Dataset	#features	#instances	#classes
colic	22	368	2
Sick	30	3772	2
Hypothyroid	30	3772	4
diabetes	8	768	2
spectf_test	44	267	2
mammographic	6	961	2

Table 2. Error rates of the classifiers on experimental disease datasets.

Dataset	HC-BN	LAGD-BN	SA-BN	Naïve Bayes	IB1	C4.5	CCBN(AND)	CCBN(OR)
colic	.179	.185	.155	.220	.188	.147	.168	.179
Sick	.022	.021	.023	.165	.038	.012	.021	.022
Hypothyroid	.010	.010	.008	.047	.085	.042	.008	.008
diabetes	.247	.247	.257	.237	.298	.262	.236	.236
spectf_test	.234	.223	.227	.335	.283	.253	.208	.216
mammographic	.169	.169	.170	.175	.253	.176	.168	.168
Avg.	.144	.143	.140	.197	.191	.149	.137	.139

BN is shown to reflect the effect of threshold δ.

Figure 1 shows that on *hypothyroid*, *colic*, *diabetes* and *mammographic*, CCBN is better than that of LAGD-BN, even for $\delta = 0$. Such an improvement should owe to random start which has depressed local suboptima. Also, with $\delta = 0.00001$ or $\delta = 0.0001$, CCBN shows much more improvement over LAGD-BN for all datasets except *sick*. This supports our claim that our method is much more efficient for feature selection due to the consideration of feature interaction. However, it should be noted that CCBN's advantage over LAGD-BN is not consistent. On *hypothyroid*, CCBN's performance is worse than LAGD-BN's when $\delta = 0.01$. On *sick*, CCBN has no win. We will discuss about this more in next section.

Also, when studying the impact of δ on CCBN's performance, we found that for δ with range from 0 to 0.001, the change of error rate for each dataset is less than 10%. This means that our method is insensitive to the threshold value δ. This improvement may benefit from random start.

Although classification accuracy is our focus, ROC is also introduced in our experiment's evaluation. According to **Table 3**, CCBN method has highest average area under ROC (AUR). Although its AUR is not highest on colic, *spectf-test* and diabetes, the difference is slight. Also, it is apparent that BN methods have higher AUR than other traditional methods.

4.3. DISCUSSION

By comparing CCBN with other classification algorithms, experiments proved that CCBN can achieve better performance, but there are still some issues that should be addressed for improvement.

First, the error rate of CCBN (AND) on four of six datasets, is better than that of the other methods used in our experiments. On *colic* and *mammogrpahic*, however, this does not hold, although CCBN (AND) is still the second best. This suggests that there may be some properties

Second, although on most dataset CCBN proves to be insensitive to threshold value, threshold value still shows big impact on *spectf-test*. Also, the best error rate for different datasets happens on different thresholds. It is obvious that the ideal choice of a threshold is the turning point that provides the lowest error rate. Exploring the meaning of the threshold, for example, from a statistical aspect, is another interesting topic for future work. It is likely that this turning point may not be independent of the dataset (problem) to be analyzed. If it is true, we hope to propose some feasible approaches that could tune the threshold automatically for a specific problem.

5. APPLICATION ON HYPERTENSION PROBLEM

In this section, a real-case from medical diagnosis is presented. We will show that our proposed method is capable of finding the high risk factors for hypertension diagnosis and prediction. Also, classification result on hypertension is evaluated against other algorithms.

5.1. Problem Description

As one of the most common disease in the world, hypertension is considered to be present when a person's systolic/diastolic blood pressure is consistently 140/90 mmHg or greater [34]. It is affecting estimated 600 million people worldwide [35], and that number is expected to increase to 1.56 billion by 2025 [36]. Therefore, it is meaningful to develop a computer-aid system for its diagnosis and prediction.

In our experiments, the data from the hypertension study were specifically collected by Hubei Provincial Center for Disease Control and Prevention in China. The investigation was conducted in compliance with the IRB

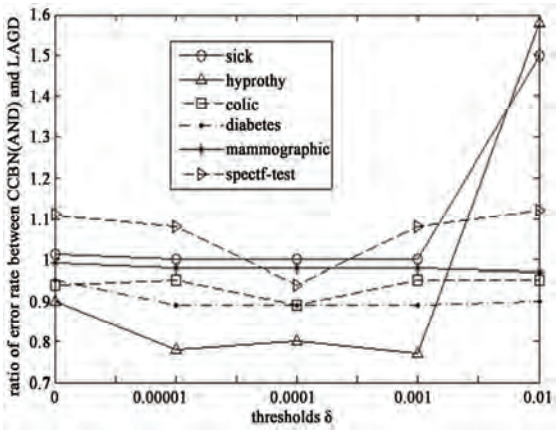


Figure 1. The ratio of error rate between CCBN (AND) and LAGD for different thresholds on six UCI datasets.

Table 3. AUR of the classifiers on experimental disease datasets.

Dataset	HC-BN	LAGD-BN	SA-BN	Naïve Bayes	IB1	C4.5	CCBN(AND)	CCBN(OR)
colic	.874	.873	.890	.842	.795	.813	.88	.876
Sick	.958	.959	.960	.925	.804	.951	.965	.964
Hypothyroid	.482	.483	.519	.282	.500	.197	.703	.703
diabetes	.807	.807	.800	.819	.662	.751	.799	.799
spectf_test	.765	.772	.799	.849	.620	.592	.795	.746
mammographic	.897	.897	.897	.894	.744	.855	.898	.898
Average	.797	.798	.810	.768	.688	.693	.840	.831

approved study protocol and the law in China. It included a representative sample of 3053 adults: 1438 males, 1610 females and five without sex information. The participants ranged in age from 35 to 92 years old and average age was 51 years old. The prevalence of hypertension was 26.9%. The 155 features in the database include sample population characteristics, heal- th, health care, habits of life, and etc. Obviously irrelevant and sparse (more than 50% missing data) features were discarded from further investigation. The features used in the experiments were recommended by the clinician for diagnosis. The resulting dataset contained 32 potential risk factors such as age, waistline, body-mass index (BMI) and etc (detailed information about some important attributes are presented in **Table 4**).

5.2. Classification Results

In order to assess the effectiveness of the proposed approach, its performance was evaluated against that of some off-the-shelf algorithms. The performance of the different algorithms is listed in **Table 5**, which contains Error rate, Sensitivity (SN), Specificity (SP), Positive predictive value (PPV), and Negative predictive value (NPV). The result shows that CCBN with 'AND' strategy has the same performance as that with "OR" strategy for hypertension classification. Both CCBNs have the lowest error rate on the hypertension classification. Besides error rate, CCBN methods also have the best specificity, PPV and AUR. On sensitivity, although Naive Bayes has a slightly better performance than CCBNs, the difference is not significant. Therefore, we can conclude that CCBN has better performance on hypertension classification than other methods.

5.3. Risk Factor Analysis

For the prediction of hypertension, CCBN tried to find those likely high risk factors. These high risk factors enabled the classifiers to achieve their best performance and also would give us important guide to hypertension prevention and treatment.

In experiment, CCBN marked age, BMI and blood pressure as the high risk factors in hypertension development. This result is consistent with many clinical studies.

As we know, lifestyle plays an important role in the development and treatment of hypertension. For the dietary intake, fish, meat, soybean products, eggs and diary are selected by CCBN, which is coherent with current knowledge about hypertension. According to exercise, the original data have three attributes: frequency of exercise in a week, taking exercise every day or not, number of days in a week that have more then 30 minutes spent on exercise. In CCBN, the first two attributes are ignored, because the third attribute has more or at least not less information about the impact of exercise on hypertension than other two attributes.

Among the attributes about social background, nationality and occupation are removed from BN. Remarkably education background and marital status are also selected as key factors in prediction of hypertension. CCBN shows that higher education will lead to less risk of hypertension, and the single is more susceptible to hypertension than the married. It is reasonable because people with higher education are more aware of the jeopardy of hypertension and have more knowledge about the prevention of hypertension, which in turn leads to good lifestyle. Similarly for marital status, the married is more likely to lead regular lifestyle and have more healthy diet than the single.

6. CONCLUSIONS

Bayesian network is a powerful learning technique to tackle the uncertainty in medical problems, but learning

Table 4. Summaries of mean and standard deviations of some important features in hypertension dataset.

Features	Mean	S.D	Missing (%)
age	51.43	12.32	0.00
sex	1.53	0.50	0.00
cigarette smoker	1.63	0.49	0.04
Grossincome families	3.65	1.19	0.01
excessive alcohol intake	1.69	0.46	0.07
body mass index (BMI)	22.14	3.03	0.03
family history of hypertension	1.12	0.33	0.10
physical inactivity	1.98	1.25	0.08
marital status	1.24	0.71	0.01
education	2.35	0.87	0.00

Table 5. Performance of classifiers for the hypertension dataset based on 10-fold CV.

Algorithms	Error rate	Sensitivity	Specificity	PPV	NPV
CCBN(AND)	.163	.553	.942	.779	.851
CCBN(OR)	.163	.553	.942	.779	.851
NaïveBayes	.198	.555	.892	.656	.845
C4.5	.186	.520	.922	.711	.839
IB1	.329	.309	.805	.369	.759
HC-BN	.164	.553	.940	.772	.851
SA-BN	.164	.550	.942	.777	.850
LAGD-BN	.164	.555	.941	.776	.852

its structure is a great challenge, which may hinder its wide application in medicine. In this paper, the CCBN algorithm is proposed for medical diagnosis. Our method firstly constructs the skeleton of BN with consideration of feature interaction, and then performs the efficient search for an optimal BN network. Case studies on disease datasets and application on hypertension predictiton show that the proposed method is a promising tool for computer-aided medical decision system, because it has better classification accuracy and consistency than other traditional methods.

REFERENCES

[1] Mitchell, T.M. (1997) Machine learning. McGraw-Hill, New York.

[2] Jordan, M.I. (1995) Why the logistic function? A tutorial discussion on probabilities and neural networks technology. *MIT Computational Cognitive Science Report* 9503, Massachusetts.

[3] Aha, D., Kibler, D. and Albert, M. (1991) Instance-based learning algorithms. *Machine Learning*, **6(1)**, 37-66.

[4] Quinlan, J.R. (1993) Programs for machine learning. *San Mateo*, Morgan Kaufmann, California.

[5] Lisboa, P.J.G., Ifeachor, E.C. and Szczepaniak, P.S. (2000) Artificial neural networks in biomedicine. Springer-Verlag, London.

[6] Cooper, G.F. and Herskovits, E. (1992) A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, **9(4)**, 309-347.

[7] Lucas, P. (2001) Bayesian networks in medicine: A model-based approach to medical decision making. *Proceedings of the EUNITE Workshop on Intelligent Systems in Patient Care*, Vienna, 73-97.

[8] Antal, P., Verrelst, H., Timmerman, D., Van Huffel, S., De Moor, B., Vergote, I. and Moreau, Y. (2000) Bayesian networks in ovarian cancer diagnosis: Potentials and limitations. 13*th IEEE Symposium on Computer-Based Medical Systems*, Texas Medical Center, Houston, 103-108.

[9] van Gerven, M., Jurgelenaite, R., Taal, R., Heskes, T. and Lucas, P. (2007) Predicting carcinoid heart disease with the noisy-threshold classifier. *Artificial Intelligence in Medicine*, **40(1)**, 45-55.

[10] Pearl, J. (1988) Probabilistic reasoning in intelligent systems: Networks of plausible inference. 2nd Edition, Morgan Kaufmann, San Francisco.

[11] Spirtes, P., Glymour, C. and Scheines, R. (2000) Causation, prediction, and search. 2nd Edition, the MIT Press, Cambridge.

[12] Pradhan, M., Henrion, M., Provan, G., Favero, B.D. and Huang, K. (1996) The sensitivity of belief networks to imprecise probabilities: An experimental investigation. *The Artificial Intelligence Journal*, **85(1-2)**, 363-397.

[13] Cheng, J., *et al*. (2002) Learning Bayesian networks from data: An information theory based approach. *The Artificial Intelligence Journal*, **137(1-2)**, 43-90.

[14] Cooper, G. and Herskovits, E. (1992) A Bayesian method for the induction of probabilistic networks form data. *Machine Learning*, **9(4)**, 309-347.

[15] Heckerman, D. (1998) A tutorial on learning with Bayesian networks, learning in graphical models. Kluwer Academic Publishers, Dordrecht, 301-354.

[16] Chickering, D. (1996) Learning Bayesian networks is NP-complete. In: Fisher, D. and Lenz, H., Eds., *Learning from Data*: *Artificial Intelligence and Statistics*, **4(1)** 121-130.

[17] Friedman, N., Nachman, I. and Peer, D. (1999) Learning Bayesian network structure from massive datasets: The 'sparse candidate' algorithm. *Uncertainty in Artificial Intelligence*, **15(3)**, 206-215.

[18] Tsamardinos, I., Brown, L. and Aliferis, C. (2006) The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, **65(1)**, 31-78.

[19] Meinshausen, N. and Buhlmann, P. (2006) High dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, **34(3)**, 1436-1462.

[20] Heckerman, D., Geiger, D. and Chickering, D. (1995) "Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, **20(3)**, 197-243.

[21] Tsamardinos, I. and Aliferis, C. (2003) Towards principled feature selection: Relevancy, filters and wrappers. *The* 9*th International Workshop on Artificial Intelligence and Statistics*, Florida, 334-342.

[22] Guyon, I. and Elisseeff, A. (2003) An introduction to variable and feature selection. *Journal of Machine Learning Research*, **3(7-8)**, 1157-1182.

[23] Dash, M. and Liu, H. (2003) Consistency-based search in feature selection. *Artificial Intelligence*, **151(1-2)**, 155-176.

[24] Hall, M. (1999) Correlation based feature selection for machine learning. Ph.D. Dissertation, University of Waikato, New Zealand.

[25] Kononenko, I. (1994) Estimating attributes: Analysis and extension of RELIEF. *European Conference on Machine Learning*, Catania, 171-182.

[26] Jakulin, A. and Bratko, I. (2003) Analyzing attribute dependencies. PKDD, Ljubljana.

[27] Jakulin, A. and Bratko, I. (2004) Testing the significance of attributes interactions. *International Conference on Machine Learning*, **20**, 69-73.

[28] Zhao, Z. and Liu, H. (2007) Searching for interacting features. *Proceedings of International Joint Conference on Artificial Intelligence*, Nagoya.

[29] Blake, C., Keogh, E. and Merz, C.J. (1998) UCI repository of machine learning databases. University of California, Irvine. http://www.ics.uci.edu/~mlearn/MLRepository.html

[30] Aha, D.W., Kibler, D. and Albert, M.K. (1991) Instance-based learning algorithms. *Machine Learning*, **6(1)**, 37-66.

[31] Witten, I.H. and Frank, E. (2005) Data mining-pracitcal machine learning tools and techniques with JAVA implementations. 2nd Edition, Morgan Kaufmann Publishers, California.

[32] Bouckaert, R.R. (2008) Bayesian network classifiers in weka for version 3-5-7. *Artificial Intelligence Tools*, **11(3)**, 369-387

[33] Cortes, C. and Mohri, M. (2003) AUC optimization vs. error rate Minimization. *Advances in Neural Information Processing Systems*, **11(10)**, 356-360.

[34] Chobanian, A.V., Bakris, G.L., Black, H.R., Cushman, W.C., Green L.A. and Izzo, J.L. (2003) The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *The JNC 7 Report, Journal of the American Medical Association*, **289(19)**, 2560-2572.

[35] (2001-2002) Cardiovascular diseases – Prevention and control. *WHO Chemical Vapor Deposition Strategy Conference*, Nottingham.

[36] Kearney, P.M., Whelton, M., Reynolds, K., *et al*. (2005) Global burden of hypertension: Analysis of worldwide data. *Lancet*, **365(9455)**, 217-223.

J. Biomedical Science and Engineering, 2010, 3, 496-500 **JBiSE**
doi:10.4236/jbise.2010.35069 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

A stage-scanning laser confocal microscope and protocol for DNA methylation sequencing

Vaithilingam Vaishnavi, Litty Varghese, Baquir Mohammed Jaffar Ali

Life Sciences Division, AU-KBC Research Centre, Anna University, Chennai, India.
Email: jaffar@au-kbc.org

Received 25 September 2009; revised 20 October 2009; accepted 25 October 2009.

ABSTRACT

Recent understanding of the role of epigenetic regulation in health and disease has necessitated the development of newer and efficient methods to map the methylation pattern of target gene. In this article we report construction of a stage-scanning laser confocal microscope (SLCM) and associated protocol that determines the methylation status of target gene. We have adapted restricted Sanger's sequencing where fluorescine labeled primers and dideoxy guanine fraction alone are used for target amplification and termination at cytosine positions. Amplified ssDNA bands are separated in 6% denaturing PAGE and scanned using SLCM to sequence the positions of methylated cytosines. We demonstrate that our methodology can detect < 100 femtomoles of DNA, and resolve the position of cytosine within ± 2 nucleotide. In a calibration run using a designer DNA of 99 bases, our methodology had resolved all the 11 cytosine positions of the DNA. We have further demonstrated the utility of apparatus by mapping methylation status in the Exon-1 region of a gene, E-Cadherin, in the plasma DNA sample of a healthy subject. We believe our approach constitute a low cost alternative to conventional DNA sequencers and can help develop methylation based DNA biomarkers for the diagnosis of disease and in therapeutics.

Keywords: Scanning Confocal Microscope; Fluorescence Detection; Methylation; Bisulfite Sequencing; DNA Sequencer

1. INTRODUCTION

Cytosine methylation is required for mammalian development [1]. Alterations of CpG island DNA hypermethylation and chromatin modification have been widely documented in human cancers [2,3] and various other diseases [4,5]. Disease-specific DNA methylation alterations present in tissues and plasma samples can serve as diagnostic markers for risk assessment, progression, early detection, treatment prediction and monitoring [6]. Therefore determination of methylation status of target gene is becoming increasingly important in diagnosis and treatment. Most laboratories use high performance liquid chromatography (HPLC) for the detection of 5-methylcytosine residues [7]. However, HPLC requires large amount of DNA and significant experimental experience to obtain reproducible results. Both points are impedimental to the large-scale analysis of clinical samples [8]. Bisulfite genomic sequencing provides an efficient method for detecting individual 5-methyl cytosine (5-MeC) residues in genomic DNA [9]. Analysis of cytosine methylation from the bisulfite- treated and PCR-amplified DNA can be determined by either cloning or sequencing individual molecules or by direct PCR sequencing [10]. In cloning approach, it often requires sequencing of 50-100 individual clones to provide an accurate estimate of methylation in the population of molecules. Whereas in sequencing, standard sequencing is performed using DNA sequencers.

In most situations, methylation status of promoter is taken as indicator to decide gene is expressed or repressed. It is determined by methylation-specific PCR (MSP) which detects the presence of methylated DNA in CpG islands of the promoter [11]. In MSP, primers are designed to bind specifically to bisulfite converted methylated or unmethylated DNA, leading to specific amplification. Advantage of this technique is its high relative sensitivity, ease of design, and low complexity of the reaction, but it is limited to determining the methylation status only in the primer binding region. Therefore, a new strategy is required to know complete methylation status of target gene.

Complete information of methylation status can usually be derived from sequencing the DNA. The current state-of-the-art technology for high-throughput DNA sequencing utilizes capillary array DNA sequencers with laser-induced fluorescence detection [12-15]. In the case of methylation mapping of the genome, it is redundant to

sequence the bisulfite converted DNA when sequence information of the genome is already known. Any approach which can by-pass the complete sequencing requirement can be more efficient, faster and resource conserving.

In this communication, we report the customized instrument and protocol for fluorescence detection of DNA bands, reporting methylation status in a target gene. An automated stage-scanning laser confocal microscope (SLCM) has been constructed and calibrated using fluorescently labeled oligo DNA of known lengths. Avoiding detail sequencing of the target gene and thus eliminating the use of DNA sequencer has remained a major objective in our endeavor. In addition, our novel approach in the protocol enabled mapping the methylation status of the target gene in detail that can match sequencing of the gene.

2. METHODS

2.1. Sample Preparations for Bisulfite Sequencing

2.1.1. DNA Extraction and Sodium Bisulphite Conversion

Genomic DNA from human blood plasma was isolated from 1 mL of serum collected from a normal subject using standard phenol/chloroform method and resuspended in 100 µL of water [16]. About 1 µg of DNA was denatured by NaOH followed by treatment with sodium bisulfite to convert all unmethylated cytosine to uracil [17]. Bisulfite treated DNA samples were purified using the Wizard DNA clean-up system (Promega Corp. USA), followed by NaOH treatment and ethanol precipitation. Finally it is resuspended in 20 µl of water.

2.1.2. Amplification of Target DNA

Following two DNA templates were designed and synthesized for calibration of the instrument. *Template A*: 5'**C**GA T**C**G TAT T**C**G G**C**G TTT GTT TT**C** GTT**C**GG **C**GT TTT **C**GG TTA GTT ATG GGT TTT TGG AGT **C**GT AGT TTT T**C**G G**C**G TTG TTG TTG TTG TTG TAG GTA TTT3' containing 11 CpG spots, and *Template B*: 5'**C**GA TAG TAT TAG G**C**G TTT GTT TTA GTT AGG AGT **C**TT GGT TAG TTA TGG GTT TTT GGA GT**C** GTA GTT TTT **C**GG GTT GTT GTT GTT GTT GTA GGT ATT T3' containing 5 CpG spots, mimicking methylated cytosine positions. Using MethPrimer [18] primer without CpG dinucleotides and devoid of regions with poly T ≥ 9 is designed to amplify ssDNA Templates A and B, and target in Exon-1 region of E-cadherin. The resulting primer synthesized with Fluorescein label (6-FAM) at 5' is given by: *Flu*-5' AAA TAC CTA CAA CAA CAA CAA CAA C3' (Ocimum Biosolutions, Hyderabad). Restricted Sanger's sequencing reaction specifically developed for this [19] was performed using SILVER SEQUENCE™ DNA sequencing system (Promega Corp. USA) on these templates. In that dideoxy guanine fraction alone is used so as to terminate the amplicons at all cytosine positions which corresponds to methylated cytosines. The optimized thermal cycling condition giving best results were as follows: 95℃ for 2 min, 35 cycles of 95℃ for 30 sec, 42℃ for 30 sec and 72℃ for 60 sec.

2.1.3. Electrophoresis Separation

Control ssDNA templates of various sizes containing Fluorescine tag at 5' end, and the bisulfite PCR products were separated on 6% denaturing PAGE in 89 mM Tris, 89 mM boric acid, 2 mM ethylenediamine tetra acetic acid (EDTA) at field strength of 10 V/cm for various time interval. Before loading the sample into the gel, an aliquot of the PCR mixture prepared above was added to 2 µL of glycerol and heated at 95℃ for 3 min to denature the amplicons. Concentration of control ssDNA templates were gradually decreased in different lanes to determine lowest possible concentration that can be detected and resolved by the apparatus.

2.2. Stage-Scanning Laser Confocal Microscope (SLCM)

The optical layout of the SLCM setup is schematically represented in **Figure 1**. The apparatus is built around the Nikon TE2000 microscope frame. The sample stage consists of a motorized XY translator having 100 mm displacement in each axis (Model: DPM Translator with Micro Position Controller, Holmarc, India). A custom built stage plate with 90 mm by 90 mm square slot constitute platform for gel scan. The 473 nm DPSS laser, CW/100 mW (Model: Ceil, Laser Quantum, UK) is used for fluorescence excitation. The laser

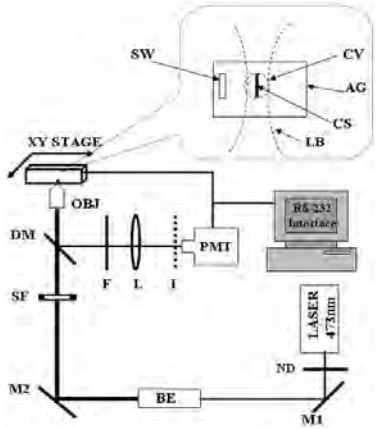


Figure 1. Optical lay-out of stage-scanning laser confocal microscope. OBJ: Objective lens; SF: Spatial filter; DM: Dichroic mirror; M1 and M2: Visible reflecting mirrors; BE: 5x beam expander; ND: Neutral density filter; F: Emission filter; L: convex lens; I: Iris; PMT: Photomultiplier Tube; LB: laser beam; CV: confocal volume; CS; confocal volume section; SW: Sample well; AG: acrylamide gel.

beam is expanded to fill the back-focal plane of the 20 X objective (Nikon). A rectangular slit is introduced as a spatial filter to section the confocal volume to a thin rectangular confocal volume-section. It is aligned parallel to the DNA bands so as to efficiently excite fluorescine tagged to the DNA and minimize the background noise in emission signal. The emission is collected through an iris acting as a confocal spatial filter at left port of the microscope. The photomultiplier tube (Hamamatsu H7468) serves as the detector.

2.3. Data Acquisition and Control

XY-Stage and PMT are controlled through RS232 interface. Linear scan is made at a speed of 0.9 mm/s and fluorescence signal is acquired by PMT set at 550 V, 100 ms integration with 0.01 ms dead time interval. Typical scan of the full length gel is about 80 mm for ~ 89 sec, giving ~11data points per mm scan. The laser power at the back-focal plane of the objective is ~3.5 mW.

2.4. Resolution and Sensitivity

Ability to resolve ssDNA bands is determined by resolving 20 to 43 bases oligo DNA for lower lengths and 73 and 76 bases oligo DNA for longer length. Sensitivity of detection is determined by linear dilution of control ssDNA templates. The optimized scanning and detection protocol given in the above section was followed throughout the measurement of sensitivity & resolution as well as for mapping the methylation pattern in ssDNA templates and target gene in human genomic DNA.

3. RESULT AND DISCUSSION

3.1. Characterization of SLCM

3.1.1. Sensitivity

Figure 2(a) demonstrates the ability of the apparatus to resolve less than 100 femtomole of fluorescine tagged DNA. The quantity of ssDNA detected is less than typical amount of DNA available in sequencing PCR products in laboratory conditions. Further, perusal of instrument parameters for SLCM suggests that there is more room to improve the fluorescence detection which may bring down the sensitivity level even further.

3.1.2. Resolution

Following the same parameters for detection, **Figure 2(b)** gives the resolution of four short length oligos differing in two bases. In **Figure 2(c)**, resolution of three base differences in longer length oligo is demonstrated. To further demonstrate the ability to resolve multiple bands, ssDNA ladder of eight bands is scanned and given in **Figure 2(d)**. Scrutiny of DNA band separation in **Figure 2** suggests that the ability to resolve further is limited by short mobility of bands. In the current configuration, we have practical limitation of 80 mm scan of the sample which limits the maximum distance over which ssDNA bands can run. In other words, longer scan length will be in a position to resolve the bands to single base resolution. Since methyl sequences can occur at minimum of two bases difference, we have not further explored the ability of the instrument to resolve better.

3.2. Methylation Mapping in Genomic DNA

Efficient, unbiased PCR amplification from methylated and unmethylated genomic DNA is essential for any PCR amplification-dependent method designed to detect methylation following bisulphite conversion. Amplicons generated from methylated genomic DNA remains CpG rich relative to the amplicons from unmethylated genomic DNA, and are often amplified less efficiently. Amplicons from bisulphite-converted genomic DNA typically have homopolymer stretches of 9 or more Ts (As), resulting in poor (or no) amplification. A broadened signal have been observed during electrophoresis for these amplicons due to enzyme "slippage" causing n + 1 and n – 1 sequences [20]. By selecting amplicons containing no more than 5 consecutive Ts (or As), our PCR was almost always successful. Achieving these two requirements–avoiding homopolymer stretches and designing a primer in a non-CpG region often restricts primer selection to a single possibility for any amplicons within CpG Island.

Following specifically developed restricted Sanger's sequencing protocol [19], two ssDNA templates having known numbers of cytosines were subjected for methyl sequencing. **Figure 3(a)** gives resolution of all eleven positions present in the ssDNA Template-A, and **Figure 3(b)** gives all the five positions present the in ssDNA Template-B. In **Figure 4(a)**, methyl sequencing results on the bisulfite treated human genomic DNA at DNA target

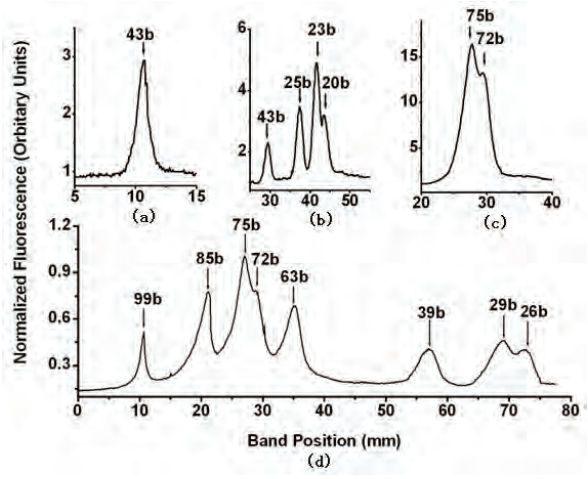


Figure 2. Demonstrates sensitivity and resolution of SCLM. (a) Detection of 100 femtomoles of ssDNA of 43bases; (b) Resolution of oligos of length 20, 23, 25 and 43 bases; (c) Resolution of oligos of length 72 and 76 bases; (d) Resolution of set of 8 ssDNA Templates length varying from 26 bases to 99 bases.

on Exon-1 using methylation insensitive primer is given. As a control for resolution of DNA bands, three ssDNA templates of known bases run concurrently on parallel lane is given in **Figure 4(b)**. In **Figure 4(c)**, band position of known control DNA ladder is plotted against the base number. The data is fit to a polynomial to interpolate unknown base information corresponding to position information. For pedagogical reason all possible CpG positions within 100 bases are plotted (red diamonds). From the precise band position of methyl sequenced DNA in **Figure 4(b)**, and the known corresponding base information, unknown nucleotide sequence is obtained through interpolation of fit in **Figure 4(c)**. It is observed that the target gene had only 30 and 64 base positions from 3' end of the primer found to be methylated. It is noted that only two cytosines are methylated out of 15 available CpG positions. Ability to retrieve all amplicons in template DNA (**Figures 3(a)** and **3(b)**) suggests that the target sequence in human genomic DNA is partially methylated at the CpG islands in the promoter region. We emphasis that in prototype resolution of DNA bands in this manner can be *software-controlled* by inputting the calibration parameters into the system, giving raise to instant reading of target DNA band sequence. Information on methylation sequence position together with quantitative gene expression data can shed more light on the mechanism of epigenetic regulation. More significantly, ability of SLCM to obtain quantitative information on DNA methylation as well as to track its position precisely anywhere in the target sequence can open up many possibilities including development of DNA methylation as biomarkers.

4. CONCLUSIONS

It is demonstrated that a customized SLCM apparatus can efficiently resolve ssDNA separation in acrylamide gel. We have applied this to map the methylation pattern of target gene. Our SLCM built on low-cost but high sensitive photon detector enabled discrimination of DNA bands of less than 100 femtomoles. Perusal of the parameters of excitation and detection reveal tremendous scope to engineer the sensitivity and resolution in the same configuration. We note that the geometry of fluorescence excitation achieved by using a spatial filter is confocal *volume-section* instead of *volume*. It has considerably improved the signal-to-noise. We have also outlined how the current specification can be improved with extension of scan length and modification. Further, XY scan of gel can enable mapping multiple lanes, thereby sequencing long length of target in a gene or multiple targets in the same genome. We emphasize that normally methylation status away from promoter region or more specifically, away from primer region could not be obtained in any methods except complete sequencing

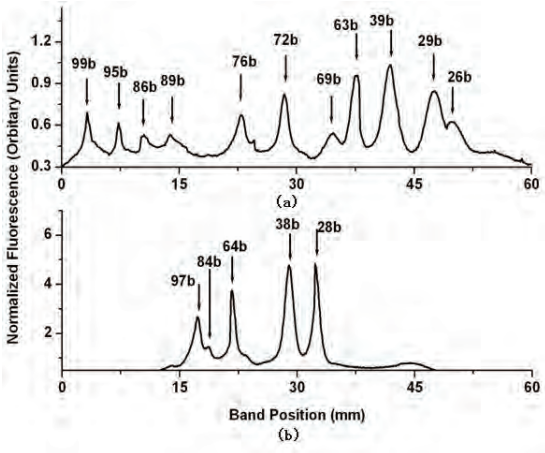


Figure 3. Resolution of methyl sequencing bands in ssDNA Templates. (a) Separation of methylation bands in ssDNA Template-A and; 9b) Separation of bands in ssDNA Template-B.

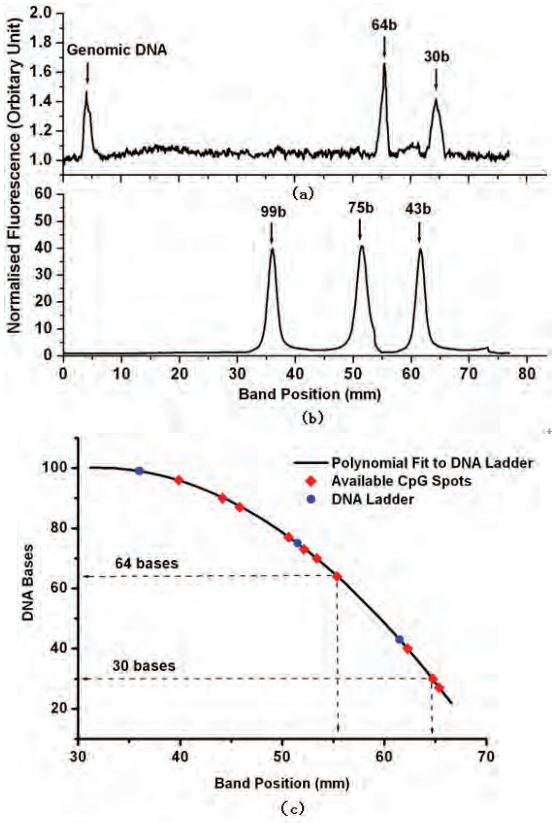


Figure 4. (a) Methyl sequencing of target gene in human genomic sample using SLCM. Human genomic DNA amplified at the target sequence in Exon-1 region of E-cadhrein, mapped using methyl insensitive in SLCM; (b) Resolution of three known length ssDNA templates runs concurrently with genomic DNA as control; (c) Calibration of DNA band separation distance versus base number. Blue solid circles represents control ladder of known number of bases. Continuous line is fit to the data and red diamonds represents possible CpG positions within the target region.

of the gene. Striking outcome of the approach presented here is that along with specifically developed methyl sequencing protocol, simple ways to methyl sequence the target gene and access information on the status of methylation away from the promoter region is demonstrated.

5. ACKNOWLEDGEMENTS

We thank S. Navaneethakrishnan for his help. This work was supported by Department of Science and Technology, Government of India Grant No. DST/TSG/PT/2006/63.

REFERENCES

[1] Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69(6)**, 915-926.

[2] Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell*, **128(4)**, 683-692.

[3] Toyooka, S. and Shimizu, N. (2004) Models for studying DNA methylation in human cancer: A review of current status. *Cancer*, **1(1)**, 37-42.

[4] Post, W.S., *et al.* (1999) Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. *Cardiovascular Research*, **43(4)**, 985-991.

[5] Anselmo, N.P., *et al.* (2006) Epigenetic alterations in human brain tumors in a brazilian population. *Genetics and Molecular Biology*, **29(3)**, 413-422.

[6] Laird, P.W. (2003) The power and the promise of DNA methylation markers. *Nature Reviews Cancer*, **3(6)**, 253- 266.

[7] Eick, D., Fritz, H.J. and Doer, W. (1983) Quantitative determination of 5-methylcytosine in DNA by reverse-phase high-performance liquid chromatography. *Analytical Biochemistry*, **135(1)**, 165-171.

[8] Fraga, M.F. and Esteller, M. (2002) DNA methylation: A profile of methods and applications. *BioTechniques*, **33(3)**, 632-649.

[9] Jean-Michel, D., Jorg, T., Helene, J. and Ivo, G.G. (2004) De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Analytical Biochemistry*, **333 (1)**, 119-127.

[10] Grigg, G.W. and Clark, S.J. (1994) Sequencing 5-methylcytosine residues in genomic DNA. *Bioessays*, **16(6)**, 431-436.

[11] Herman, J.G., Graff, J.R., Myohanen, S., Nelkin, B.D. and Baylin, S.B. (1996) Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences*, **93(18)**, 9821-9826.

[12] Stach, D., *et al.* (2003) Capillary electrophoretic analysis of genomic DNA methylation levels. *Nucleic Acids Research*, **31(2)**, e2.

[13] Michaela, W.D.S., Hans-Christian, K., Manfred, W. and Oliverl J.S. (2004) Determination of the DNA methylation level in tumor cells by capillary electrophoresis and laser-induced fluorescence detection. *Electrophoresis*, **25(6)**, 839-845.

[14] Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Research*, **22(15)**, 2990-2997.

[15] Warnecke, P.M., Stirzaker, C., Melki, J.R., Millar, D.S., Paul, C.L. and Clark, S.J. (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Research*, **25(21)**, 4422-4426.

[16] Sambrook, J. and Russell, D.W. (2000) Molecular cloning A laboratory manual. 3rd Edition, Chapter 6.

[17] Marianne, F., *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of National Academy of Sciences*, **89(5)**, 1827-1832.

[18] http://www.urogene.org/methprimer/index1.html

[19] Vaishnavi, V., Aarthi, R., Smitha, S. and Jaffar Ali, B.M. (2010) A simplified method to determine methylated cytosines in a target gene. *IEEE Explore Conference Proceedings*, in Press.

[20] Victoria, L.B., Kristina, I.M., Achim, E.K., Kenneth, J.L. G.Z. and John, W.B., (2006) Methylation-dependent fragment separation: Direct detection of DNA methylation by capillary electrophoresis of PCR products from bisulfite-converted genomic DNA. *Analytical Biochemistry*, **354(2)**, 266-273.

J. Biomedical Science and Engineering, 2010, 3, 501-508 **JBiSE**
doi: 10.4236/jbise.2010.35070 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Diffusive modelling of glioma evolution: a review

Alexandros Roniotis[1,2], Kostas Marias[1], Vangelis Sakkalis[1], Michalis Zervakis[2]

[1]Foundation for Research and Technology-Hellas, Institute of Computer Science, Heraklion, Greece;
[2]Technical University of Crete, Department of Electronic and Computer Engineering, Chania, Greece.
Email: {roniotis, kmarias, sakkalis}@ics.forth.gr; michalis@display.tuc.gr

Received 26 January 2010; revised 27 January 2010; accepted 24 February 2010.

ABSTRACT

Gliomas, the most aggressive form of brain cancer, are known for their widespread invasion into the tissue near the tumor lesion. Exponential models, which have been widely used in other types of cancers, cannot be used for the simulation of tumor growth, due to the diffusive behavior of glioma. Diffusive models that have been proposed in the last two decades seem to better approximate the expansion of gliomas. This paper covers the history of glioma diffusive modelling, starting from the simplified initial model in 90s and describing how this have been enriched to take into account heterogenous brain tissue, anisotropic migration of glioma cells and adjustable proliferation rates. Especially, adjustable proliferation rates are very important for modelling therapy plans and personalising therapy to different patients.

Keywords: Glioma; Brain Tumor; Diffusive Models; Proliferation; Invasion

1. INTRODUCTION

Glioma is a type of cancer of central nervous system that starts in the brain or in the spine. It is called a glioma because it arises from glial cells. The most common site of gliomas is the brain. Gliomas constitute more than 50% of all brain cancer cases. Despite extended research in this area, patients are rarely given more than 12 months survival time [1].

The diagnosis of gliomas can be done by MRI, CT, angiogram or biopsy, with MRI being the most common method. After glioma is diagnosed, treatment is directly necessary. However, the most important problem with diagnosis and treatment is that gliomas are characterized by invasiveness. This means that there are cells diffused beyond the imaged tumor, which cannot be visualized by common imaging techniques. Thus, even if the clinician allows some safety margin during resection around the imaged tumor, cancer is expected to recur.

These dispiriting results have forced researchers worldwide to work with understanding the glioma growth procedure and its special pathology. One of the rapidly emerging fields in glioma study is tumour growth modeling; researchers have been working on finding mathematical models that efficiently describe the glioma growth procedure.

1.1. Mathematical Modelling

A mathematical model uses mathematical language to describe a system. In the case of glioma, a mathematical description of how glioma grows is under research. At first someone has to bear in mind the basic features that any typical mathematical model in biology must possess [2]. These are schematically described in **Figure 1**. Firstly, the model should be initiated within a realistic biological state. Additionally, the modeled biological processes should be understood and discretized as much as possibly, meaning that steps and real biological parameters should be isolated. Continuing, it is essential to allocate a mechanism that could simulate these steps and incorporates these parameters. Specifically, this mechanism could be described by an equation. Going further, the next step is to study the model mathematically and come up with solutions that include realistic boundary and initial conditions. Lastly, after having acquired the theoretical results it is of great importance to go back in biological process with predictions, comments and suggestions for experiments that will either ascertain or disprove the developed model. At this level, model success is highly dependent on combining experimentation and theory together. Because, even if the experimental results indicate that the model is incorrect, this is the right way to reach a successful conclusion. In final consideration, mathematics is very important in biology, however, they must be treated with seriousness. If mathematics is used for solving any biological process, without thoroughly studying the biological background, it is very possible to come up with solutions that not only do not contribute to corroborated conclusions, but also do harm. As stated in [2], the theoretical literature abounds with many such articles.

Recently, with the explosion of biological sequence data, many biological sequence databases have redun-

dant sequences which can cause problems for data analysis. These redundant sequences cannot provide valuable information for analysis but detracts from the statistical significance of interesting hits. Moreover, processing these redundant sequences often requires more time and computational resources. Removing redundant sequences is undoubtedly very helpful for performing statistical analysis and accelerating extensive database searching [1]. And it is also a way to obtain the real protein families and their representatives from a large sequences dataset. Therefore, it is necessary to develop an appropriate algorithm to remove redundant sequences from a biological sequence database.

Hobohm and Sander's algorithm is a widely used algorithm in many redundant sequence removing programs. Hobohm and Sander's algorithm was firstly introduced by U. Hobohm *et al.* of EMBL laboratory in 1992. In 1998, Lissa Holm and Chris Sander developed a program based on this algorithm to generate a non-redundant protein database NRDB90 [2]. After that, other researchers developed some programs for removing redundant sequences on the basis of Hobohm and Sander's algorithm, such as CD-HIT and PISCES.

A lot of research is currently taking place in mathematically modeling the glioma growth procedure. An efficient mathematical model for gliomas could help researchers and clinicians to get a better understanding of tumour growth pathology. They could also to predict the aggressiveness of a patient's tumour and, thus, to better define the margins of a tumour, so as to use them when applying resection or radiotherapy. Moreover, by importing therapy parameters into the model, the clinician can predict which therapy scheme is expected to yield better results for the patient.

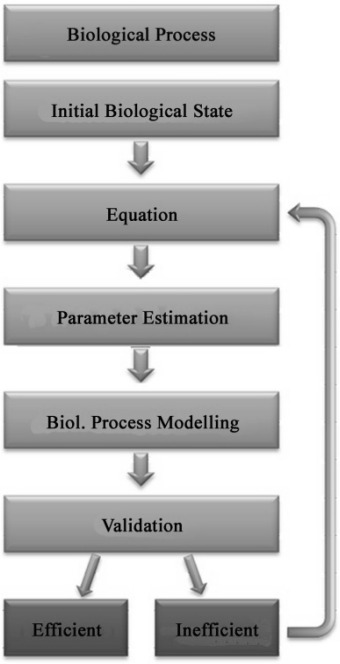


Figure 1. Mathematical modeling for a general biological process.

1.2. General Cancer Modelling

Studying tumour expansion and simulating this according to mathematical models, has been an area of studies in cancer since late 90s' [3-5]. Tumour growth has been studied by a series of models. The first models focused on tumour behavior in time. More specifically, the first proposed temporal models were based on either exponential, logistic or Gompertz laws [6]. As expected, these models were followed by spatial growth models in later years. Thus, one such deterministic model has been used to simulate cancer growth as a wave phenomenon, taking into account mitosis and nutritient depletion [7]. Moreover, deterministic models taking into account immune response [8] or mitotic rates changes [9] have been proposed.

The research efforts for applying these models in infiltrative cancers, such as glioma, failed because cell motility hadn't been included in the model. Thus, some stochastic and cellular automata have been introduced for glioma simulation, taking into account cell cycle, lack or abundance of nutritional elements in the surrounding area of cells and therapeutic regimen [10]. However, the mostly used models for the glioma case are the diffusive models that simulate the spatiotemporal change of glioma cell density, using partial differential equations. The combination of biomechanics and diffusive models is one of the latest advances in glioma modeling, for combining diffusion glioma cells and displacement of tissue caused by glioma growth [11].

2. THE DIFFUSION- REACTION MODEL

Unlike solid tumors, for which simple exponential or geometric expansion represents expansion of tumor volume, the glioma growth rate cannot be determined as the classical doubling rate [12], because gliomas can migrate and proliferate. In order to simulate glioma expansion, scientists have proposed the application of the diffusion-reaction equation, which is currently mostly used. The first that proposed a diffusive model for glioma growth was Murray in 1989 [13]. Murray derived the equation by exploiting the mass balance equation, by imagining cells as internal sources for producing the diffused cells. Murray proposed the diffusion-reaction formalism as:

$$\frac{\partial c(\mathbf{x},t)}{\partial t} \rightarrow -div\left(\underline{J}(\mathbf{x},t)\right) + \underbrace{S\left(c(\mathbf{x},t)\right) - T(c(\mathbf{x},t))}_{f(c(\mathbf{x},t))} \quad (1)$$

where

$c(\mathbf{x}, t)$ denotes the glioma cell concentration in position $\mathbf{x}$ at time t.

$\underline{J}(\mathbf{x}, t)$ is the diffusion flux of cell that follows Flick's law (*i.e.* $\underline{J} = -D\nabla c$ where ∇ and *div* are the gradient and divergence operators respectively).

$S(c)$ denotes the source factor, representing the glioma cell reproduction.

$T(c)$ denotes the treatment factor, representing the glioma cell loss due to treatment. This is zero, when no treatment is applied.

The initial state of the model, $c(x, 0)$, is defined as the initial distribution of cancerous cells.

$f(c) \equiv S(c) - T(c)$ is the net proliferation rate.

This equation was the basis of the later most important works for glioma modeling. The general procedure for glioma modelling, as derived by Murray's diffusive model, is presented in **Figure 2**.

The solution of the diffusion-reaction equation requires the application of numerical schemes, since there is no direct formula of its solution. The solution has to be approximated iteratively till time point of interest is reached.

2.1. Net Proliferation Rates for untreated gliomas

In 1995, Tracqui studied the evolution of cell concentration, by using two characteristic of tumour growth: proliferation and invasion [3]. Tracqui proposed that the cells proliferate at exponential rate ρ, *i.e.*:

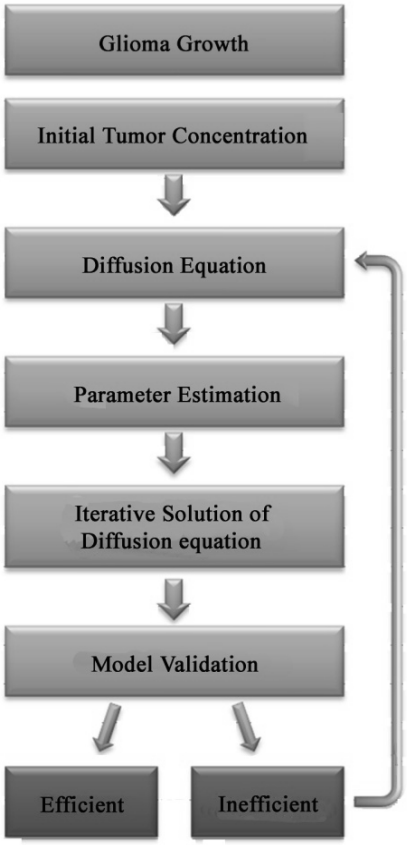


Figure 2. Generalised mathematical modeling of glioma growth.

$$f(c) = \rho \tag{2}$$

So, Tracqui changed Equation 1, to

$$\frac{\partial c}{\partial t} \rightarrow \nabla \cdot (D\nabla c) + \rho c \tag{3}$$

where ρ denotes the proliferation rate of cells. Other proposed models [14], instead of geometrical rate, used either Verhulst law

$$f(c) = \rho c \frac{c_m - c}{c_m} \tag{4}$$

or Gompertz law

$$f(c) = \rho c \ln \frac{c_m}{c} \tag{5}$$

where $\boldsymbol{c_m}$ is the maximum value that concentration can reach. **Eq.2** has been mainly used for simulating untreated gliomas.

2.2. First Estimation of D and ρ

From the very beginning of diffusive models, one of the key issues was the estimation of parameters that are being used. The diffusion coefficient $\boldsymbol{D}$ and the proliferation rate $\boldsymbol{\rho}$ are the basic parameters for the diffusive models. The first estimation of them is places in 1995 [1], when Silbergeld studied biological data and introduced two groups of glioma cells: the common ones and the resistant-to-first-chemotherapy ones. Parameter D was firstly defined either at $\boldsymbol{D = 10^{-2} cm^2/day}$, with the percentage of cells resistant to chemotherapy being at 8%, or at $\boldsymbol{D = 10^{-3} cm^2/day}$ without resistant cells, while ρ was defined at $\boldsymbol{\rho = 10^{-2}/day}$.

2.3. Resection Modeling

Getting this further, one of the next steps was to the model cancer evolution after ectomy [3,15]. This was firstly simulated in 1993 by setting the concentration of the ectomized area equal to zero and, then, allowing the surrounding malignant cells proliferate and diffuse until the sphere reaches 6cm diameter. An example of ectomy, reproduced from [3], is given in **Figure 3**.

2.4. Low Grade Gliomas

Up to 1996, diffusive models studied high-grade gliomas due to their remarkably fast invasion. However, studying low-grade gliomas was important as well. Hence, in 1996, Woodward [16] suggested that speed of growth in low-grade tumours should be 10% of the respective one in high-grade gliomas, yielding satisfactory results.

After some years, in 2003, Mandonnet [17] proposed that low-grade gliomas grew slowly, but linearly. This is mathematically derivable by **Eq.3**, because the expanding velocity of a population, which follows only the diffusion and growth laws of (3), can be calculated as

$2\sqrt{\rho D}$. Mandonnet *et al.* used clinical data reproduced from 27 patients to estimate that the average tumour velocity was 2 mm per year.

2.5. Brain Heterogeneity

Up to 2000, researchers didn't take brain anatomy into account. However, taking into account brain matter is of foremost importance, since it was observed that migration in white brain tissue is faster than in grey tissue [4]. Thus, modeling of gliomas should take brain heterogeneity into account. Indeed, in 2000, Swanson innovated by introducing the problem and incorporating white and gray matter differentiation in the diffusion coefficient D of **Eq.3**. More specifically, the equation continued to hold, but for variable D, *i.e.*:

$$\frac{\partial c}{\partial t} \rightarrow \nabla \cdot (D(\mathbf{x})\nabla c) + \rho c \qquad (6)$$

Diffusion coefficient D varies according to position, with $D(\mathbf{x}) = D_g$ or D_w, *i.e.* being constant for $\mathbf{x}$ in grey and white brain tissue respectively. Moreover, in order Swanson to apply the model, an brain atlas was required providing all the information about white and grey matter areas. Indeed, BrainWeb database [18] was available for extracting this information. An example of Swanson's simulation, reproduced from [4], is given in **Figure 4**.

2.6. Anisotropic Cell Migration

In 2005, Jbabdi *et al.* [19] introduced brain tissue anisotropy in diffusive modeling. As observed, glioma cell migration is facilitated along the directions of white matter fibers [20-21]. This observation can be supported by the diffusion tensor Magnetic Resonance imaging (DT MRI) that gives a very good 3D reconstruction of white matter fibers. Jbabdi defined **Eq.7**, as

$$\frac{\partial c}{\partial t} \rightarrow \nabla \cdot (\boldsymbol{D}(\mathbf{x})\nabla c) + \rho c \qquad (7)$$

where $\boldsymbol{D}$ is the diffusion tensor that describes cell diffusion rate at point $\mathbf{x}$, *i.e.* a 3 × 3 symmetric positive definite matrix that reflects local anisotropy.

2.7. Biomechanical Deformation

In 2005, Clatz *et al.* from INRIA [11] developed a model that simulates Jbadbi's model, but taking also into account the biomechanical deformations that occurs in brain due to tumour expansion. This model uses a predictor of the mass effect induced by both the tumor proliferation and infiltration. **Figure 5** shows an example of a simulation with Clatz's model, with the deformation of tissues presented. The image is reproduced from [11].

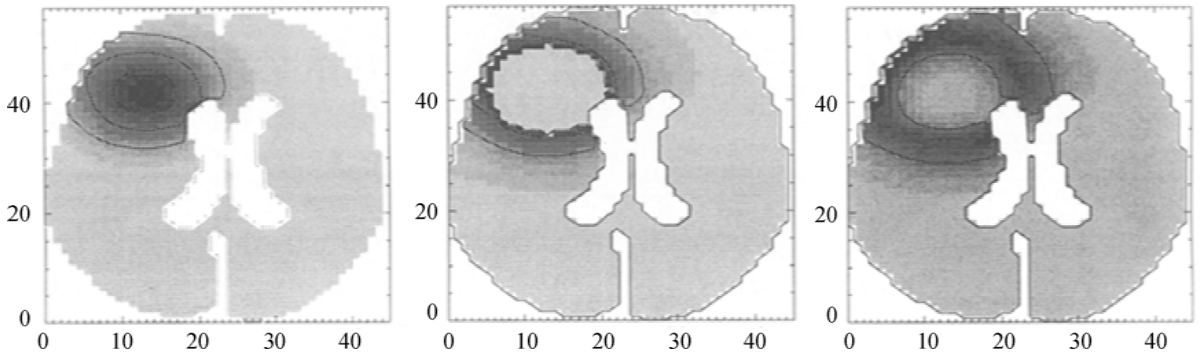
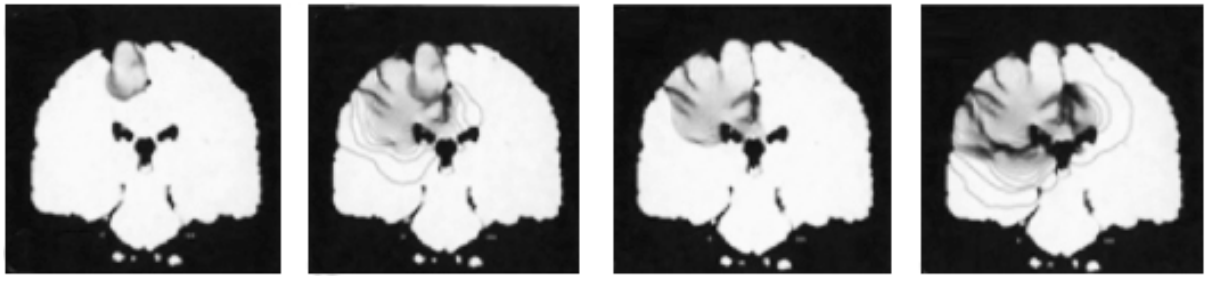


Figure 3. (Reproduced from [3])-tracqui's simulation of a tumour resection. The parameter values of the model are $\boldsymbol{\rho}$ = 0.012 /day, $\boldsymbol{D}$ = 10^{-7} cm^2/s.



(a) (b) (c) (d)

Figure 4. (Reproduced from [4])- Simulation of tumour invasion of a high-grade glioma in the superior cerebral hemisphere using Swanson's model: (a) (b) at diagnosis; (c) (d) at death; (a) (c) as seen by Swanson's standard threshold of detection; (b) (d) as calculated out to 1.25% of the threshold (boundary) cell concentration defining the sensitive threshold of detection.

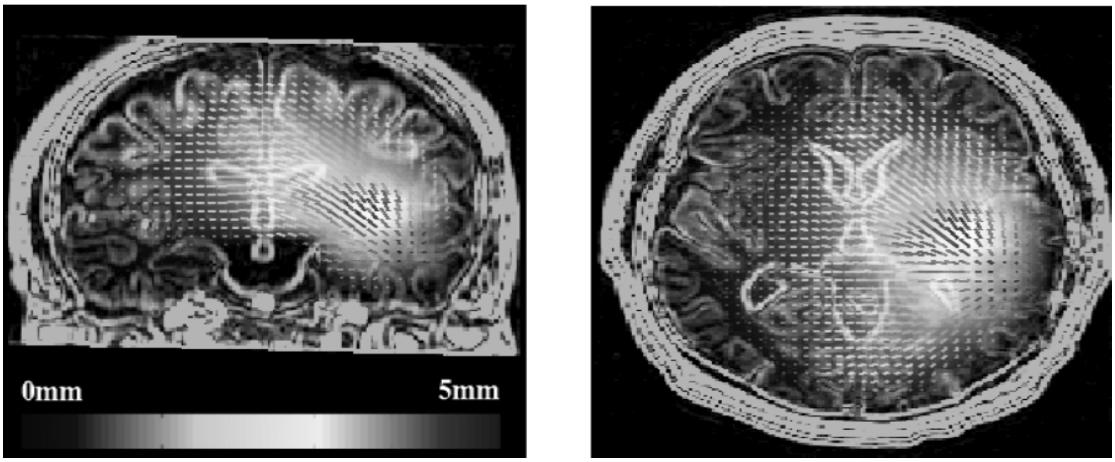


Figure 5. (Reproduced from [11])-Simulation of displacement of the tissues induced by the tumor mass effect, using Clatz's model.

3. Parameter estimation in modern models

According to [22], **Figure 6** can be used as a guiding index for defining parameters D and ρ according to glioma grade, velocities of growth and D/ρ. Indeed, this log to log graph includes all parameter values found up to Jan. 2007 for both low-and high-grade gliomas. Low-grade gliomas are sited in bottom left rectangle (LGG), for 2 mm/yr average velocities. Respectively, high grade gliomas (HCG) are positioned in the large rectangle, defined by D/ρ of 2 to 20 cm^2 and average velocities from 10mm/yr to 200mm/yr. On the left part gliomas with detectable mass, which can be cured with surgery, are placed.

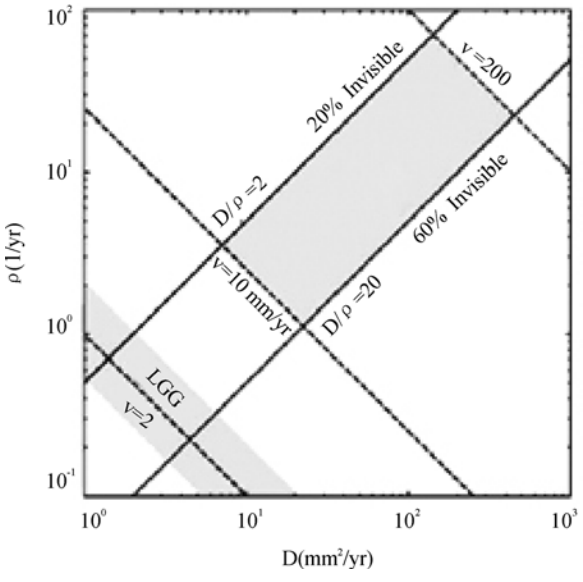


Figure 6. (Reproduced from [22])-Log-log graph of D and ρ, for high-and low-grade gliomas.

Continuing, for the case of heterogenous brain matter for high-grade gliomas, it is suggested in [4] that a typical value for ρ is ρ=0.0012 /day. This value for low-grade gliomas can be defined at ρ = 0.00012 /day. Then by assuming that $D_w = 5D_g$, $D_w = 10^{-2}$ mm^2/day and $D_g = 2 \cdot 10^{-3}$ mm^2/day can be used [19].

It is noteworthy to present the following table, reproduced from [23], where there is a list of the proposed values of each parameter that the latest glioma models use.

4. CHEMOTHERAPY MODELING

Chemotherapy modeling is a quite blurred part of glioma models, which is mostly studied currently by researchers. Parameters that have to be taken into consideration

Table 1. (Reproduced from [23])-The parameters of the diffusive model as proposed in the last 2 decades.

Parameter (namely)	Parameter Symbol	Value	Reference
Growth Rate	ρ	0.012 (1/day)	[24]
Diffusion Coefficient (Gray matter)	D_G	0.0013 (cm^2/day)	[3]
Diffusion Coefficient (White matter)	D_W	5 D_G (cm^2/day)	[4]
Initial number of tumor cells	N_0	10^5 cells	[16,25]
CT threshold density	-	400 (cells/mm^2)	[4]
CT threshold radius	-	1.5 cm	[24]
Cell death rate (chemotherapy)	k	0.0196 (1/day)	[26]
Number of fractions/ day	n	1-conventional (CR) 2-hyperfractionated (HF)	[27,28]
Time interval (step)	ΔT	1 day (CR) 6 hours (HFR)	[29,30]

should be extracted by histopathological and biostatistical data of specific patients [31]. Swanson has introduced a generalised net proliferation rate of chemotherapy as:

$$f(c) = \rho c - G(t)c \quad (8)$$

where $G(t)$ is the temporal profile of the chemotherapy treatments, assuming a loss proportional to the strength or amount of therapy at a given time. Swanson sets $G(t) = k$ when the chemotherapy is being administered and $G(t) = 0$ otherwise. k is actually a measure of effectiveness of chemotherapy. Thus, in order therapy to be effective and size of tumor to decrease, k should be larger than ρ, so that the net proliferation rate $f(c)$ is negative. This means the number of the dying cells is larger than the new born cells.

An example of chemotherapy application on real clinical MRI data, using (8) can be reproduced by [32]. In this example, k was intentionally set to the high value of $G(t) = 0.024$, so as the cancer to have shrinking effect. The data (18 MRI slices) has been acquired by Universität des Saarlandes Klinikum (Germany) within the scope of the Contracancrum project [33]. The modeling results are given in **Figure 7**. The 3-dimensional representation of the initial and after-chemotherapy states are presented, accompanied by a sampled series of 4 MRI slices.

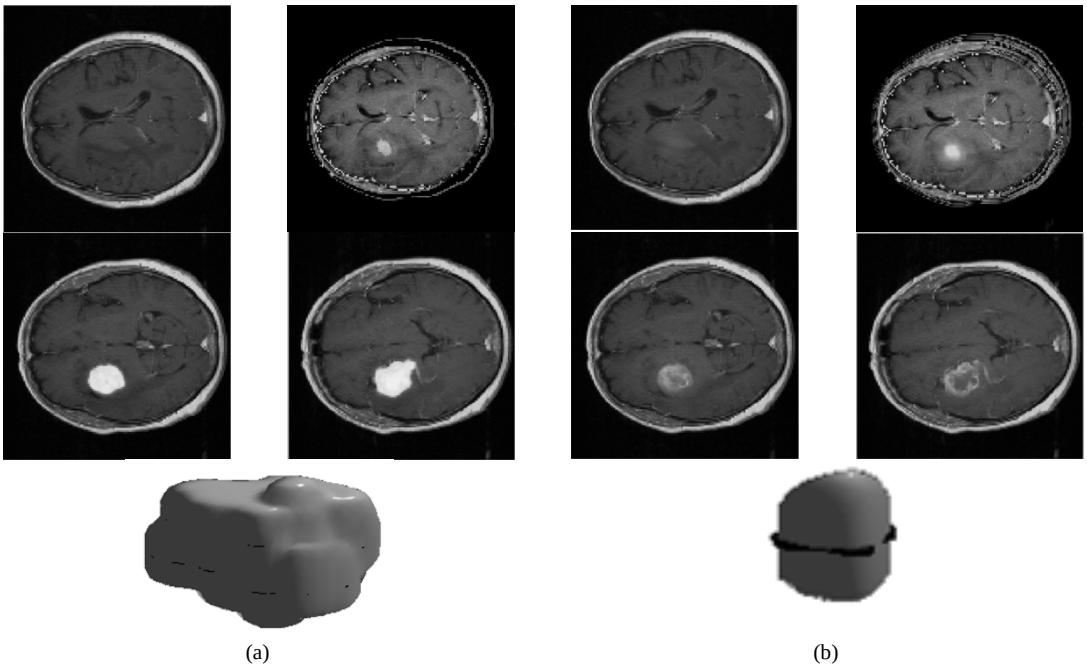
4.1. Mathematical Frameworks

The increasing interest of researchers on diffusive models was remitted by the total absence on specific guidelines on how to design them mathematically. In 2009, Roniotis *et al.* [34] published a mathematical framework on designing diffusive models by using Finite Differences, for the needs of the ContraCancrum Project [33]. The framework includes heterogeneous tissue, anisotropic migration of cells along white fibers and is 3D. Moreover, chemotherapy formalism of **Eq.8** can also be incorporated in the model.

Different schemes of finite differences have been developed, namely forward Euler, backward Euler, Crank Nikolson and θ-methods. Moreover, the same model has been also developed with finite elements. The accuracy of the different schemes has been tested on simplified cases which suggested that the backward Euler scheme (Finite Differences) yields the best results. An example of applying this framework on real clinical data is displayed in **Figure 8**, reproduced from [34].

5. CHALLENGES FOR FUTURE MODELS

Simulations of applying therapy can be performed by adjusting the proliferation term according to glioma cell proliferation rates in the diffusion equation. Moreover, consecutive chemotherapy sessions have been simulated by works. However, the main issue governing these applications is the estimation of the most efficient model parameters. This requires study of pharmacokinetics and study of real chemotherapy sessions on patients.



(a) (b)

Figure 7. (Reproduced from [32])-MRI and 3D representation of glioma in. (a) initial state; (b) after chemotherapy simulation. 4 slices out of 18 slices are presented.

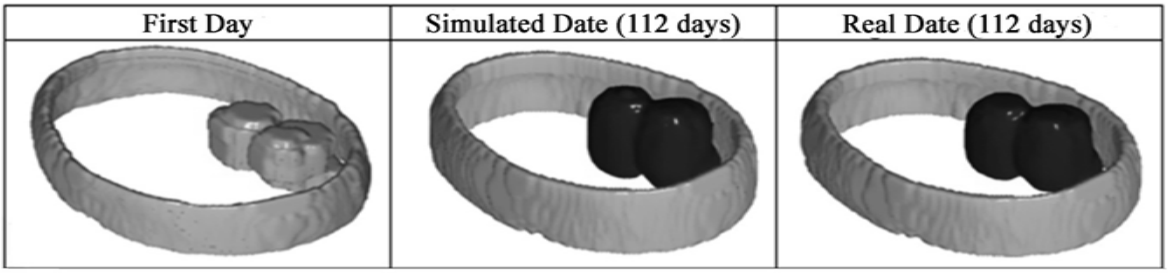


Figure 8. (Reproduced from [34])-results of modeling (left column) initial data, (central column) simulated data after 112 days, by using the backward euler scheme and (right column) real data on 112th day.

An ideal model will have the ability to suggest paths of invasion, unseen by the doctor and help the clinician predict how the tumor is going to behave after different cures. Personalized parameter estimation for treatment term, proliferation rate and initial state are very important for the model and are being studied by most researchers working on diffusive models globally.

Another important issue regarding modern diffusive models is the estimation of parameters of anisotropic diffusion. This is already incorporated in the models, but the methods used for extracting diffusion tensors have not been validated. This is mostly due to the use of DT-MRI, that is a modern technique and is hard to accompany all medical data. An estimation of tensors based on atlas could be studied.

6. CONCLUSIONS

Mathematical diffusive modeling for simulating glioma growth and invasion is a currently developing scientific field. An efficient 3-dimensional model of glioma growth will constitute a powerful tool for clinicians, since they could predict how glioma is going to develop in time. The most important part of this is the differentiation of the predicted outcome according to the different model input parameters. These parameters are changing with different therapy treatments. Thus the clinician can make an easier decision on which therapy scheme yields the best predictions.

7. ACKNOWLEDGEMENTS

This work was supported in part by the EC ICT project ContraCancrum, Contract No: 223979.

REFERENCES

[1] Silbergeld, D.L., Rostomily, R.C. and Alvord, Jr.E.C. (1991) The cause of death in patients with glioblastoma is multifactorial: Clinical factors and autopsy findings in 117 cases of supratentorial glioblastoma in adults. *Journal of Neuro-Oncology*, **10**, 179-185.

[2] Swanson, K.R., True, L.D. and Murray, J.D. (2003) On the use of quantitative modeling to help understand PSA dynamics and other medical problems. *American Journal of Clinical Pathology*, **119(1)**, 14-17.

[3] Tracqui, P. (1995) From passive diffusion to active cellular migration in mathematical models of tumour invasion. *Acta Bibliotheoretica*, **43**, 443-464.

[4] Swanson, K.R, Alvord, E.C. and Murray, J.D. (2000) A quantitative model for differential motility of gliomas in grey and white matter. *Cell Proliferation*, **33(5)**, 317-330.

[5] Burgess, P.K., Kulesa, P.M., Murray, J.D. and Alvord, E.C, (1997) The interaction of growth rates and diffusion coefficients in a three-dimensional mathematical model of gliomas. *Journal of Neuropathology & Experimental Neurology*, **56**, 704-713.

[6] Wheldon, T.E. (1986) Mathematical models in experimental and clinical oncology. In: Ingrain, D. and Bloch, R.F., Eds, *Mathematical Methods in Medicine*, John Wiley & Sons, Chichester, 1-32.

[7] Adam, J.A. and Maggelakis, S.A. (1990) Diffusion regulated growth characteristics of a spherical prevascular carcinoma. *Bulletin of Mathematical Biology*, **52**, 549-582.

[8] Lefever, R., Hiemaux, J. and Meyers, P. (1989) Evolution of tumours attacked by immune cytotoxic cells: The immune response dilemma. In: Goldbetor, A., Ed., *Cell to Cell Signaling: From Experiments to Theoretical Models*, Academic Press, Boston, 315-333.

[9] Sherratt, J.A. and Nowa, M.A. (1992) Oneogenes, anti-oncogenes and the immune response to cancer a mathematical model. *Proceeding of the Royal Society of London*, ***B* 248**, 261-271.

[10] Dionysiou, D., Stamatakos, G.S., Uzunoglu, N.K., Nikita, K.S., Marioli, A. (2004) A four-dimensional simulation model of tumour response to radiotherapy in vivo: Parametric validation considering radiosensitivity, genetic profile and fractionation. *Journal of Theoretical Biology*, **230(1)**, 20. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=15275995&query_hl=2

[11] Clatz, O., Sermesant, M., Bondiau, P., Delingette, H., Warfield, S.K., Malandain, G. and Ayache, N. (2005) Realistic simulation of the 3-D growth of brain tumours in MR images coupling diffusion with biomechanical deformation. *IEEE Transactions on Medical Imaging*, **24**, 1334-1346.

[12] Leighton, J. and Tchao R. (1984) The propagation of cancer, a process of tissue remodeling. *Cancer and Metastasis Reviews*, **3**, 81-97.

[13] Murray, J.D. (1989) Mathematical biology. Springer-Verlag, Heidelberg.

[14] Marusic, M., Bajzer, Z., Freyer, J.P. and Vuk-Palovic, S. (1994) Analysis of growth of multicellular tumour spheroids by mathematical models. *Cell Proliferation*, **27**, 73-94.

[15] Kreth, F.W., Warnke, P.C., Scheremet, R., *et al*. (1993) Surgical resection and radiation therapy versus biopsy and radiation therapy in the treatment of glioblastoma multiforme. *Journal of Neurosurgery*, **78**, 762-766.

[16] Woodward, D.E., Cook, J., Tracqui, P., *et al*. (1996) A mathematical model of glioma growth: The effect of extent of surgical resection. *Cell Proliferation*, **29**, 269-288.

[17] Mandonnet, E., Delattre, J.Y., Tanguy, M.L., *et al*. (2003) Continuous growth of mean tumour diameter in a subset of grade II gliomas. *Annals of Neurology*, **53**, 524-28.

[18] Cocosco, C.A., Kollokian, V., Evans, A.C., *et al*. (1997) Brainweb: Online interface to a 3D simulated brain database. *Neuroimage*, **5**, S425.

[19] Jbabdi, S., Mandonnet, E., Duffau, H., *et al*. (2005) Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging. *Magnetic Resonance in Medicine*, **54**, 616-24.

[20] Belien, A.T., Paganetti, P.A. and Schwab, M.E. (1999) Membrane-type 1 matrix metalloprotease (mt1-mmp) enables migration of glioma cells in central nervous system white matter. *Journal of Cell Biology*, **144**, 373-384.

[21] Giese, A., Bjerkvig, R., Berens, M.E. and Westphal, M., (2003) Cost of migration: Invasion of malignant gliomas and implications for treatment. *Journal of Clinical Oncology*, **21**, 1624-1636.

[22] Harpold, H.L.P., Alvord, E.C. and Swanson, Jr.,K.R. (2007) Visualizing beyond the tip of the iceberg: The evolution of mathematical modeling of glioma growth and invasion. *Journal of Neuropathology and Experimental Neurology*, **66(1)**, 1-9.

[23] Powathil, G., Kohandel, M., Oza, A. and Milosevic, M. (2007) Mathematical modeling of brain tumours: Effects of radiotherapy and chemotherapy, *Physics in Medicine and Biology*, **52**, 3291-3306.

[24] Cook, J., Woodward, D.E., Tracqui, P. and Murray, J.D. (1995) Resection of gliomas and life expectancy. *Journal of Neuro-Oncology*, **24**, 131.

[25] Murray, J. D. (2003) Mathematical biology I and II interdisciplinary applied mathematics. 3rd Edition, Springer, Berlin.

[26] Stupp, R., *et al*. (2005) Radiotherapy plus concomitant and adjuvant Temozolomide for glioblastoma. *The New England Journal of Medicine*, **352**, 987-996.

[27] Walker, M.D., Strike, T.A and Sheline, G.E (1979) An analysis of dose-effect relationship in the radiotherapy of malignant Gliomas. *International Journal of Radiation Oncology • Biology • Physics*, **5**, 1725-31.

[28] Werner-Wasik, M., Scott, C.B., Nelson, D.F., Gaspar, L.E., Murray, K.J., Fischbach, J.A., Nelson, J.S., Weinstein, A.S. and Curran, W.J.Jr. (1996) Final report of a phase I/II trial of hyperfractionated and accelerated hyperfractionated radiation therapy with carmustine for adults with supratentorial malignant gliomas (radiation therapy oncology group study 83-02). *Cancer*, **77**, 1535-43.

[29] Shibamoto, Y., Nishimura, Y., Tsutsui, K., Sasai, K., Takahashi, M. and Abe, M. (1997) Comparison of accelerated hyperfractionated radiotherapy and conventional radiotherapy for supratentorial malignant glioma, Japan. *Journal of Clinical Oncology*, **27**, 31-36.

[30] Nieder, C., Nestle, U., Niewald, M., Walter, K. and Schnabel, K. (1999) Hyperfractionated reirradiation for malignant glioma. *Frontiers of Radiation Therapy and Oncology*, **33**, 150-157.

[31] Blakenberg, F.G., Teplitz, R.L., Ellis, W., Salamat, M.S., Min, B.H., Hall, L., *et al*. (1995) The influence of volumetric tumor doubling time, DNA ploidy, and histologic grade on the survival of patients with interracial astrocytomas. *American Journal of Neuroradiology*, **16**, 1001-1012.

[32] Roniotis, A. (2010) Glioma growth modeling (internal report-submitted for review in the IEEE transactions on information and technology in biomedicine: Spectroscope issue on biomedicine information). FORTH-ICS, Heraklion.

[33] Marias, K., Sakkalis, V., Roniotis, A., Farmaki, C., Stamatakos, G., *et al*. (2009) Clinically oriented translational cancer multilevel modeling: The contracancrum project. *World Congress on Medical Physics and Biomedical Engineering*, Munich.

[34] Roniotis, A., Marias, K., Sakkalis, V., Tsibidis, G. and Zervakis, M. (2009) A complete mathematical study of a 3D model of heterogeneous and anisotropic glioma evolution. *IEEE Engineering in Medicine and Biology Society*, **1(185)**, 2807-10.

J. Biomedical Science and Engineering, 2010, 3, 509-516 **JBiSE**
doi:10.4236/jbise.2010.35071 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Advanced decision support for complex clinical decisions

Brain Keltch[1], Yuan Lin[1], Coskun Bayrak[2]

[1]Department of Applied Science, University of Arkansas at Little Rock, USA;
[2]Department of Computer Science, University of Arkansas at Little Rock, USA.
Email: bwkeltch@ualr.edu; yxlin1@ualr.edu; cxbayrak@ualr.edu

Received 1 March 2010; revised 11 March 2010, accepted 13 March 2010.

ABSTRACT

A Physician's decision-making skills are directly related to the patient's positive outcomes. Therefore, a wealth of medical knowledge and clinical experience are key assets for a physician to have. The goal here is to use historical clinical data and relationships processed by Artificial Intelligence (AI) techniques to aid physicians in their decision making process. Presenting this information in a Clinical Decision Support System (CDSS) is an effective means to consolidate decision results. The CDSS provides a large number of medical support functions to help clinicians make the most reasonable diagnosis and choose the best treatment measures. Initial results have shown great promise in accurately predicting Fibrosis Stage in Hepatitis patients. Utilizing this tool could mitigate the need for some liver biopsies in the more than 170 million Hepatitis patients worldwide. The prototype is extendable to accommodate additional techniques (for example genetic algorithms and logistics regression) and additional medical domain solutions (for example HIV/AIDS).

Keywords: Fibrosis; Clinical Decision Support; Decision Tree; Neural Network

1. INTRODUCTION

1.1. CDSS Definition

In hospital information systems (HIS), there are typically two main systems: Hospital Management Information Systems (HMIS) and Clinical Information Systems (CIS) [1]. HMIS support the hospital administration and transaction processing services while the CIS is used to support the clinical staff activities, to collect and dispose of clinical medical information, and to accumulate rich clinical knowledge. The CIS also provide clinical advice, support clinics, assistant clinical decision-making and to enhance staff efficiency. Clinical decision support systems (CDSS) are part of the CIS. It is an information system which uses expert systems and artificial intelligence (AI) technology to support clinical decision. It makes integrated diagnostic and medical advice bases on the collected patients' information, providing reference for the clinical medical officers.

1.2. Key Functions

Clinical decision support systems vary greatly in their complexity, function, and application. A Recent study [2] on health care information management four key functions of CDSS were outlined as follows:

1) Administrative: Supporting clinical coding and documentation, authorization of procedures, and referrals.

2) Managing clinical complexity and details: Keeping patients on research and chemotherapy protocols; tracking orders, referrals follow-up, and preventive care.

3) Cost control: Monitoring medication orders; avoiding duplicate or unnecessary tests.

4) Decision support: Supporting clinical diagnosis and treatment plan processes; and promoting use of best practices, condition-specific guidelines, and population-based management.

Our project will focus on item four, the decision support function and, in particular, utilization of historical laboratory data and outcome data processed through artificial intelligence tools. The combination of historical data and predictive tools provides valuable information in the hands of physicians as they develop a course of treatment for a patient.

2. BACKGROUND: AI TECHNIQUE IN CDSS

Decisions about medical treatment are best made by a trained and experienced physician. These decision makers can benefit from historical data and artificial intelligence tools. Computer scientists have dreamed of creating an "electronic brain" [3]. Computer scientists and doctors alike have been captivated by the potential such a technology might have in medicine [4]. With intelligent computers able to store and process vast stores of knowledge, the hope was that they would become perfect 'doctors in a box', assisting or surpassing clinicians

with tasks like diagnosis [3].

Today the importance of diagnosis as a task requiring computer support in routine clinical situations receives much less emphasis. The strict focus on the medical setting has now broadened across the healthcare spectrum, and instead of AI Medical systems, it is more typical to describe them as CDSS [3].

In our project, we evaluated and compared two techniques that will be core forecasting tools in a CDSS, which are Data Mining and Neural Networks.

The main purpose of doing data mining and knowledge discovery on the medical database is to predict disease and disease classification. Classification and prediction are two forms of data analysis which can be used to describe the model of the important data type or predict the future trends of the data [5].

Commonly used data mining algorithms are: association rules, decision trees, rough sets, statistical analysis, neural networks, support vector machines, fuzzy clustering, Case-Based Reasoning (CBR), Bayesian forecasting and visualization technology [6]. The common methods used in auxiliary diagnosis of clinical disease are 1) Bayes discriminate analysis 2) artificial neural network 3) decision tree.

However within the context of the study, the focus will be concentrated on decision tree and neural network.

2.1. Decision Tree

The decision tree is a very efficient machine learning classification algorithm. It is the origin in the concept of learning systems CLS, and then progress to ID3 method. In the end, it evolved to c5.0 which can handle continuous attributes. Well-known decision tree methods are CART and Assistant [5].

Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. In these tree structures, leaves represent classifications and branches represent conjunctions of features that lead to those classifications. One of the biggest advantages of the method is that the learning process does not require the user to understand a lot of background knowledge [6].

Nonetheless, data mining is a complex process to identify the useful information from large data sets. Although it is common to focus on the development, analysis and application of algorithms, the data selection and data pre-processing are the most timeconsuming activity in the entire data mining process, which affects the process and results [6].

2.2. Neural Network

Artificial Neural Networks have been proven to build efficient rule extraction/classification and forecasting applications. They provide a powerful non-linear machine learning techniques and are able to extract relevant features from large data sets. They are able to utilize and compare equally quantitative and qualitative data which is common in the clinical environments [7]. Neural Networks can handle redundant features as weights are learned from the training data.

The primary disadvantage of neural networks is that they will always arrive at a solution for any data set. This means that the quality of the resulting model is highly dependent on the quality and breath of the training data. It is easy to over-train the model, so that it can only predict the training data set. This can be minimized by assuring that convergence stopping strategies are effective and that the training data set is representative of the solutions space. Also, the resulting neural network solution is simply a node structure with inter-node weights. This requires validation of the output by additional statistical methods (*i.e.* decision trees) that are more understandable by subject matter experts.

3. AI ASSISTED CLINICAL DECISION SUPPORT SYSTEM

This study will focus on the demonstration and incorporation of neural network and decision tree techniques into a Clinical Decision Support System. These two AI techniques were selected because of their complementary attributes. Neural networks provide little definition of their predictive result, while decision tree output provides clear connections to historical data. Both methods will provide different information to the physician. Other AI methods should be considered in future work.

The database we plan on using for this study was collected at Chiba University hospital in Japan, and is a Practice of Knowledge Discovery in Databases (PKDD) 2005 Discovery Challenge dataset [11]. The data set contains patient data, laboratory data, and liver biopsies data on 771 hepatitis B and C patients. The goal will be to evaluate whether laboratory examinations can be used to estimate the stage of liver fibrosis. If this is possible, physicians may be able to use laboratory examinations as substitutes for biopsies and to aid in the treatment scenario. Liver biopsy is an invasive procedure and entails risk to patients. Decision tree and neural network methods based on a historical dataset will aid physicians in the development of treatment plans for Hepatitis patients.

The overall architectural representation of the system is shown in **Figure 1**.

The system is designed to be utilized by a physician to assist them in the development of a treatment plan for Hepatitis B and Hepatitis C patients. This system addresses an important question for the treatment – "Should a liver biopsy procedure be conducted?" This procedure provides important information on the advancement

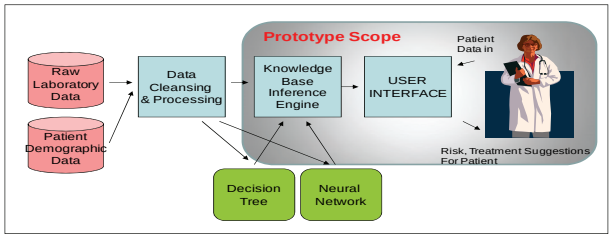


Figure 1. AI assisted clinical decision support system.

of the disease and fibrosis formation which are the key outcome measures. These fibrosis stage results determine the treatment protocol. This study attempts to predict fibrosis stage base on laboratory and patient data. The methodology and approach of processing data, applying AI techniques, and development of the resulting knowledge base can be utilized as a pattern for other medical treatment needs represented in a CDSS.

3.1. Data Processing and Cleaning

The sample hepatitis data set for our study is derived from the ECML/PKDD 2005 Discovery Challenge found at [11]:

The hepatitis dataset contains the results of laboratory examinations taken on the patients of hepatitis B and C, who were admitted to Chiba University Hospital in Japan. Hepatitis A, B and C are virus infections that affect the liver of the patient. Hepatitis B and C chronically inflame the hepatocyte, whereas hepatitis A acutely inflames it. Hepatitis B and C are especially important because they have a potential risk of developing liver cirrhosis or hepatocarcinoma. An indicator that can be used to know the risk of cirrhosis or hepatocarcinoma is fibrosis of hepatocyte. For instance, liver cirrhosis is characterized as the terminal stage of liver fibrosis.

We utilized three tables for our study, as show in **Tables 1-3**, below:

The MID field provided a common link between the three tables. Since the question we were addressing is to evaluate whether laboratory examinations can be used to estimate the stage of liver fibrosis our goal was to obtain one table that contained patient information, fibrosis stage and laboratory data. We chose to utilize in-hospital laboratory data, as it was more complete and we felt it would have better controls on quality and consistency.

Because of the large volume of data all files were translated into MS Access tables to process. The following steps were utilized:

1) Select all records from **Table 2** - 649 records.

2) Link to **Table 1** based on Patient ID - 649 records.

3) Link to the Laboratory Examination **Table 3** based on Patient ID. Select records where laboratory examinations were available for at least 90% of our biopsy patients, resulting in 16 examination parameters.

4) Link to and Match appropriate lab examination results (**Table 3**) with patient information and fibrosis stage information (the combined **Tables 1** and **2**). Select the laboratory examination that was earlier that the biopsy, but not more than 30 days earlier. This resulted in 425 records with complete data on the 16 laboratory examination results, patient information, and fibrosis stage biopsy results.

Table 1. Basic information of patients: pt_e030704.csv (total 771 records).

Item	Meaning	Remark
MID	Identification of the patient (masked)	unsigned integer
PT Sex	Sex of the patient	M(male) or F (female)
PT BirthDate	Birth date of the patient	YYYYMMDD

Table 2. Results of biopsy: bio_e030704.csv (total 694 records).

Item	Meaning	Remark
MID	Identification of the patient (masked)	unsigned integer
BIOPSY Lo_No	Identification of the specimen	unsigned integer
BIOPSY Hepatitis_Type	Type of hepatitis	B or C
BIOPSY Exam_Date	The date when biopsy was performed	YYYYMMDD
BIOPSY Hepatitis_Subtype	Subtype of Hepatitis	(Data concerning this field were dropped for simplicity)
BIOPSY Facility	Name of the facility where the specimen was collected	(Data concerning this field were dropped for simplicity)
BIOPSY Fibrosis	Biopsy report about progress of fibrosis	Discrete values: 0(F_0; no fibrosis)-4(F_4; severe)
BIOPSY Activity	Biopsy report about activity of virus	Discrete values: 0(A_0; no activity)-3(A_3; severe), or FALSE (activity could not be specified) Note that FALSE does NOT represent 'no activity', but represent 'activity could not be specified'. According to the donator, FALSE should be treated as a missing value.

Table 3. Results of in-hospital examinations: ilab_e030704.csv.

Item	Meaning	Remark
MID	Identification of the patient (masked)	
ILAB Exam_Date	Date of the examination	YYYYMMDD
ILAB Exam_No	ID for each examination performed repeatedly on the same day (n-th examination)	
ILAB Exam_Name	Name/code of the examination	
ILAB Exam_Result	Result of the examination	

3.2. Methods and Analysis

The objective is to use the decision tree and neural networks to predict fibrosis stage from patient data and laboratory data. We have a data set (see data preparation) of 424 historical that we will use to train and validate a decision tree model and a neural network model. **Table 4** is a map of the data we will be using.

3.2.1. Decision Tree Analysis

Data Mining is an analytic process designed to explore data (usually large amounts of data - typically business or market related) in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. Decision Tree's are one methodology utilized in the data mining set of tools. The ultimate goal of data mining is prediction - and predictive data mining is the most common type of data mining.

For the project, Weka 3.4 [9] which is a collection of machine learning algorithms for data mining tasks is used. The algorithms can either be applied directly to a dataset or called from your own code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes. It also provides Decision Tree capability that we will use in this study.

Table 4. Data utilized for neural network.

INPUT LABORATORY DATA	
Abbreviation	Description
ALB	Albumin
ALP	alkaline phosphatase
CHE	Cholinesterase
CL	Chloride
CRE	Creatinine
D-BIL	bilirubin, direct
G-GTP	gamma-glutamyltranspeptidase
G.GL	gamma-globulin
I-BIL	bilirubin, indirect
K	Potassium
T-BIL	bilirubin, total
T-CHO	cholesterol, total
TP	protein, total
TTT	thymol turbidity test
UA	uric acid
UN	blood urea nitrogen
INPUT PATIENT DATA	
Abbreviation	Description
AGE	Patient Age at Time of Test
SEX	Male = 1 or Female = 2
HEP_B_C	Hepatitis B = 1 or Hepatitis C = 2
OUTPUT FIBROSIS STAGE DATA	
Abbreviation	Description
FIBSTAGE	Fibrosis stage (0 to 4)

3.2.1.1. Decision Tree Testing Procedure

Step 1: Data preprocessing

The input data file needs to be organized in the form of ARFF in order to be processed in the Weka environment. In the file, all the values for the attributes needed to be filled in. If there is any missing value for an attribute, a "?" is used for substitution.

The **Figure 2** shows the attributes and the patient's information whose Masked ID is 1. In the input file, even the data for the predicted attribute Biopsy Fibrosis needs to be completed for the training data set. For the analysis 392 patients' information was used to develop the decision tree module and predict the other 32 patients' fibrosis stage.

The remaining 32 patients' data was formatted into a test ARFF file. In this file the fibrosis is substituted with a "?". **Figure 3**, below shows the patient's information whose MID is 907.

Step 2: Build the Model and get Decision Tree

We used the C4.5 algorithm to construct the decision tree. The root node is the Biopsy Fibrosis with five branches which present the 5 levels where the different laboratory values and patient characteristic are assigned one level at a time. **Figure 4** shows a graphical presentation of the complete decision tree model of the data set. The **Figure 5** shows a graphical presentation of the branch of the decision tree model where CHE <= 4.48.

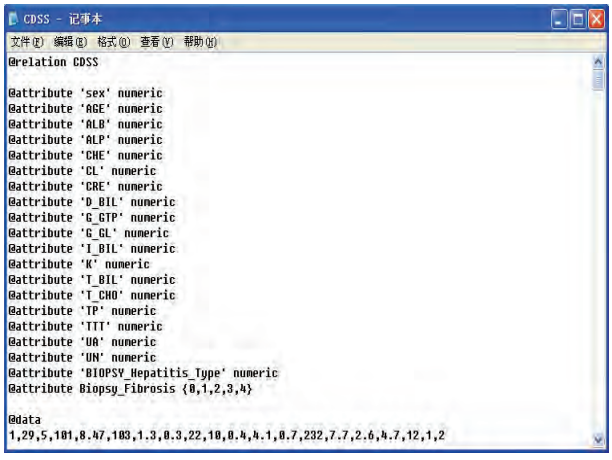


Figure 2. The WEKA training data file.

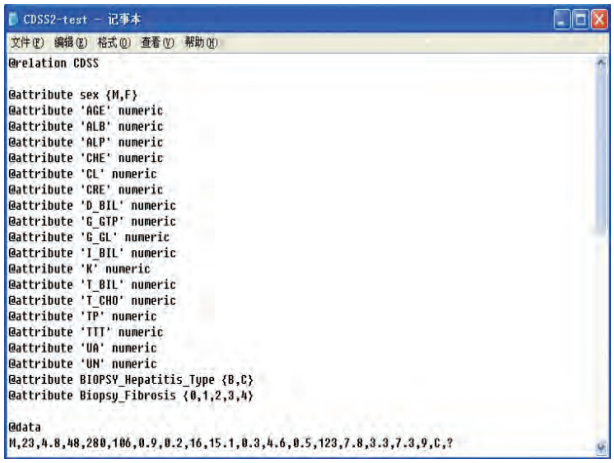


Figure 3. WEKA test data file.

Step 3: Prediction

The result of predicting the values with the constructed decision tree model is shown in **Table 5**. The table contains 32 patients' fibrosis stage values; the first column is the actual fibrosis stage, α, from the biopsy; the second is the predict values, β, by the decision tree model; and the third column shows the difference, $\gamma = \alpha - \beta$.

The decision tree model results showed an accuracy of 37.5% (12/32) of correct fibroid prediction. Predicting fibrosis values within a range of +or – one of the actual fibroid stage showed an accuracy of 91% (29/32).

3.2.1.2. Suggestions to Improve Decision Tree Accuracy

1) An increase in the number of training data examples will increase the correctness. Hence we can get a more powerful model, if we have a comprehensive training data set.

2) It is possible that obtaining additional laboratory attributes will increase the correctness. We were able to use only 16 laboratory attributes to predict the result. This was limited by the original data availability. We deleted some laboratory data since the incomplete data may also affect the test result.

3) The parameter setting for the decision tree algorithm utilized by Weka software is entered manually. We used the default value for the test because of limited time for the analysis. We may be able to improve results by testing additional parameters.

3.2.2. Neural Network Analysis

Artificial neural networks (ANNs) are systems that are constructed to use some organizational principles resembling those of the human brain. They are information processing systems that demonstrate the ability to learn, recall, and generalize from training patterns or data. ANNs are good at tasks such as pattern matching and classification, data clustering, and forecasting.

Table 5. Decision tree prediction quality.

Known Fibrosis Stage (α)	Predicted Fibrosis Stage (β)	Difference = (Known-Predicted) γ = α – β
1	0	1
1	0	1
2	2	0
2	2	0
1	1	0
4	3	1
3	2	1
3	2	1
1	2	−1
1	1	0
1	0	1
0	1	−1
3	3	0
2	4	−2
1	1	0
2	1	1
1	1	0
1	1	0
4	3	1
2	3	−1
1	2	−1
2	1	1
2	3	−1
1	1	0
1	1	0
1	4	−3
2	1	1
2	2	0
1	3	−2
1	1	0
1	0	1
1	2	−1

For the experiment we used a freeware tool called Neuro 3 [10] which uses the back propagation neural network (BPN). The term backpropagation refers to the training method by which the weights of the network connection are adjusted. The calculations procedure is feedforward, from input layer through hidden layers to output layer. During training, the calculated outputs are compared with the desired values, and then the errors are backpropagated to correct all weight factors.

All Training factors are defined by the users, including

- Number Hidden Layers
- The Threshold Value
- Transfer functions
- Learning Rate
- Momentum Coefficient
- Maximum Iterations
- Convergence Criteria

Network and training parameters are stored in scenario files and on projects spreadsheets which may be copied and pasted to other spreadsheet programs.

3.2.2.1. Neural Network Testing Procedures

The process for testing the predictive capabilities of the Neural Network includes the following steps:

Step 1. Divide the historical data set into a Training Data Set and a Prediction Data Set.

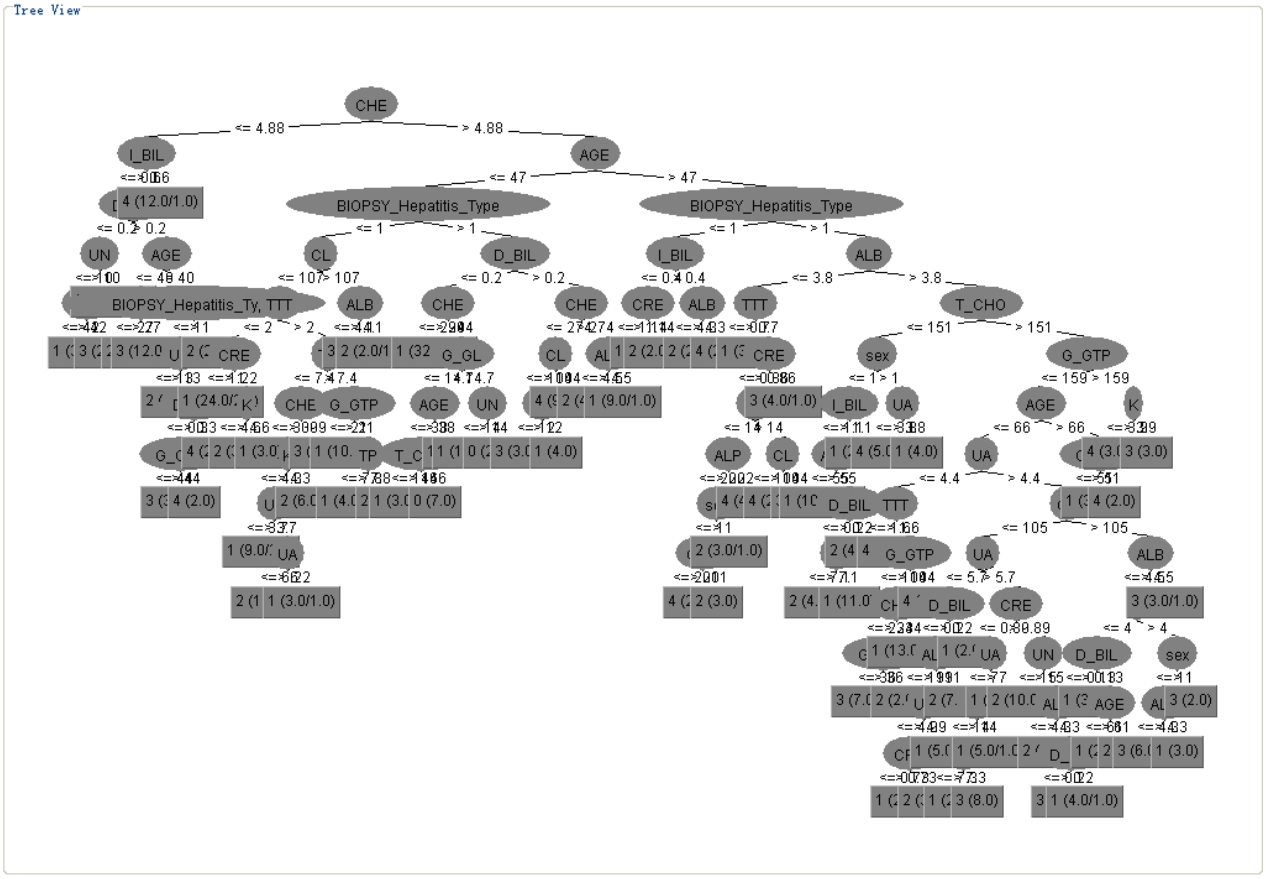


Figure 4. Decision tree.

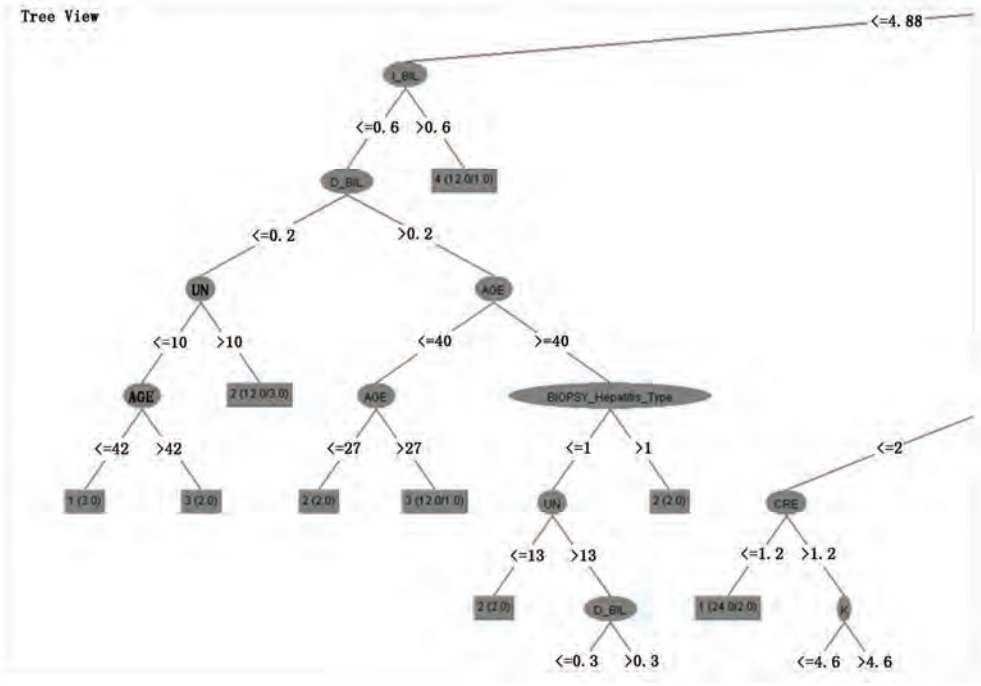


Figure 5. Section of the decision tree.

Step 2. Define the Neural Net architecture. The number of input nodes is defined by our 19 input (three patient data parameters and 16 laboratory parameters) parameters. And we have one output node (fibroid stage). The number of nodes in the hidden Layer is shown in the **Figure 6**.

Step 3. Define the run parameters. For this study we will only adjust the number of iterations that the error will be calculated and the weight adjusted. This error calculation and backpropagation will be executed 10,000 or 30,000 times.

Step 4. Run the neural network model on the training data set. The Neuro 3 tool provides a R-squared value and Sum of Errors for the training data set that provides some indication of the goodness of the model. A example of the Neuro 3 screen is shown in **Figure 7**.

Step 5. Utilize the trained neural network in Neuro 3 to predict the fibrosis stage for the prediction data set. Evaluate the goodness of fit by evaluating the percent of predictions that were correct, and that were within plus or minus 1, 2, or 3.

Step 6. Adjust the parameters in Step 1 through Step 5 until the optimal fit is achieved.

Step 7. Utilize the final selected model and train this model with all 424 historical data elements. Identify any deficiencies in the process or items that would be helpful for future work.

3.2.2.2. Testing and Validation Results

Eleven neural network models were generated. The best fitting model is Run Number 8, a 19-4-1 architecture trained with 10,000 iterations. This model provides fairly good predictive capabilities with 56% of predictions being correct ($\gamma = \alpha - \beta = 0$) and 90% being within +/- one fibroid stage ($\gamma = \alpha - \beta = \pm 1$). The run statistics are show below in **Table 6**.

The process of testing and validation of results revealed the following future considerations:

First, the increase in the amount of training data would greatly improve the predictive capability of the tool. This can be observed comparing Run Number 1 to Run Number 6. The application should allow for inclusion of additional data.

Second, additional data elements may increase the predictive capabilities of the tool. We only had sufficient data coverage for 16 laboratory data elements. Additional laboratory data elements should be acquired and evaluated utilizing these techniques.

Third, the analysis tools and selection of neural network architecture and run parameters was selected manually. This was a limitation of time available for this study. Automated analysis techniques and hybrid techniques combining neural network and genetic algorithms should be considered. An illustration of the sensitivity of predictive capabilities to hidden layer nodes shows in the **Figure 8** below. The vertical axis shows the percent of patients that the neural network predicted the correct fibrosis stage. The horizontal axis shows the number of hidden nodes specified in the neural network model, indicating that the accuracy of prediction is very sensitive to the number of hidden nodes in the model.

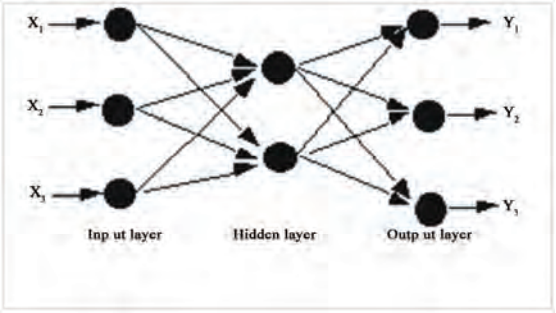
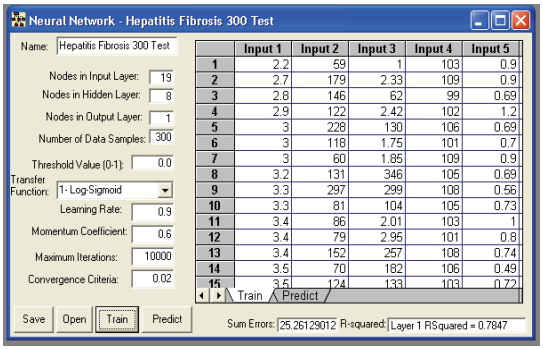


Figure 6. Architecture of a BPN neural network.



	Input 1	Input 2	Input 3	Input 4	Input 5
1	2.2	59	1	103	0.9
2	2.7	179	2.33	109	0.9
3	2.8	146	62	99	0.69
4	2.9	122	2.42	102	1.2
5	3	228	130	106	0.69
6	3	118	1.75	101	0.7
7	3	60	1.85	109	0.9
8	3.2	131	346	105	0.69
9	3.3	297	299	108	0.56
10	3.3	81	104	105	0.73
11	3.4	86	2.01	103	1
12	3.4	79	2.95	101	0.8
13	3.4	152	257	108	0.74
14	3.5	70	182	106	0.49
15	3.5	124	133	103	0.72

Train | Predict

Figure 7. Neuro 3 screen.

4. CONCLUSIONS

We have completed the development of a prototype that utilizes publically available patient data to address a single clinical decision–the prediction of fibrosis stage shown in **Figure 9**. The near-term customers for our project prototype are clinicians treating Hepatitis B and C patients.

Currently the key diagnostic tool in assessing the degree of liver disease in these patients is a liver biopsy. This procedure is invasive and requires the physician to

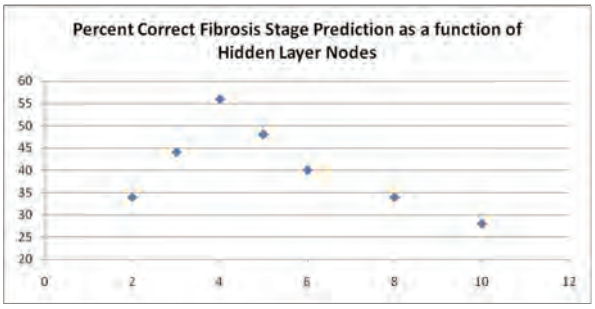


Figure 8. Correctness as a function of the number of hidden nodes.

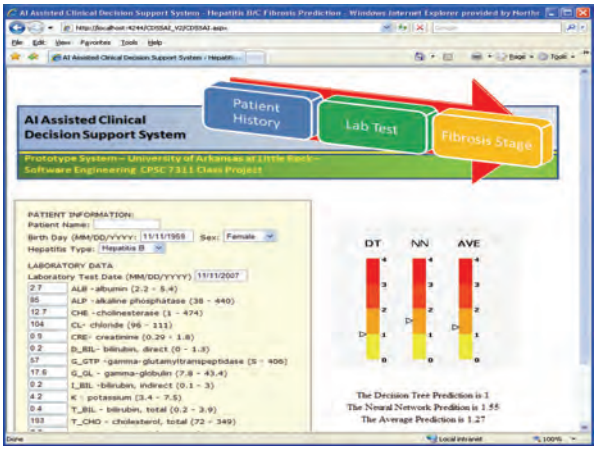


Figure 9. Prototype web application.

extract a small amount of the patient's liver using a fine needle and a suction syringe. While complications rates are low, < 3%, when they do occur they have large impacts. Also these procedures are expensive ($ 2,000 - $ 4,000 per test). Within the past several years several non-invasive tests have been developed as alternatives to biopsies. These tests generally require use of ultra sound equipment and/or specialized/proprietary blood test. These include: FibroScan, FIBROSpect II, FibroTest, FibroTest-ActiTest, and HCV-FibroSure. Cost for these test range from $ 400 to $ 700 [8].

The advantage of our method is that only standard liver panel blood tests are required to make the prediction of liver fibrosis. This provides low cost testing without any additional impact to the patient. Beyond the specific application in our prototype, linking AI tools

Table 6. Run statistics, run 8 is used for prediction.

Run Number	Neural Network Architecure	Training Data Set Number	Prediction Data Set Number	Numer of Iterations	Training R-squared	Training Sum of Errors	Prediction % Correct	Prediction % +/− 1	Prediction % +/− 2	Prediction +/− 3
1	19-8-1	300	124	10,000	0.81	25.5	30	71	92	100
2	19-8-1	300	124	30,000	0.86	19.6	27	72	84	97
3	19-6-1	300	124	10,000	0.72	32.2	38	76	90	99
4	19-6-1	374	50	30,000	0.66	43.1	40	76	92	98
5	19-6-1	374	50	10,000	0.68	39.3	34	74	92	98
6	19-8-1	374	50	10,000	0.74	34.7	34	82	98	100
7	19-10-1	374	50	10,000	0.78	31.3	28	64	94	98
8	19-4-1	374	50	10,000	0.56	50.3	56	90	96	100
9	19-2-1	374	50	10,000	0.39	64.1	34	90	100	100
10	19-5-1	374	50	10,000	0.62	45.1	48	84	96	100
11	19-3-1	374	50	10,000	0.47	57.4	44	88	98	100

with clinical decision support systems has wide applicability for other medical and healthcare solutions.

We have utilized Visual Studio 2008 for our prototype. With continued modifications beyond the prototype, we will utilize SQL Server 2008. Additionally we are utilizing two open source analytic tools for the AI component of our project. Namely, we are using "Neuro 3" a Visual Basic application for the neural network application and "Weka" for the decision tree portion of our project.

REFERENCES

[1] Begg, R. (2009) Artificial intelligence techniques in medicine and health care. *Concepts, Methodologies, Tools, and Applications*, Sugumaran, V., Ed., **48(12)**, 1-99.

[2] Perreault, L.E. and Metzger, J.B. (1999) A pragmatic framework for understanding clinical decision support. In: Middleton, B., Ed., Clinical decision support systems. *Journal of Healthcare Information Management*, **13(2)**, 5-21.

[3] Enrico, C. (2003) Guide to health informatics. 2nd Edition, Chapter 19, *Artificial Intelligence in Medicine*: *An Introduction*. Hodder Arnold, Arnold.

[4] Robert, S.L. and Lee, B.L. (1959) Reasoning foundations of medical diagnosis: Symbolic logic, probability, and value theory aid our understanding of how physicians reason. *American Association for the Advancement of Science*, **130(3366)**, 9-21.

[5] Špečkauskienė, V. and Lukoševičius, A. (2009) Methodology of adaptation of data mining methods for medical decision support: Case study. *Data Mining*, **9(2)**, 228-235.

[6] Han, J. and Kamber, M. (2001) Data mining concepts and techniques. 2nd Edition, Academic, Kluwer.

[7] Taylor, J.G. (1996) Neural networks and their applications. John Wiley and Sons, New York.

[8] http://janis7hepc.com/biopsies1.htm

[9] http://www.cs.waikato.ac.nz/ml/weka/

[10] http://www.keltch.com/neuro3.html

[11] http://lisp.vse.cz/challenge/ecmlpkdd2005/

J. Biomedical Science and Engineering, 2010, 3, 517-524 **JBiSE**
doi:10.4236/jbise.2010.35072 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Time dependent dispersion of nanoparticles in blood vessels

Francesco Gentile[1], Paolo Decuzzi[2]

[1]Center of Bio-Nanotechnology and -Engineering for Medicine, Università Magna Graecia, Catanzaro, Italy;
[2]Department of Nanomedicine and Biomedical Engineering, the University of Texas Health Science Center, Houston, USA.
Email: gentile@unicz.it

Received 25 September 2009; revised 20 October 2009; accepted 25 October 2009.

ABSTRACT

The dispersion of intravasculary injected nanoparticles can be efficiently described by introducing an effective diffusion coefficient D_{eff} which quantifies the longitudinal mass transport in blood vessels. Here, the original work of Gill and Sankarasubramanian was modified and extended to include 1) the variation over time of D_{eff}; 2) the permeability of the blood vessels and 3) non-Newtonian rheology of blood. A general solution was provided for D_{eff} depending on space (ζ), time (τ), plug radius (ξ_c) and a subset of permeability parameters. It was shown that increasing the vessel plug radius (thus hematocrit) or permeability leads to a reduction in D_{eff}, limiting the transport of nanoparticles across those vessels. It was also shown that the asymptotic time beyond which the solution attains the steady state behaviour is always independent of the plug radius and wall permeability. The analysis presented can more accurately predict the transport of nanoparticles in blood vessels, compared to previously developed models.

Keywords: Nanoparticle Transport; Casson Fluid; Permeable Blood Vessels; Drug Delivery

1. INTRODUCTION

The study of solute dispersion in capillaries dates back to the celebrated works of Taylor and Aris [1,2], who first studied the effect of shear stress on the transport in laminar flows. They provided a solution for the classic advection/diffusion equation

$$\frac{\partial C}{\partial t} + \mathbf{u}\cdot\nabla C = D_m \nabla^2 C \tag{1}$$

in the long term steady state limit, in terms of a constant effective coefficient of diffusion as

$$D_{eff} = D_m\left(1+\frac{P_e^2}{192}\right), \tag{2}$$

which includes the molecular diffusion contribution ($\propto D_m$) and the convective contribution ($\propto P_e$). In the **Eqs.1** and **2** above, P_e ($P_e = R_e \times u_0 / D_m$) is the Peclet number for a capillary with radius R_e and centerline velocity u_0, C is the local solute concentration; **u** is the fluid velocity vector; D_m is the Brownian or molecular diffusion coefficient and ∇ and ∇^2 are the gradient and Laplacian operators, respectively. The solution of Taylor and Aris is valid under the simplifying assumptions of 1) quasi-steady dispersion and 2) unidirectional flow. In particular, it is strictly valid beyond the asymptotic time $t_{st} = 1/2 \times R_e^2/D_m$. Notice that sub-micrometric particles with a molecular diffusivity D_m typically ranging between 10^{-11} and 10^{-9} m^2/s, in large vessels ($R_e \cong 10^{-2}$ m) would have t_{st} of the order of 10^5-10^7 s, whereas in small capillaries ($R_e \cong 10^{-6}$ m) t_{st} would fall in the range 10^{-3}-10^{-1} s.

Considerable efforts were expended in the attempt of relaxing the above assumptions. Gill [3] extended Taylor's formulation to obtain the local concentration C by means of a series expansion about the mean concentration, leading to the Generalized Dispersion Model (GDM), founding upon the rephrased convective-diffusive equation

$$\frac{\partial \Psi_m}{\partial \tau} = \sum_{i=0}^{\infty} K_i(t)\frac{\partial^i \Psi_m}{\partial \varsigma^i} \tag{3}$$

where $K_i(t)$ are suitable functions of time; Ψ_m is the normalized concentration averaged over a cross section of the capillary as explained in the sequel, ζ and τ are the longitudinal and time coordinates respectively. Sankarasubramanian and Gill [4] further developed the GDM including the effect of wall permeability to the solute (*i.e.* nanoparticles). In 1993, Sharp derived explicit expressions for the constant steady state coefficient D_{eff} for a non-Newtonian fluid considering, in particular, a Casson-like fluid [5]. Dash *et al.* [6] and Nagarani *et al.* [7] combined the model of Sharp and the GDM to obtain the unsteady dispersion in a Casson-like fluid, introducing solute adsorption to the walls. More recently, Decuzzi *et al.* [8] revisited the theory of Taylor and Aris incorporating the effects of wall permeability for the working fluid (plasma) and deriving a novel and more general

expression for D_{eff} being

$$D_{eff} = D_m\left[1+\frac{P_{e_0}^2}{192}\times f\left(\Omega,\Pi,\tilde{z}\right)\right] \quad (4)$$

where P_{e_0} is the Peclet number at the entrance of the capillary ($\tilde{z}=0$), and f is a function of the permeability parameter Π, pressure parameters Ω, and longitudinal coordinate $\tilde{z}$ along the capillary, as described in the sequel. In 2008, Gentile *et al*. [9] expanded the solution in [8] to include a Casson-like model for the fluid. Noticeably, the models presented in [8] and [9] are valid in the limit of large times of dispersion or, equivalently, at the steady state. No explicit dependency on time was introduced and the solution was deduced in terms of the longitudinal space coordinate solely.

In this work, the transport formulation proposed in [9] was further developed to account for the time dependency of the problem. The transport of nanoparticles was investigated and the effective diffusion coefficient D_{eff} derived. D_{eff} would in general depend upon the permeability of the capillary and the rheology of blood as in [9], but this dependency was extended to all times, thus also comprising the initial regime of dispersion. The model presented herein comprises, in the limits, well established schemes of diffusion.

2. MATERIALS AND METHODS

A circular capillary with radius Re and length l was considered as in **Figure 1**. A Casson-like fluid was considered with capillary walls permeable to the fluid, impermeable and not adsorbent to the solute (*i.e*. nanoparticles). In the following of the paper, the Generalized Dispersion Model was recalled and revised.

2.1. The Governing Equations

Following [4], the dispersion of a solute in a cylindrical capillary was described by the normalized advection-diffusion equation

$$\frac{\partial\Psi}{\partial\tau}+\upsilon\frac{\partial\Psi}{\partial\varsigma}=\left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right)+\frac{1}{P_{e_0}^2}\frac{\partial^2}{\partial\varsigma^2}\right)\Psi \quad (5)$$

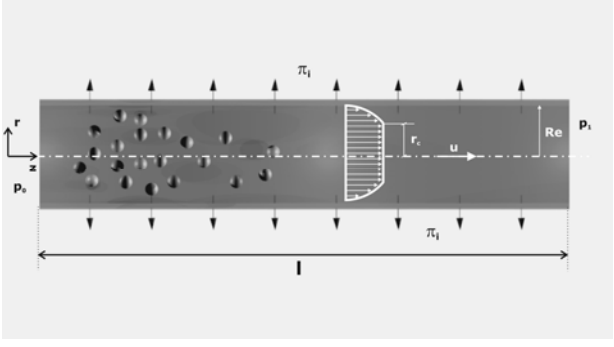


Figure 1. Longitudinal transport of molecules or nanoparticles in a blood capillary with a blunted velocity profile.

with the non dimensional terms being

$$\Psi=\frac{C}{C_0};\quad \upsilon=\frac{u}{u_0};\quad \rho=\frac{r}{R_e};$$
$$\varsigma=\frac{D_m z}{R_e^2 u_0};\quad \tau=\frac{D_m t}{R_e^2}; \quad (6)$$

where C is the local concentration of the solute and C_0 a reference concentration, u_0 is the initial center line velocity at the inlet and u the velocity distribution within the capillary with radius R_e, D_m is the molecular diffusivity of the solute, r and z are the radial and longitudinal coordinates as from the frame of reference in **Figure 1**, and t stays for the dimensional time. In **Eq.6** P_{e_0} (= $R_e \times u_0/D_m$) is the characteristic Peclet number defined as above. It was assumed that the particles are sufficiently small to have the same velocity of the dislodging fluid so that the diffusion/advection problem and the fluid-dynamic problem may be treated separately. The solution of **Eq.5** for Ψ can be derived *exactly* as

$$\Psi=\sum_{i=0}^{\infty} f_i\left(\rho,\varsigma;\tau\right)\frac{\partial^i\Psi_m}{\partial\varsigma^i} \quad (7)$$

where the functions f_i were related to the i-th derivative of Ψ_m as shown in the sequel. The mean concentration Ψ_m was defined as

$$\Psi_m = 2\int_0^1 \Psi\rho d\rho. \quad (8)$$

From **Eqs.5** and **7**, it follows that Ψ_m has to satisfy the relation

$$\frac{\partial\Psi_m}{\partial\tau}=\sum_{i=0}^{\infty} K_i\frac{\partial^i\Psi_m}{\partial\varsigma^i} \quad (9)$$

where the dispersion coefficients K_i were defined properly as function of time as to give

$$K_i(\varsigma,\tau)=\frac{\delta_{i2}}{P_{e_0}^2}+2\frac{\partial f_i}{\partial\rho}(1,\varsigma;\tau)- \\ =-2\int_0^1 f_{i-1}(\rho,\varsigma;\tau)\upsilon(\rho,\varsigma)\rho d\rho \quad (10)$$

with the understanding that $f_0 = 1$ and $f_1 = 0$. Here δ_{i2} denotes the delta of Kronecker. The dispersion problem was thus reduced to estimating f_i and K_i for each i. The auxiliary functions f_i must satisfy the differential equations

$$\frac{\partial f_n}{\partial\tau}=\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f_n}{\partial\rho}\right)-\upsilon f_{n-1}+\frac{f_{n-2}}{P_{e_0}^2}-\sum_{i=0}^{\infty}K_i\ f_{n-i}. \quad (11)$$

Relations **Eq.10** and **Eq.11** are coupled, and their solution becomes untractable of $i > 2$. Nevertheless it was shown [10] that all terms involving a coefficient higher than $i = 2$ in **Eq.10** can be neglected, in that K_2 is more than two orders of magnitude greater than K_3. **Eq.9** thus

reduces to the simplified relation

$$\frac{\partial \Psi_m}{\partial \tau} = K_1 \frac{\partial \Psi_m}{\partial \varsigma} + K_2 \frac{\partial^2 \Psi_m}{\partial \varsigma^2} \tag{12}$$

where K_1 and K_2 represent the convective and diffusive term, respectively. Notice that K_i and f_i depend upon the velocity field in the capillary $\upsilon(\rho, \zeta)$.

2.2. The Velocity Distribution in Permeable Capillaries

The velocity distribution in the capillary was given for a Casson-like fluid by [9]

$$\upsilon = -\frac{d\chi}{d\varsigma} \times \begin{cases} 1 - \frac{8}{3}\xi_c^{1/2} + 2\xi_c - \frac{1}{3}\xi_c^2 & \text{for } \rho < \xi_c \\ -1 + \rho^2 + \frac{8}{3}\left(1 - \rho^{3/2}\right)\xi_c^{1/2} - 2(1-\rho)\xi_c & \text{for } \rho > \xi_c \end{cases} \tag{13}$$

where ξ_c is the ratio between the plug radius r_c and the radius of the capillary R_e ($\xi_c = r_c/R_e$) and $d\chi / d\zeta$ is the pressure gradient along ζ. From **Eq.13** the non dimensional flow rate was derived through integration over the cross section to give

$$\Theta = 2\pi \int_0^1 \upsilon(r)\rho d\rho = -\frac{\pi}{2}\frac{d\chi}{d\varsigma} A(\xi_c), \tag{14}$$

where

$$A(\xi_c) = 1 - \frac{16}{7}\sqrt{\xi_c} + \frac{4}{3}\xi_c - \frac{1}{21}\xi_c^4, \tag{15}$$

and the mean fluid velocity could be written as

$$\Upsilon = \frac{\Theta}{\pi} = -\frac{1}{2}\frac{d\chi}{d\varsigma} A(\xi_c). \tag{16}$$

In the limit of a Newtonian fluid ($\xi_c \to 0$), **Eq.16** yields the expected value $\Upsilon = 0.5$.

In permeable capillaries, the fluid flows laterally across the walls inducing a continuos reduction in mean fluid velocity along the capillary. Following [8,9], the normalized mean fluid velocity Υ was expressed as a function of the hydraulic conductivity L_p, the interstitial fluid pressure π_i, the inlet and the outlet vascular pressures p_0 and p_1, giving

$$\Upsilon = -\frac{1}{2}\frac{d\chi}{d\varsigma} A(\xi_c) = \frac{\cosh\left(\kappa\varsigma\Gamma(\xi_c)\right) - \Omega\cosh\left(\Gamma(\xi_c) - \kappa\varsigma\Gamma(\xi_c)\right)}{1 - \Omega\cosh\left(\Gamma(\xi_c)\right)} \frac{A(\xi_c)}{2}, \tag{17}$$

where Ω is a non dimensional pressure parameter

$$\Omega = \frac{p_0/\pi_i - 1}{p_1/\pi_i - 1}; \tag{18}$$

$\Gamma(\xi_c)$ is the permeability parameter given by

$$\Gamma(\xi_c) = \frac{4l}{R_e}\sqrt{\frac{\eta}{R_e}L_p}\frac{1}{\sqrt{A(\xi_c)}} = \frac{\Pi}{\sqrt{A(\xi_c)}}. \tag{19}$$

Notice that differently from [8], the permeability parameter Γ is not fixed and varies with ξ_c. Substituting back the **Eqs.17** and **13** to **Eqs.10** and **11**, the coefficients K_i were appropriately derived.

2.3. The Initial and Boundary Conditions

It was assumed that a bolus of nanoparticles is introduced instantaneously and uniformly at the initial time t = 0 into the capillary, that is

$$\Psi(\rho, \varsigma; 0) = \Psi_m(\varsigma; 0). \tag{20}$$

In addition, the walls are impermeable to the solute and no absorption occurs to lead to

$$\left.\frac{\partial \Psi}{\partial \rho}\right|_{\rho=1} = 0, \tag{21}$$

symmetry at the centerline imposed

$$\left.\frac{\partial \Psi}{\partial \rho}\right|_{\rho=0} = 0, \tag{22}$$

and finally mass conservation was translated in mathematical terms as

$$\Psi|_{\varsigma\to\infty} = \left.\frac{\partial^i \Psi}{\partial \varsigma^i}\right|_{\varsigma\to\infty} = 0; \quad \Psi_m|_{\varsigma\to\infty} = \left.\frac{\partial^i \Psi_m}{\partial \varsigma^i}\right|_{\varsigma\to\infty} = 0; \quad i > 0. \tag{23}$$

The above relations should be also rephrased in terms of f_i to solve **Eq.11**, giving [4,6,7]

$$\int_0^1 f_i \rho d\rho = \delta_{0i}, \left.\frac{\partial f}{\partial \rho}\right|_{\rho=1} = 0, \left.\frac{\partial f}{\partial \rho}\right|_{\rho=0} = 0. \tag{24}$$

2.4. Solution for K_1 and f_1

Imposing n = 1 in **Eq.11**, it was derived

$$\frac{\partial f_1}{\partial \tau} = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial f_1}{\partial \rho}\right) - \upsilon - K_1, \tag{25}$$

and multiplying by ρ and integrating with respect to ρ from 0 to 1, invoking the first of **Eq.24**, it followed that

$$K_1 = -2\int_0^1 v\rho d\rho = -\Upsilon(\varsigma). \tag{26}$$

From **Eq.26**, it was deduced that the convective term K_1 equals the mean velocity Υ that is not constant along the capillary. Also notice that assuming a frame of reference moving with Υ, K_1 would be zero as in [6]. f_1 was found as a solution of the partial differential **Eq.25**

that can be decomposed as the sum the steady state solution $f_{1s}(\rho, \zeta)$ and the transient term $f_{1t}(\rho, \zeta; \tau)$.

$$f_1(\rho,\varsigma;\tau) = f_{1s}(\rho,\varsigma) + f_{1t}(\rho,\varsigma;\tau). \tag{27}$$

Substitution of the steady state term f_{1s} into **Eq.25** yield

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f_{1s}}{\partial\rho}\right) = \upsilon - \Upsilon; \tag{28}$$

which holds in the core of the capillary ($\rho < \xi_c$), where the velocity is blunted, and in the cell free layer ($\rho > \xi_c$), where the velocity varies with ρ. At the interface, $\rho = \xi_c$, continuity imposed that $f_{1s}(\rho = \xi_c^-) = f_{1s}(\rho = \xi_c^+)$ which, together with the boundary conditions **Eq.24**, allowed the deconvolution of f_{1s} as

$$f_{1s}(\rho,\varsigma) = \frac{\cosh(\kappa\varsigma\Gamma(\xi_c)) - \Omega\times\cosh(\Gamma(\xi_c) - \kappa\varsigma\Gamma(\xi_c))}{1-\Omega\cosh(\Gamma(\xi_c))} \times \begin{cases} B(\rho) \text{ for } \rho < \xi_c \\ C(\rho) \text{ for } \rho > \xi_c \end{cases} \tag{29}$$

where B and C are solely functions of ρ:

$$\begin{aligned} B(\rho;\xi_c) &= \frac{-8085 + 21600\xi_c^{1/2} - 15092\xi_c + 1430\xi_c^4}{194040} \\ &+ \frac{147\xi_c^6 + 1155(21 - 64\xi_c^{1/2} + 56\xi_c - 14\xi_c^2 + \xi_c^4)\rho^2}{194040} \\ &- \frac{2310\xi_c^4\ln(\xi_c)}{194040}, \end{aligned} \tag{30}$$

$$\begin{aligned} C(\rho;\xi_c) = &-\frac{1}{48}(2 - 6\rho^2 + 3\rho^4) \\ &+ \frac{4}{1617}(45 - 154\rho^2 + 88\rho^{7/2})\xi^{1/2} \\ &-(7 + 10\rho^2(2\rho - 3))\xi + \frac{(2\rho^2 - 3)}{336}\xi^4 \\ &+ \frac{\xi^6}{1320} - \frac{\xi^4}{84}\ln(r). \end{aligned} \tag{31}$$

The transient term f_{1t} depends upon f_{1s} and was readily derived as [6]

$$\sum_{n=0}^{\infty}\frac{J_0(\lambda_n\rho)\int f_{1s}(\rho,\varsigma)\rho d\rho}{J_0(\lambda_n)^2}e^{-\lambda_n^2\tau}J_0(\lambda_n\rho) \tag{32}$$

where J_0 and J_1 are the Bessel function of first type and order zero and one, respectively, and the eigenvalues λ_n were found as the roots of the equation $J_1(\lambda_n) = 0$.

2.5. Solution for K_2

Imposing n = 2 in **Eq.11**, multiplying by ρ and integrating with respect to ρ from 0 to 1, K_2 was obtained as

$$K_2(\varsigma;\tau) = \frac{1}{Pe_0^2} - 2\int_0^1 f_1(\rho,\varsigma;\tau)\upsilon(\rho,\varsigma)\rho d\rho \tag{33}$$

notice that, differently from the original formulation by Gill and Sankarasubramanian [4,10], the auxiliary functions K_2 would in general depend also on the longitudinal coordinate ζ and, in particular, the problem would be determined if the velocity field in the capillary is known. In the limit of large time K_2 is found as

$$\begin{aligned} &K_2(\varsigma)\big|_{\tau\to\infty} \\ &= \frac{1}{Pe_0^2} + \frac{1}{192}\frac{\cosh(\kappa\varsigma\Gamma(\xi_c)) - \Omega\times\cosh(\Gamma(\xi_c) - \kappa\varsigma\Gamma(\xi_c))}{1-\Omega\cosh(\Gamma(\xi_c))} \\ &\times\Bigg(1 - \frac{5888}{1555}\xi_c^{1/2} + \frac{558368}{56595}\xi_c - \frac{6144}{715}\xi_c^{3/2} + \frac{128}{45}\xi_c^2 \\ &+ \frac{244}{21}\xi_c^4 - \frac{272128}{3773}\xi_c^{9/2} + \frac{385312}{2205}\xi_c^5 - \frac{4096}{21}\xi_c^{11/2} \\ &+ \frac{11464}{165}\xi_c^6 + \frac{55808}{1155}\xi_c^{13/2} + -\frac{6976}{165}\xi_c^7 + \frac{430331}{66885}\xi_c^8 \\ &- \frac{512}{147}\xi_c^{17/2} + \frac{64}{21}\xi_c^9 - \frac{872}{1155}\xi_c^{10} + \frac{4}{147}\xi_c^{12} - \frac{8}{147}\xi_c^8\ln(\xi_c)\Bigg) \end{aligned} \tag{34}$$

thus recovering the results derived in [9]. Incidentally notice that **Eq.34** represents the most general formulation for the non dimensional coefficient of diffusion K_2 in that it comprises an extensive subset of solutions, depending on the rheological parameters ξ_c, Γ and Ω. In particular, as Γ (or, equivalently, Π) goes to zero (impermeable capillary) **Eq.34** coincides with the relation given in [5], whereas as the rheological parameter ξ_c goes to zero the result given by [8] is recovered. The classical solution of Taylor and Aris [1,2] is found when both $\Gamma(\Pi)$ and ξ_c are null.

3. RESULTS AND DISCUSSION

The most important coefficient for estimating the transport of nanoparticles is the normalized effective diffusion coefficient

$$K_2 = \left(\frac{D_{eff}}{D_m} - \frac{A(\xi_c)^2 - 1}{A(\xi_c)^2}\right)\frac{A(\xi_c)^2}{Pe_0^2} \tag{35}$$

in that it gives a measure of the propensity of the particles to spread about their center of mass along the capillary. Differently from all the schemes proposed so far, the K_2 presented in **Eq.35** changes with ζ due to the variation of the mean fluid velocity along the permeable vessel. In **Figure 2** the relation $192(K_2 - P_e^{-2})$ was plotted as a function of ζ and τ in the case of large permeability of the walls ($\Pi = 8$, $\Omega = -2$) and for a Newtonian

fluid ($\xi_c = 0$). Generally K_2 increases with time and attains the steady state value after the early stage of dispersion which corresponds to $\tau = 0.5$. A central position of the vessel was observed where $K_2 - P_e^{-2} = 0$, implying that in such area dispersion is solely driven by pure molecular diffusion. The decrease of K_2 with ζ strongly depends upon the permeability of the capillary (Π) and the plug radius of the fluid (ξ_c). In **Figure 3** the 3D plot of the relation $192(K_2 - P_e^{-2})$ as a function of time τ and position along the capillary ζ was displayed showing the effects of Π and ζ_c varying between 0 and 4 and 0 and 0.4 respectively, and for a constant $\Omega = -2$. In **Figure 4**, the contourplots corresponding to **Figure 3** were reported. As time increases, the solution for K_2 tends to a constant asymptotic value. Noticeably, the time beyond which dispersion turns to be time independent is always less than 0.5, regardless Π and ξ_c Therefore, the permeability parameter and the plug radius have a negligible effect upon the process of diffusion along with time but do effect on the steady state behavior of the system. In particular, when both Π and ξ_c are larger than zero the reduction in dispersion (D_{eff} or K_2) is dramatic, and in large portions of the capillary the transport of the nanoparticles is mostly diffusion limited. This is easily explained observing that longitudinal transport is enhanced by radial velocity gradients (shear diffusion), thereby either an increase of the core region of the capillary with a flat velocity profile (thus ξ_c) or a reduction in the velocity amplitude due to an augmented permeability (thus Π), generates a decrease in K_2, as thoroughly discussed in [8,9].

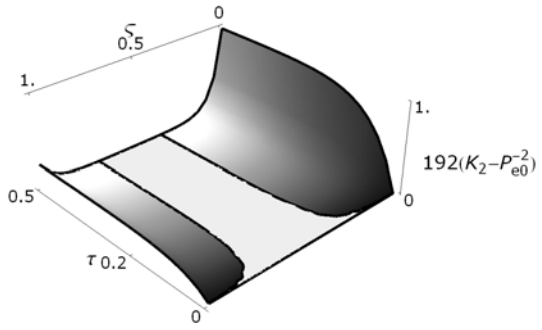


Figure 2. The dimensionless effective diffusion K_2 as a function of the normalized position (ζ) and time (τ) for a fixed plug radius $\xi_c = 0$ and for a permeable capillary ($\Pi = 8$, $\Omega = -2$).

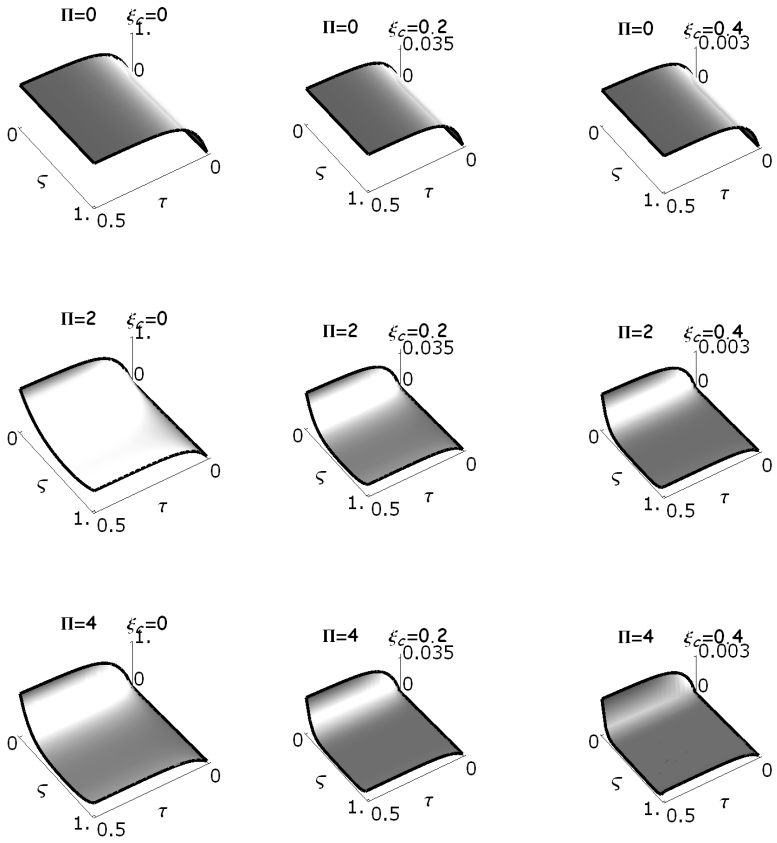


Figure 3. 3D plots of the dimensionless effective diffusion coefficients K_2 as a function of the normalized position (ζ) and time (τ), for Π and ξ_c varying between 0 and 4, and 0 and 0.4 respectively, and for a constant $\Omega = -2$.

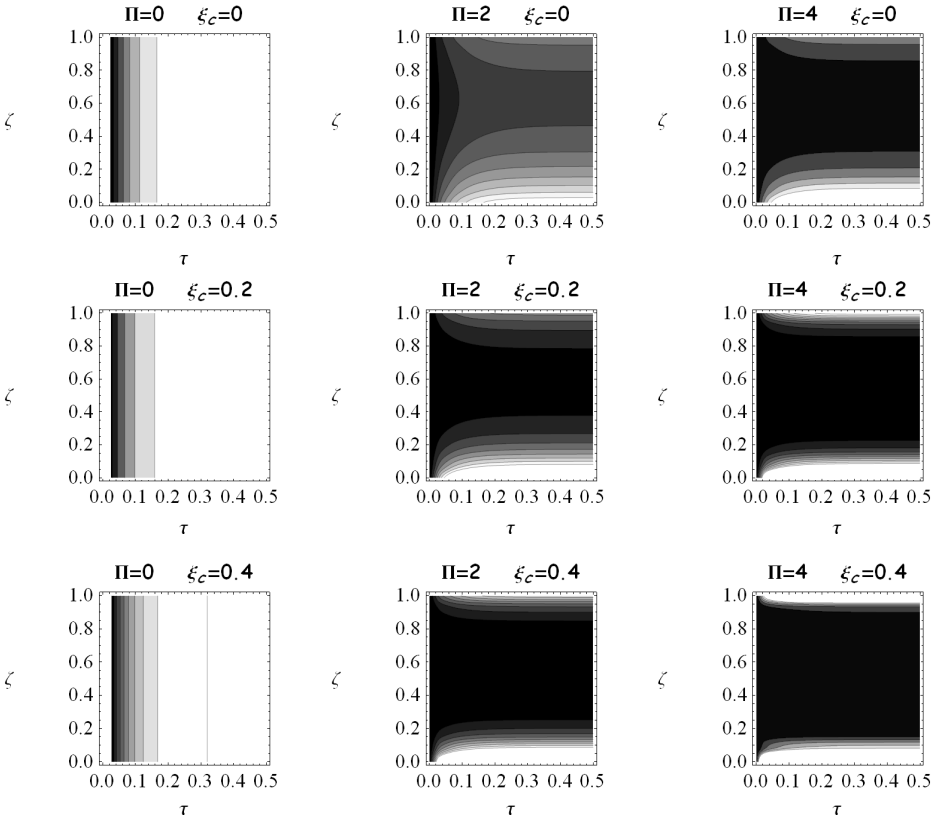


Figure 4. Contour plots of the dimensionless effective diffusion coefficients K_2 as a function of the normalized position (ζ) and time (τ), for Π and ξ_c varying between 0 and 4, and 0 and 0.4 respectively, and for a constant $\Omega = -2$.

Given K_2, the effective diffusion coefficient D_{eff} was deduced as

$$D_{eff} = \frac{R_e^2 u_0^2}{A(\xi_c)^2 D_m} K_2 + \frac{A(\xi_c)^2 - 1}{A(\xi_c)^2} D_m, \quad (36)$$

or, equivalently

$$\frac{D_{eff}}{D_m} = \frac{P_{e_0}^2}{A(\xi_c)^2} K_2 + \frac{A(\xi_c)^2 - 1}{A(\xi_c)^2}. \quad (37)$$

Eq.37 shows that any enhancement in effective diffusion over the Brownian diffusion (D_m) is proportional to the product $P_{e_0} \times K_2$ and would strongly depend on the local hydrodynamics and capillary size.

The dimensionless effective diffusion D_{eff}/D_m as a function of the rheological parameter ξ_c, for different values of P_e and for a fixed $\Pi = 0$ was shown in **Figure 5**. As expected, confirming the results derived in [9], larger P_e and smaller ξ_c lead to larger D_{eff}/D_m ratios. **Figure 6** illustrated the ratio D_{eff}/D_m over time, for an impermeable channel ($\Pi = 0$) and for different values of ξ_c. **Figure 7** reported the same diagram of **Figure 6** for a permeable channel ($\Pi = 2$, $\Omega = -2$). In all cases, a steady state value was attained for τ larger than 0.5. Notice that for $\Pi = 0$ and for ξ_c moving from 0 to 0.4, at large times the classical solutions of Taylor and Aris [1,2] ($\xi_c = 0$), and Sharp [4] ($\xi_c = 0.2, 0.4$) are recovered (**Figure 6**).When the permeable solution was instead considered (**Figure 7**), the steady state values recapitulated the results given by Decuzzi *et al.* [8] ($\xi_c = 0$) and Gentile *et al.* [9] ($\xi_c = 0.2, 0.4$).

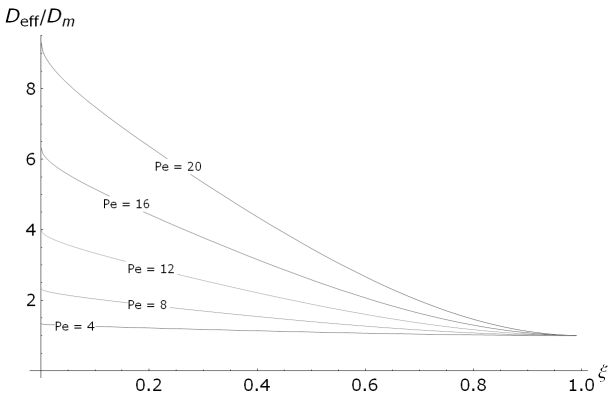


Figure 5. The dimensionless effective diffusion (D_{eff}/D_m) as a function of the rheological parameter ξ_c, for different values of Pe and for a fixed $\Pi = 0$.

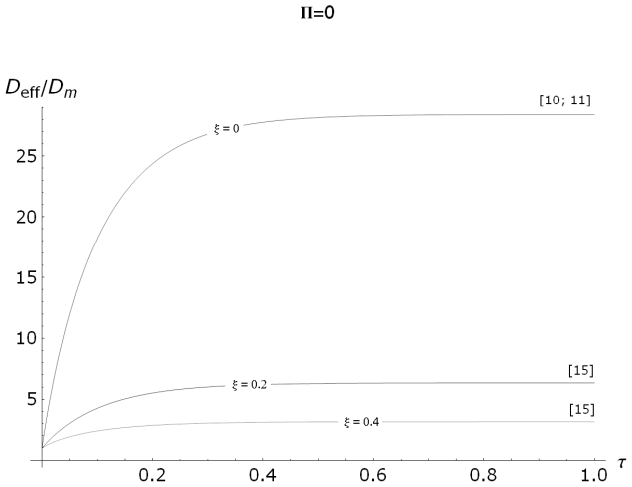


Figure 6. The ratio (D_{eff}/D_m) over time, for a permeable channel ($\Pi = 0$) and for different values of ξ_c.

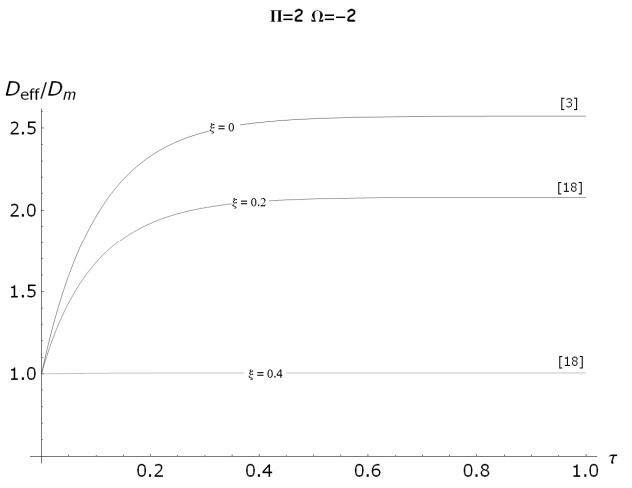


Figure 7. The ratio (D_{eff}/D_m) over time, for a permeable channel ($\Pi = 2$; $\Omega = -2$) and for different values of ξ_c.

Table 1. Average dimensions and velocities of blood vessels (Decuzzi, 2006 [8]). P_e is calculated for $D_m = 6 \times 10^{-13}$ m^2/s.

Vessel	L[mm]	R_e[mm]	U[mm/s]	P_e
Aorta	50	25	400	1.6×10^{10}
Artery	1.5-2	4	100	6.67×10^{8}
Arteriole	1.5-2	0.02-0.1	5	$1.67 - 8.33 \times 10^{5}$
Capillary	0.5	0.005-0.001	0.1-1	833-41667
Venules	1	0.02-0.05	0.5	$1.66 - 4.16 \times 10^{4}$
Vein	1-14	2-5	50	$1.6 - 4.1 \times 10^{8}$
Vena Cava	40-50	30	100	5×10^{9}

Recalling that the width of the plug radius ξ_c scales with R_e as $\xi_c \sim 1\text{-}3 \times R_e^{-0.8}$ [9] and considering the data of **Table 1**, moving from capillaries to arterioles and venules P_{e_0} significantly increases and the ratio D_{eff}/D_m augments accordingly despite a reduction of the cell free layer area. **Figure 8** showed the minimum value that D_{eff}/D_m would assume in a vessel at the steady state as a function of Re and of the plug radius ξ_c for an impermeable vessel ($\Pi = 0$). **Figure 9** reported the same diagram for a permeable vessel ($\Pi = 5$, $\Omega = -2$). It was observed

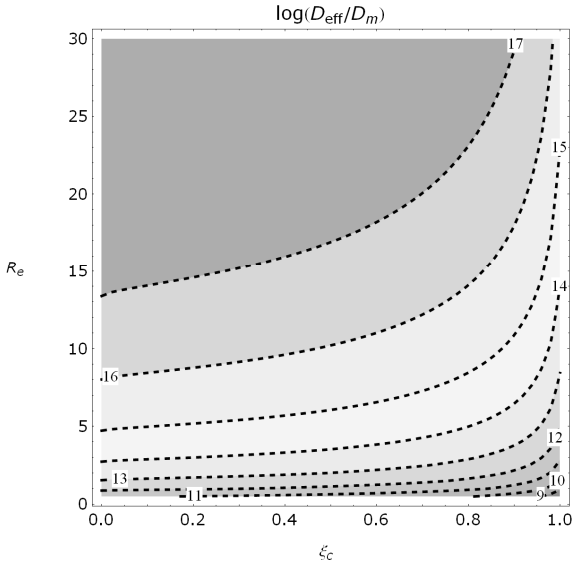


Figure 8. The minimum value that D_{eff}/D_m would assume in a vessel at the steady state as a function of Re and of the plug radius ξ_c; for an impermeable vessel ($\Pi = 0$).

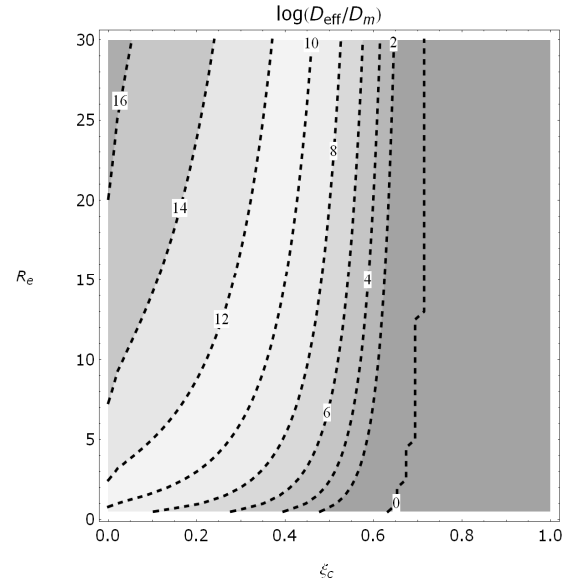


Figure 9. The minimum value that D_{eff}/D_m would assume in a vessel at the steady state as a function of Re and of the plug radius ξ_c; for a permeable vessel ($\Pi = 5$, $\Omega = -2$).

that the effect of the radius of the vessel (or equivalently of P_{e0}, see **Table 1**) dominates over that of the plug radius, meaning that in large capillaries, where ξ_c is large, yet the longitudinal diffusion increases up to 10^6 times with respect to small vessels. And this effect is dramatically amplified considering leaky or fenestrated capillaries. It was argumented in [8,9] that in a capillary network passively transported molecules or nanoparticles would follow the path with the largest effective diffusion. Therefore, nanoparticles and molecules would in a larger percentage stay in the macrocirculation (high D_{eff}) rather than in the microcirculation (small D_{eff}) or highly permeable vessels (even smaller D_{eff}), as for instance in the angiogenic tumor vasculature. This would constitute a barrier to the rational systemic administration of therapeutic and contrast agents. The correct design of nanoparticles could constitute an effective way to overcome this barrier. It was demonstrated, either experimentally [11-13] and theoretically [14], that particles having different sizes or shapes also have different margination properties, that is the attitude of "spontaneously" drifting towards the walls of the blood vessels. In particular, in considering the case of spherical particles, it was shown that the delivery efficiency is not size independent, instead larger particles would perform better than smaller ones under the effect of a gravitational or magnetic field [11]; whereas, below a characteristic diameter, the settling of nanoparticles would be mostly size independent [13]. In considering shapes other than spherical, it was demonstrated that, in the range of physiological relevant values of shear rates, inertial discoidal particles perform better than quasi-hemispherical and significantly better than spherical, and this circumstance would suggest the use of discoidal carriers in drug delivery [12,14].

In sight of the above findings, it is understandable that tailoring the shape and size of nanovectors inasmuch that they would tend to accumulate in the 'cell free layer', could significantly increase the efficiency of delivery.

4. CONCLUSIONS

The Generalized Dispersion Model firstly introduced by Gill and Sankarasubramanian was revised to account for blood rheology and vessel permeability. The non dimensional coefficient of diffusion was derived as a function of time, of the plug radius ξ_c and of a subset of permeability parameters, Π and Ω. It was observed that an enhancement in permeability or a blunted velocity profile (high hematocrit) dramatically reduces vascular transport. It was seen that an augmented permeability at the vessels walls does not influence the time in correspondence of which the dispersion process attains the steady state. Evidence was given that freely administrated drugs or nanoparticles very harshly would leave the macrocirculation in favour of leaky capillaries of tumor districts. Strategies for the avoidance of this physiological barrier were proposed.

REFERENCES

[1] Taylor, G. (1953) Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London, A*, **219(1137)**, 186-203.

[2] Aris, R. (1956) On the dispersion of a solute in a fluid flowing through a tube. *Proceedings of the Royal Society of London, A*, **235(1200)**, 67-77.

[3] Gill, W.N. (1967) A note on the solution of transient dispersion problems. *Proceedings of the Royal Society of London, A*, **298(1967)**, 335-339.

[4] Sankarasubramanian, R. and Gill, W.N. (1973) Taylor diffusion in laminar flow in an eccentric annulus. *Proceedings of the Royal Society of London, A*, **333**, 115-132.

[5] Sharp, M.K. (1993) Shear-augmented dispersion in non-Newtonian fluids. *Annals of Biomedical Engineering*, **21(4)**, 407-415.

[6] Dash, R.K., Jayaraman, G. and Mehta, K.N. (2000) Shear augmented dispersion of a solute in a casson fluid flowing in a conduit. *Annals of Biomedical Engineering*, **28(4)**, 373-385.

[7] Nagarani, P., Sarojamma, G. and Jayaraman, G. (2004) Effect of boundary absorption in dispersion in casson fluid flow in a tube. *Annals of Biomedical Engineering*, **32(5)**, 706-719.

[8] Decuzzi, P., Causa, F., Ferrari, M. and Netti, P.A. (2006) The effective dispersion of nanovectors within the tumor microvasculature. *Annals of Biomedical Engineering*, **34(4)**, 633-641.

[9] Gentile, F., Ferrari, M. and Decuzzi, P. (2008) The transport of nanoparticles in blood vessels: The effect of vessel permeability and blood rheology. *Annals of Biomedical Engineering*, **2(36)**, 254-261.

[10] Gill, W.N. and Sankarasubramanian, R. (1970) Exact analysis of unsteady convective diffusion. *Proceedings of the Royal Society of London, A*, **316**, 341-350.

[11] Decuzzi, P., Gentile, F., Granaldi, A., Curcio, A. and Indolfi, C. (2007) Flow chamber analysis of size effects in the adhesion of spherical particles. *International Journal of NanoMedicine*, **2(4)**, 1-8.

[12] Gentile, F., Chiappini, C., Fine, D., Bhavane, R.C., Peluccio, M.S., Cheng, M., Liu, X., Ferrari, M. and Decuzzi, P. (2008) The margination dynamics of non spherical inertial particles in a microchannel. *Journal of BioMechanics*, **41(10)**, 2312-2318.

[13] Gentile, F., Curcio, A., Indolfi, C., Decuzzi, P. and Ferrari, M. (2008) The margination propensity of spherical particles for vascular targeting in the microcirculation. *Journal of Nanobiotechnology*, **6(9)**, 1-9.

[14] Lee, S.Y., Ferrari, M. and Decuzzi, P. (2009) Design of bio-mimetic particles with enhanced vascular interaction. *Journal of Biomechanics*, **42(12)**, 1885-1890.

J. Biomedical Science and Engineering, 2010, 3, 525-528 **JBiSE**
doi:10.4236/jbise.2010.35073 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

In vitro evaluation of a new resilient, hard-carbon, thin-film coating as a bearing material for ventricular assist devices
——In Vitro Bearing Evaluation of BioMedFlex

Nicole A. Mielke[1], Alex L. Massiello[1], David J. Horvath[1], Stephen M. Benefit[1], Darren Burgess[2], Leonard A. R. Golding[1], Kiyotaka Fukamachi[1]

[1]Department of Biomedical Engineering/ND20, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA;
[2]BioMedFlex, LLC, Denver, North Carolina, USA.
Email: fukamak@ccf.org

Received 1 December 2009; revised 14 December 2009; accepted 20 December 2009.

ABSTRACT

Our aim was to evaluate the potential use of BioMedFlex® (BMF), a new resilient, hard-carbon, thin-film coating, as a blood journal bearing material in Cleveland Heart's continuous-flow left and right ventricular assist devices (VADs). BMF is not classified as a diamond-like carbon (DLC) and differs from other thin-film carbon coatings by its high flexural strength, radiopacity, and wear resistance. A 2- to 4-µm-thick BMF adhesion layer was deposited on the VAD journal bearing surfaces. A commercial DLC coating used in other clinical blood pump applications was used as a control. Durability and reliability of the BMF coating was verified in severe pump start/stop testing using 20 BMF-coated journal bearing pairs. The BMF-coated surfaces showed no coating failures, whereas 57% of the DLC bearing pairs developed scratches through the carbon coating, documenting that BMF can provide a durable coating in our blood journal bearing application. In conclusion, BMF has shown qualities that support its significant advantages as an alternative journal bearing material in Cleveland Heart pumps. Our plan includes biocompatibility testing with ongoing animal studies, endurance testing with submerged pumps running in saline, and assessment of batch coating processing capability.

Keywords: Heart Assist Device; Diamond-Like Carbon; Materials Testing; Wear Resistance; Journal Bearing

1. INTRODUCTION

Implantable ventricular assist devices (VADs) are a reliable treatment option for patients with terminal-stage heart failure who are unresponsive to conventional therapies. Although current designs demonstrate adequate performance, reliability remains a significant issue, as these devices are implanted for extended durations (well beyond 1 year) and may likely be intentionally implanted in patients as permanent ("destination") therapy. Because components of implantable pumps must withstand difficult *in vivo* environments and running conditions, carbon materials (diamond-like carbon [DLC] and pyrolytic carbon) have emerged as promising coating materials for pump components to enhance durability and biocompatibility characteristics. Some of these carbon coatings are chemically inert, wear and corrosion resistant, and bio- and hemocompatible. These coatings may also minimize platelet adhesions and activation, prevent thrombogenicity, and improve performance by decreasing power usage or increasing pump output.

The continuous-flow VADs of Cleveland Heart (Charlotte, NC) have been adapted from the CorAide LVD-4000 Left Ventricular Assist System (Arrow International, Reading, PA), developed at Cleveland Clinic. Detailed descriptions of the pump have previously been published [1,2]. The left and right ventricular assist devices (LVADs and RVADs, respectively) consist of three subassemblies: the volute housing, the rotating assembly (RA), and the stator assembly (**Figure 1**). The RA contains a cylindrical four-pole magnet and spins around the titanium stator housing post, which contains the motor windings. The RA is supported axially by permanent magnets and radially by a thin film of blood, forming a patented blood journal bearing which, except for startup and shutdown, makes no mechanical contact during use. The blood journal bearing materials, dimensions, clearances and geometry are shared by both VADs. The volute housings and RA impeller blades are optimized for the different afterload conditions that occur in the RVAD

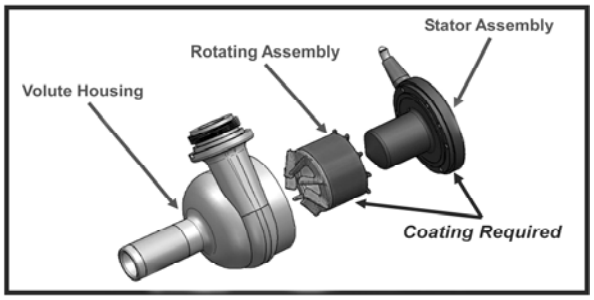


Figure 1. Depiction of an LVAD subassembly. RVADs differ only in RA impeller and in volute housing design.

vs. LVAD.

During European clinical trials of the CorAide LVAD, it was found that the original fluorinated ethylene propylene (FEP) coating on the journal bearings delaminated from its titanium substrate. The authors have since searched for the best alternative bearing material to use on both the LVAD and RVAD; this article focuses on the preliminary *in vitro* evaluation of the BioMedFlex (BMF) thin-film coating (BioMedFlex, Denver, NC).

2. MATERIAL AND METHODS

2.1. BioMedFlex Coating

The BMF coating is deposited on substrates by a proprietary plasma-assisted chemical vapor deposition process in a high-vacuum environment at temperatures below 200℃. Because this coating is created from layers of nano-crystalline diamond and nano-crystalline silicon carbide in a matrix of the noncrystalline forms of both compounds, BMF's unique characteristics are set apart from other thin-film carbon coatings used in biomedical applications [3]. Its high flexural strength, its radiopacity, and its wear resistance are surpassed only by pure diamond thin films. Several other material properties of the BMF coating that make it a good candidate for use in blood journal bearing applications include chemical resistance, dimensional stability, and low coefficient of friction. BMF does not absorb water and does not corrode in saline. These two key qualities prevent dimensional changes due to absorption or corrosion in the high-tolerance clearance and profile of the journal bearing-conditions that occurred with the FEP coating in the CorAide clinical trial. BMF is a clean, reliable, high-technology coating process that is not line-of-sight and does not require parts to be rotated in the chamber. The BMF coating is 2-4 μm thick and aggressively adheres to titanium surfaces, requiring sandblasting to remove it from the surface of coated parts. For reference, a matrix comparing the material properties of BMF vs. titanium nitride, a commonly used material in the field, is provided below as **Table 1**.

2.2. Journal Bearing Starting Conditions

Once running, the bearing allows the RA to revolve levitated from the stator housing post surfaces on a film of blood, deriving lift from its own motion and thereby eliminating any surface contact between moving parts until the pump is shut down (**Figure 2**). However, there is transient (300 ms) sliding and vibrating contact (800-900 g) load between the RA and stator assembly at start-up and again during stopping (**Figure 3**); the only mechanical loading the bearing undergoes is during its start/stop cycles. The worst-case expected loading sustained by the bearing over its lifetime was established based on the fact that, on average, the pump is started and stopped six times during manufacturing and three times during clinical implantation. During its service life with the patient, the pump may be started and stopped (in a worst case) eight times per year due to external component changes, software upgrades, or simple patient connection mistakes. These values translate to about 50 starts and stops in a 5-year period. Multiplying by a safety factor of 2, the number of starts and stops that will approximately simulate the worst-case service life is 100.

Table 1. Comparison matrix of material properties: BioMedFlex vs. Titanium nitride.

Typical values	BioMedFlex	Titanium nitride
Inert	Yes	No
Flexural strength	High triaxial	Low single axis, brittle
Optical transparency	50Å Yes, 2 μm No	No
Mechanical wear	Low	High
Conformal	Yes	No: Line of sight
Hardness (V)	3800	2500
Young's modulus (GPa)	130-180	600
Adhesion (PSI)	> 8500	< 2000
Coefficient of friction	0.10-0.15	0.65
Thickness (μm)	0.1-4	0.25-12
Surface Ra (μm)	0.10	0.20

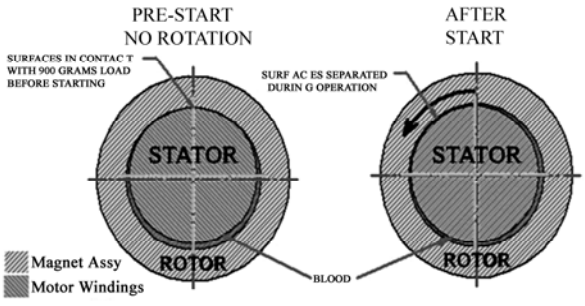


Figure 2. Journal bearing. Left: Inner surfaces of the RA are in contact with stator when pump is not running. Right: After starting, the RA is separated from the stator and rides on a film of blood during operation. Assy, Assembly.

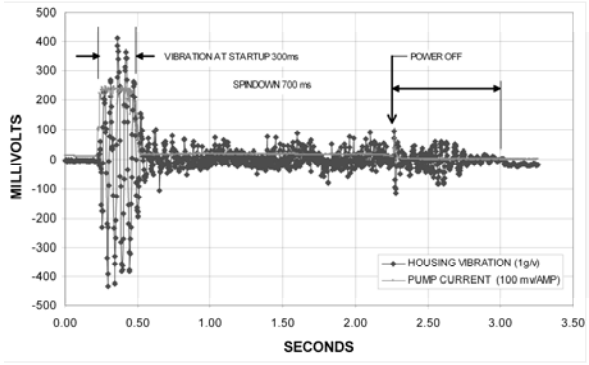


Figure 3. Pump vibration and motor current at start-up and shut-down of a Cleveland Heart VAD.

2.3. Start/Stop Test

To evaluate the durability and reliability of the coating on pump journal bearing mated surfaces, pump assemblies consisting of a BMF-coated stator and RA were exposed to 100 start/stop cycles at the most challenging bearing load conditions: maximum pump speed and wide-open flow conditions in de-ionized water. The pumps were disassembled, and the stator and RA were visually inspected for scratches, delamination, or other damage after 1, 5, 25, and 100 start/stop cycles.

A commercial DLC coating used in other blood pump applications was also applied to another set of bearing surfaces as a control. Twenty different combinations of BMF-coated journal bearing pairs and seven different DLC-coated pairs were exposed to start/stop testing to compare the two coatings for durability and reliability. The high component costs, long lead time for fabrication and coating of parts, and the early and frequent failure of the DLC-coated parts limited the number of DLC test pieces evaluated in this study, as the primary goal of the project was to develop new journal bearing materials. The success criterion was the completion of at least 100 start/stop cycles per bearing pair without damage to the coated surfaces.

3. RESULTS

BMF-coated pumps showed no coating failures (**Table 2**), whereas four of seven commercial DLC bearing pairs developed scratches through the carbon coating (**Table 3**), documenting that BMF can provide a durable coating in our blood journal bearing application. A representative scratch is shown in **Figure 4**; scratches were present on the stator post only, with no damage to the RA on commercial DLC bearing pairs.

4. DISCUSSION

Although titanium alloy has been the material of choice for implantable rotary blood pumps [4], given the bearing loads seen in our pumps at startup and shutdown, it is a very poor choice as a bearing material. It has been a priority to determine the optimum coating material and coating parameters to obtain the bearing properties needed for the Cleveland Heart LVAD and RVAD. BMF's unique thin-film carbon formulation and its layered matrix of diamond and silicon carbide have provided a new biomedical material that combines very high hardness, aggressive substrate adhesion and much higher flexural properties. BMF has proven to be more wear resistant than the commercially available DLC control.

Previous studies of this bearing design have documented

Figure 4. BMF-coated RVAD RA (left) and stator assembly (middle); commercial DLC-coated stator assembly scratches post-test (right).

Table 2. BMF bearing testing results.

BMF Bearing Testing		**BMF Coated RA**			
		RA BMF-01	RA BMF-02	RA BMF-03	RA BMF-04
BMF-Coated Stator	Stator BMF-01	Passed: No damage	Passed: No damage	Passed: No damage	Passed: No damage
	Stator BMF-02	Passed: No damage	Passed: No damage	Passed: No damage	Passed: No damage
	Stator BMF-03	Passed: No damage	Passed: No damage	Passed: No damage	Passed: No damage
	Stator BMF-04	Passed: No damage	Passed: No damage	Passed: No damage	Passed: No damage
	Stator BMF-05	Passed: No damage	Passed: No damage	Passed: No damage	Passed: No damage

BMF, BioMedFlex®; RA, Rotating Assembly.

Table 3. Commercial DLC bearing testing results.

Commercial DLC Bearing Testing		Commercial DLC-Coated Rotating Assembly				
		RA DLC-01	RA DLC-02	RA DLC-03	RA DLC-04	RA DLC-05
Commercial DLC-Coated Stator	Stator DLC-01	Passed: No damage	Passed: No damage	FAILED	FAILED	-
	Stator DLC-02	-	-	-	Passed: No damage	FAILED
	Stator DLC-03	-	-	FAILED	-	-

DLC, diamond-like coating; RA, Rotating Assembly.

that the blood film and journal bearing remain stable once the start-up period is over [5]. Because of this, damage to the bearing can occur only during starting and stopping; it is not a matter of total running time, but the number of start/stops that is critical. According to the test procedure, pump disassembly and component inspection was performed only after the 1st, 5th, 25th and 100th start/stop cycle. Generally, bearing damage occurred early on in the testing (within the first 5 start/stop cycles) for the DLC-coated parts and was progressive and cumulative. Representative bearing wear (**Figure 4**) for the commercial DLC-coated stator is compared to undamaged BMF-coated RA and stator parts post-bearing test; bearing wear was only on the stator post of the DLC-coated parts and not the RA.

With start/stop reliability demonstrated, endurance tests of indefinite duration in 37℃ saline and glycerin blood analog have been started to verify compatibility in a saline environment. After 275 and 70 days, respectively, there has been no bearing wear on two pump assemblies with BMF paired journal bearings undergoing mock circulatory *in vitro* endurance testing in an aggressive, pulsatile environment.

BMF was successfully deposited onto titanium pump surfaces and endured at least 100 start/stop cycles for each bearing pair without damage; over 57% of the commercial DLC-coated pairs failed at the same test parameters. Five of the BMF stators accumulated up to 800 start/stop cycles with no wear. Neither the testing duration nor number of start/stop cycles were increased, as the test was deemed to be adequately aggressive given the "failure" of the commercial DLC pairs. These tests demonstrated that BMF can provide a durable coating in our application, along with a viable solution to bearing FEP coating adhesion problems seen in the initial CorAide clinical trials.

Biocompatibility is being validated with ongoing *in vivo* studies of the performance of LVADs, RVADs, and biventricular assist devices in animals [6].

5. CONCLUSIONS

BMF has many qualities that support its significant advantages as an alternative journal bearing material in both Cleveland Heart pumps: 1) demonstrated bearing reliability, 2) a thin-film coating that still offers a depth of several micrometers, and 3) the capability for batch coating processing. To further test durability, BMF-coated pumps have been placed on mock circulatory endurance test with the pumps running and submerged in a saline/glycerin blood analog fluid at body temperature.

6. ACKNOWLEDGEMENTS

This project was supported by the Global Cardiovascular Innovation Center and the State of Ohio under the Third Frontier project. The authors also acknowledge support from grant BRP 5 R01 HL074896 (to K.F.) from the National Heart, Lung, and Blood Institute of the National Institutes of Health.

REFERENCES

[1] Fukamachi, K., Horvath, D.J., Massiello, A.L., Ootaki, Y., Kamohara, K., Akiyama, M., Zahr, F., Kopcak, M.W., Dessoffy, R., Chen, J.F., Benefit, S. and Golding, L.R. (2005) Development of a small implantable right ventricular assist device. *American Society for Artificial Internal Organs Journal*, **51(6)**, 730-735.

[2] Ootaki, Y., Kamohara, K., Akiyama, M., Zahr, F., Kopcak, M.W., Dessoffy, R., Massiello, A.L., Horvath, D.J., Chen, J.F., Benefit, S., Golding, L.R. and Fukamachi, K. (2005) Initial in vivo evaluation of the DexAide right ventricular assist device. *American Society for Artificial Internal Organs Journal*, **51(6)**, 739-742.

[3] BioMedFlex, LLC, Accessed November 27, 2009. http://www.BioMedFlex.com

[4] Yamazaki, K., Litwak, P., Tagusari, O., Mori, T., Kono, K., Kameneva, M., Watach, M., Gordon, L., Miyagishima, M., Tomioka, J., Umezu, M., Outa, E., Antaki, J.F., Kormos, R.L., Koyanagi, H. and Griffith, B.P. (1998) An implantable centrifugal blood pump with a recirculating purge system (Cool-Seal system). *Artificial Organs*, **22(6)**, 466-474.

[5] Malanoski, S.B., Belawski, H., Horvath, D.J., Smith, W.A. and Golding, L.R. (1998) Stable blood lubricated hydrodynamic journal bearing with magnetic loading. *American Society for Artificial Internal Organs Journal*, **44(5)**, M737-M740.

[6] Takaseya, T., Fumoto, H., Shiose, A., Arakawa, Y., Rao, S., Horvath, D.J., Massiello, A.L., Mielke, N.A., Chen, J.F., Zhou, Q., Dessoffy, R., Kramer, L., Benefit, S., Golding, L.R. and Fukamachi, K. (2009) In vivo biocompatibility evaluation of a new resilient, hard-carbon, thin-film coating for ventricular assist devices. *Artificial Organs* (in Press).

J. Biomedical Science and Engineering, 2010, 3, 529-537 **JBiSE**
doi:10.4236/jbise.2010.35074 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Predicting heat-stressed EEG spectra by self-organising feature map and learning vector quantizers
——SOFM and LVQ based stress prediction

Prabhat Kumar Upadhyay[1], Rakesh Kumar Sinha[2], Bhuwan Mohan Karan[3]

[1]Department of Electrical and Electronics Engineering, Birla Institute of Technology, Ranchi, India;
[2]Department of Biomedical Instrumentation, Birla Institute of Technology, Ranchi, India;
[3]Department of Electrical and Electronics Engineering, Birla Institute of Technology, Ranchi, India.
Email: uprabhat@rediffmail.com; upadhyay@biticrak.ae

Received 9 December 2009; revised 28 December 2009; accepted 30 December 2009.

ABSTRACT

Self-Organising Feature Map (SOFM) along with learning vector quantizers (LVQ) have been designed to identify the alterations in brain electrical potentials due to exposure to high environmental heat in rats. Three groups of rats were considered—acute heat stressed, chronic heat stressed and control groups. After long EEG recordings following heat exposure, EEG data representing three different vigilance states such as slow wave sleep (SWS), rapid eye movement (REM) sleep and AWAKE were visually selected and further subdivided into 2 seconds long epoch. In order to evaluate the performance of artificial neural network (ANN) in recognizing chronic and acute effects of heat stress with respect to the control subjects, unsupervised learning algorithm was applied on EEG data. Mean performance of SOFM with quadratic taper function was found to be better (chronic-92.6%, acute-93.2%) over the other two tapers. The effect of LVQ after the initial SOFM training seems explicit giving rise to considerable improvements in performance in terms of selectivity and sensitivity. The percentage increase in selectivity with uniform taper function is maximum for chronic and its control group (4.01%) and minimum for acute group (1.29%) whereas, with Gaussian it is almost identical (chronic-2.57%, acute-2.03%, control-2.33%). Quadratic taper function gives rise to an increase of 2.41% for chronic, 1.96% for acute and 2.91% for control patterns.

Keywords: ANN; EEG; Heat Stress; SOFM; LVQ

1. INTRODUCTION

The scientific interest of stress in relevance to health and disease began to develop in the 20th century, when Selye started his work on stress with more suited scientific analysis, and now stress has been accepted to be a state, comprised of certain psychophysiological reactions that prepare an organism for action. It is usually described to be an essential component that is enabling the organism to survive in the hostile environment and to make an effort to compensate with the altered situation of the stressful conditions. Stress has been defined as nonspecific responses of the body to any demand. Though in some respect, every demand made on the body is unique and specific, but all stress, however, have one thing in common; they increase the demand for the readjustment for performance of adaptive functions, which reestablish normally. Generally, stress is meant to be acute or at least of a limited duration. The time limited nature of this process renders its accompanying antianabolic, catabolic and immunosuppressive effects temporarily beneficial and of no adverse consequences. Chronically and excessiveness of stress system activation, on the other hand, would lead to the syndromal state that severe chronic disease of any etiology could present with anorexia, loss of weight, depression, and peptic ulcers.

Although, the problems of heat-afflicted illness are receiving increased importance in view of the current estimates of global warming and its impact on biological systems, the etiological factors that lead to heat exhaustion and heat stroke have not been well established. However, the failure of cardiovascular system had been thought to be an important factor. Inadequate acclimatization also appears to be a significant factor predisposing to the onset of heat stroke. Review of literature revealed that the afflictions and damages to the central nervous system (CNS) imposed by high environmental temperature have largely been ignored as the likely cause of heat induced mortality, although it is well known that neurochemical and cellular mechanisms of neural issues are highly temperature sensitive [1]. The conventional long term paper recording of EEG signals

following stress events such as heat stress is not of much diagnostic value. So, computer and digital signal processing tools have been used to quantify the EEG signals for all three sleep-wake stages after acute and chronic heat stress. Since, long-term EEG recordings, in addition to other two channels of electrophysiological signals, EOG (Electrooculogram) and EMG (Electromyogram), reflect the variations in sleep-wake states, so the changes in sleep parameters following acute as well as chronic states were also observed. Further, to reduce the labor involved in the manual sleep staging and to analyze psychophysiological alterations, artificial neural network's architectures were designed. Present study signifies the computerized recognition of different sleep-wake states and the changes occurred in EEG signals due to exposure to high environmental heat.

In the last decade, several works introduced the use of artificial neural network (ANN) as a tool for automated sleep scoring. Most of the system used spectral information of the signal using Fourier transformation [2]. Computerized EEG and other electrophysiological parameters monitoring reduces the problem of huge data handling. Computerization has led to more sophisticated use of EEG, even in effective disorders, where perceptual processes are significantly distorted [3-5]. Fourier transformation, as a conventional method, has been widely used for the standard quantitative analysis of the spectral decomposition and the clinical application of EEG signals [6]. The ANN programs were developed for the analysis of most of the works that rely on spectral analysis and power spectrum method to evaluate electrographic data. In an attempt to classify sleep-wake stages determined, power of FFT or power spectrum band were used for better performance of the system [3,7]. The numerical values of the power of different frequency bands were used as inputs to ANN. As multilayer perceptron neural network (MLPNN) undergoes some limitations, the performance of SOFM has also been tested to solve the problem at hand. In the present study, an effort has been made to exploit the inherent qualities of SOFM. The results obtained from computer simulations have been found to be very encouraging.

In addition to widespread application of artificial neural networks (ANNs) in diagnosis, much developmental work is being undertaken in signal processing and analysis of bioelectric signals. ANNs are widely used as to process raw electroencephalogram (EEG) data or features. Recently, ANNs have been employed successfully for many pattern recognition problems of electroencephalogram (EEG) spectral component [8-11], K-complex detection [12-13], event related potentials (ERPs) detection [14-15], classification of the evoked potentials [16-17] and the recognition of epileptic spike patterns [18-19].

Review of literature reveals that supervised ANNs have been used many times for sleep-wake state identification, but the literature on the methodology for use of unsupervised ANNs are still obscure. The SOFM as proposed by T. Kohonen [20-21] follows unsupervised learning (competitive learning) and consists of a single layer feed-forward network or lattice, the neurons of which become specifically tuned to various input patterns through a self-organizing process. The spatial location of a neuron then corresponds to a particular feature, or group of features, of the input patterns. Output neurons of a topographic map are usually arranged such that each neuron has a set of neighbours. Each node of the output layer is connected with all other nodes of the same layer with inhibitory weights and competitive learning occurs among all the nodes of the output layer. The most highly activated node which has least Euclidean distance becomes the winner for a particular pattern. It has been realized that the trained and calibrated Self-Organising Feature Map (SOFM) alone can not be used as a classifier. Hence, to increase the performance of the pattern classifier, Learning Vector Quantizers (LVQ) is applied in the trained SOFM [22]. In the present study, unsupervised ANN system using both SOFM and LVQ has been designed for classification of heat stressed spectra.

2. MATERIALS AND METHODS

2.1. Subjects and Electrode Implantation

The experiments were carried out with male Charles Foster rats of age 12-14 weeks and weight around 180-200 grams. The rats were individually housed in polypropylene cages (30 cm × 20 cm × 15 cm) with drinking water and food (Hindustan Liver Limited, India) *ad libitum*. All rats were kept in an ambient environment temperature of 23 ± 1°C from birth and the animal room was artificially illuminated with 12 : 12 hours Light : Dark cycle, changed at 7 o'clock and 19 o'clock Indian Standard Time (IST). The technique of electrode implantation for polygraphic sleep recordings have been used as suggested earlier [7].

2.2. Heat Stress Model

In order to produce the effects of heat stress, rats were subjected to the Biological Oxygen Demand (BOD) incubator at preset temperature of 38 ± 1℃ and relative humidity 45-50%, simulated with the environmental conditions of Varanasi (India) in the months of May and June. For chronic heat stress, rats were subjected to the incubator for one hour daily for 21 days of chronic heat exposure from 8.00 a.m. to 9.00 a.m. and electrophysiological signals were recorded on 22nd day whereas, acute heat stressed rats were subjected to the incubator for continuous four hours of heat exposure from 8.00 a.m. to 12.00 p.m. for a single day, just before the recording of electrophysiological signals. Respective control groups of rats were placed in the incubator at room temperature

(23 ± 1℃) and whole procedure was followed exactly similar to that of their stressed groups.

2.3. EEG Data Acquisition

Four hours of continuous recordings of EEG, electro-oculogram (EOG) and electromyogram (EMG) were performed from 12 o'clock to 16 o'clock IST on the recording day for chronic and acute heat stressed rats through the 8 channels Electroencephalograph (EEG - 8, Recorders & Medicare Systems, India). The paper recordings were performed with standard amplifier setup [23] and at the chart speed of 7.5 mm/sec. The digitized data (at sampling frequency of 256 Hz) was collected, stored and processed with the help of data acquisition system (ADLiNK, 8112 HG, NuDAQ, Taiwan) and processing software (Visual Lab.-M, Version 2.0 c, Blue Pearl Laboratory, USA).

2.4. Self-Organising Feature Map and Learning Vector Quantiser

With self-organised learning, no external teacher is required in order to adjust the weights of the ANN, but the choice of the input data set will still reflect the generalizing ability of the ANN (as for the supervised case). It presents a method of identifying similarities (or features) in a vast (unlabelled) training set. An important advantage of self-organising ANNs over their supervised counterparts is that they can be exposed to and make use of vast quantities of input data for training purposes without the need for assigning labels to each input forwarded to the ANN. Let *m*, *x*, and α be the weight vector, input vector and learning rate parameter respectively then with the following rules the weights of the winning neuron as well as its neighbours are updated.

$$w_i(t+1) = w_i(t) - \eta(t)R_i[x(t) - w_i(t)];\ for\ i \in N_c(t)$$

$$w_i(t+1) = w_i(t);\ \ for\ i \notin N_c(t)$$

where, N_c is a topological neighbourhood function centered around the winning neuron *c* and R_i is the neighbourhood taper function. Three different taper function-uniform, Gaussian, and Quadratic have been used for training the network, the mathematical expressions of which are as followed:

Linear: $R_i = 1$, for $i \in N_c$,

Gaussian: $R_i = \exp\dfrac{-d_i^2}{(N_c+1)^2}$ for $i \in N_c$,

$R_i = 1\text{-}1.05\dfrac{-d_i^2}{(N_c+1)^2}$ for $i \in N_c$.

where, d_i is the Euclidean distance between the weight vector w_i and the winning weight vector w_c. Learning rate (η) and neighbourhood size (N_c) are assumed to be time varying. Two types of monotonically decreasing functions are considered here, a piecewise linear decreasing function and an exponential decreasing function. The piecewise linear function is given by:

$$\eta(k) = \eta_0(1-\frac{k}{\mu k_{\max}}) + \eta_{\min}(\frac{k}{\mu k_{\max}});if\quad k \le \mu k_{\max},$$

$$\eta(k) = \eta_{\min};if\quad k > \mu k_{\max},$$

$$N_c(k) = N_0(1-\frac{k}{\mu k_{\max}}) + N_{\min}\frac{k}{\mu k_{\max}};if\quad k \le \mu k_{\max},$$

$$N_c(k) = N_{\min};if\quad k > \mu k_{\max}.$$

The piecewise exponential function is given by:

$$\eta(k) = \eta_0^{(1-\frac{k}{\mu k_{\max}})}\eta_{\min}^{\frac{k}{\mu k_{\max}}};\ if\quad k \le \mu k_{\max},$$

$$\eta(k) = \eta_{\min};\ if\quad k > \mu k_{\max},$$

$$N_c(k) = N_0^{\left(1-\frac{k}{\mu k_{\max}}\right)}N_{\min}^{\frac{k}{\mu k_{\max}}};\ if\quad k \le \mu k_{\max},$$

$$N_c(k) = N_{\min};\ if\quad k > \mu k_{\max}.$$

where, in both cases, η_0 and N_0 give the initial learning rate and neighborhood radius respectively, whereas $\eta_{\min}$ and $N_{\min}$ give the respective minimum values. Here, *k* and μ denote iteration number and fraction of total number of iteration respectively.

To achieve improved classification performance, training and calibration of the SOFM with a supervised learning scheme is followed. Learning vector quantization (LVQ) is one such technique, developed by Kohonen to fine-tune the weights of the trained SOFM in a supervised manner [22]. LVQ is a supervised technique that uses class information to move Voronoi vectors slightly in such a way as to improve the quality of the classifier decision regions. The initial values of w_i before the fine-tuning with LVQ, must be such that w_i represent the overall statistical density function of the input. The SOFM is suited to achieve this. Out of the three versions of the LVQ algorithm, LVQ1 has been applied in the present work, which follows the following weight update rules:

$$w_i(t+1) = w_i(t) + \eta(t)[x(t) - w_i(t)],$$

if $i = c$ and *x* is classified correctly.

$$w_i(t+1) = w_i(t) - \eta(t)[x(t) - w_i(t)]$$

if $i = c$ and *x* is classified correctly.

$$w_i(t+1) = w_i(t) + \eta(t)[x(t) - w_i(t)]\quad \text{for } i \ne c.$$

2.5. Data Preparation and Simulation

The EEG data set prepared for input space of the network was divided into training data set and test data set. The training data set consists of the raw EEG signals of SWS, REM and AWAKE states, each having matrix size of [1, 512]. So, the total size of the input vectors for one

presentation at the input layer of SOFM is [1, 1536], in which all the three sleep states have been equally presented. This constitutes one training epoch. The training set was used to repeatedly train a 2-D lattice of neurons of different sizes and varying SOFM training parameters. The input data was presented to the network 600 times with different values of learning rate parameter (η), topological neighborhood function (N) centered around the winning neuron and neighborhood taper function (h). The weights of the trained SOFM now get ordered in the input space such that they represent the underlying densities of the inputs.

The program searches the position of the weight vectors in relation to the input vectors after different iterations during training SOFM of different sizes such as 4 × 4, 8 × 8, 10 × 10, 14 × 14. The distribution of the model vectors in the n-dimensional space will approximate the probability distribution of the input vectors. The topographic organization of the map will also approximate the metric ordering relations in the input space. Thus similar inputs project near each other onto the map. The map thus forms an "elastic surface" in the input space, which approximates the probability density function of the input samples. Increasing the number of locations (neurons in the lattice) increases the accuracy of the approximation. After training a 2-D lattice of neurons by SOFM, each trained map was calibrated and learning vector quantizer (LVQ) algorithm (a supervised learning technique) was applied in each case for fine tuning the weights of the trained SOFM.

2.6. Body Temperature

Core body temperature was recorded as stress markers for both acute and chronic stress group of rats through the thermistor probe connected to 6-channel telethermometer. The marked probe at 4 cm was inserted to the rectum of the animal and kept static for one minute to record the body temperature. For acute stress group, body temperature was recorded before and after the heat exposure. While for the chronic stress group, the body temperature was recorded on every third day just before putting them into the incubator for chronic heat stress.

3. RESULTS

With the processed EEG data sets as shown in **Figure 1**, the network was simulated number of times and the performance was calculated for some of the simulations employing two decay functions of learning rate and neighborhood size, three neighborhood tapering schemes, and different number of training iterations. Use of linear decay and exponential decay for learning rate and neighborhood size did not show any fixed trend such that conclusion could be drawn precisely. A few graphs showing the mean change in Euclidian distance between weight vectors of SOFM of varying sizes at successive epochs during the training process, have been presented in **Figure 2**. The first subscript denotes decay mode of learning rate and neighborhood distance whereas, letters 1 and 2 stand for linear and exponential decay respectively. The second subscript denotes taper function, where 1, 2 and 3 represent linear, Gaussian, and quadratic functions respectively. Time span of the ordering phase and fine adjustment phase can distinctly be seen in the curves of **Figure 2**. The ordering phase occurs within the first 10%-20% of the training process and is characterized by large changes in the Euclidean distances. The fine adjustment phase is characterized by smaller changes in the distances. Substantial increase in performance was seen in the simulation of 8 × 8 SOFM in acute heat stress. Number of iterations of the input vectors required for simulation was set to 500 and 200 times the number of neurons of the Kohonen layer. The results as shown in the tables (**Tables 1** and **2**) suggest that iterating 200 × size of the SOFM produces almost identical

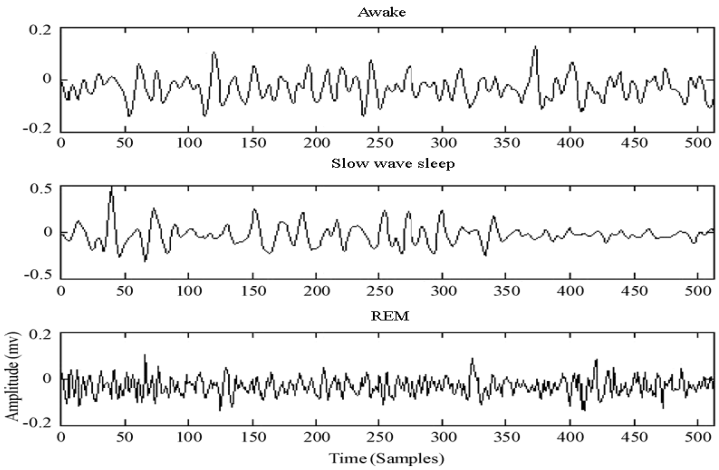


Figure 1. Processed sleep EEG of awake, slow save sleep and REM.

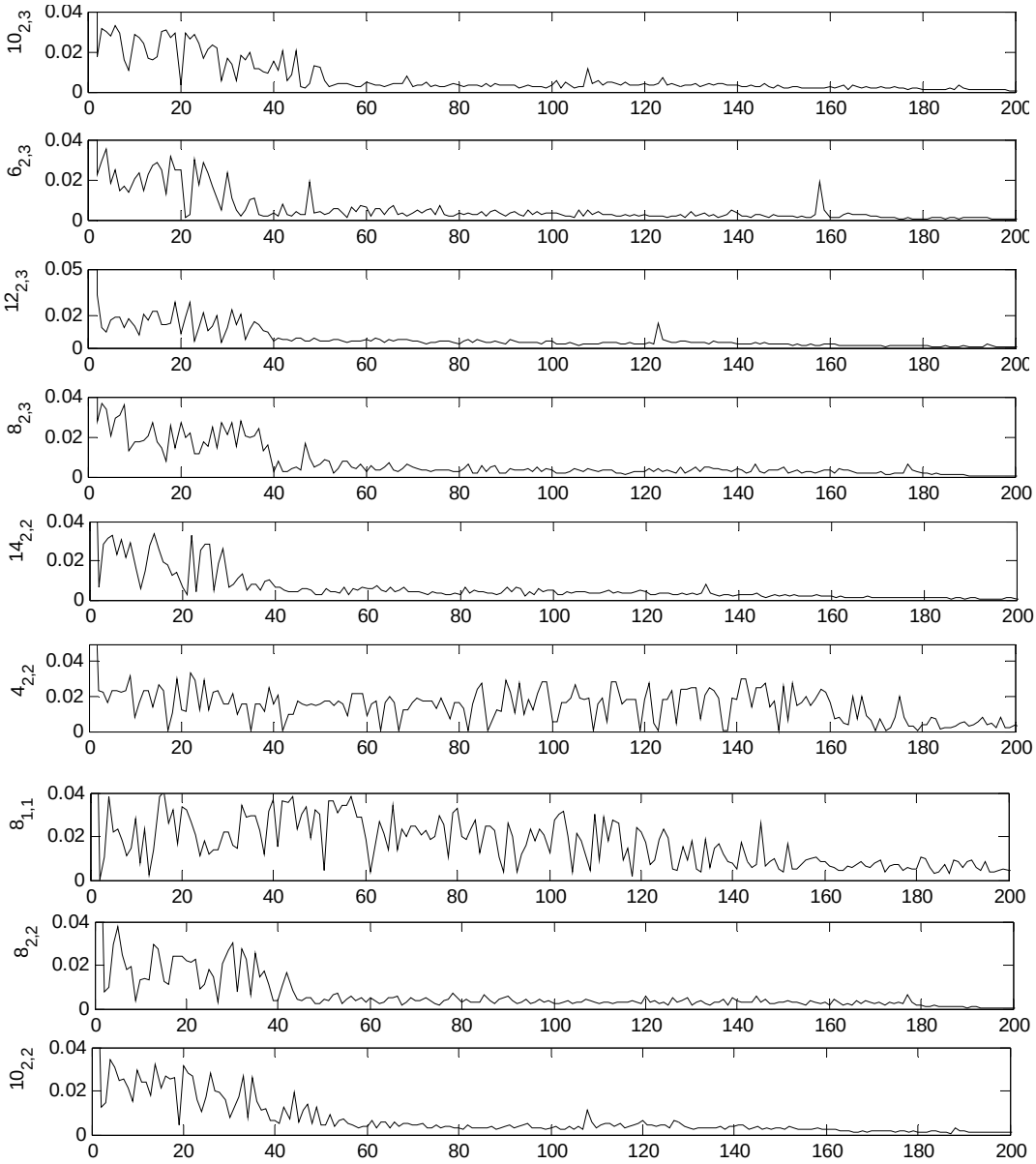


Figure 2. The mean change in euclidean distance between weight vectors during training of SOFM of varying sizes. The first subscript corresponds to decay mode of learning rate and neighbourhood distance whereas the second subscript denotes the type of taper function.

results as iterating 500 × size of SOFM. In some simulations, for both stress conditions, better network performance has been obtained by iterating 200 × size of the SOFM. For 14 × 14 SOFM used in chronic heat stress, it was found to be 94.6% whereas, mean performance of only 91.1% was obtained when the network was iterated by 500 × size of the SOFM. All simulations were performed on two different values of η_{min}, where comparatively, η_{min} = 0.001 offered better performance.

As the SOFM size decreases, the mean changes increase and the SOFM become more settled. 14 × 14 SOFM acquires the optimum set of weight vectors almost after 20% of the training whereas for 8 × 8 SOFM it is 60%. It is also obvious that with the change of network size, learning rate and neighborhood decay function, training time of the network as well as its convergence are being directly affected. Having trained the network, calibration of SOFM is accomplished, in which a class label has been assigned to each neuron ac cording to the maximum voting criteria. Each neuron is also assigned a value, which provides the confidence level with which each neuron represents that class. The boundaries between the three patterns of heat stress-chronic, acute and respective control groups (symbolized by-'ch', 'ac'

Table 1. Mean performance of SOFM in chronic heat stress under varying conditions after simulation.

Parameter	**Mean performance (%) 8 × 8 SOFM**		
Decay	Linear 90.1	Exponential 92.3	-
Iterations	500 × (SOFM Size) 91.4	200 × (SOFM Size) 92.3	-
α_{min}	0.01, 92.4	0.001, 92.8	-
N_c tapper	Uniform 90.73	Gaussian 91.38	Quadratic 92.6
Algorithm	SOFM 90.6	LVQ1 92.8	-
Parameter	**Mean performance (%) 10 × 10 SOFM**		
Decay	Linear 80.2	Exponential 79.4	-
Iterations	500 × (SOFM Size) 83.2	200 × (SOFM Size) 83.0	-
α_{min}	0.01, 80.2	0.001, 79.1	-
N_c tapper	Uniform 80.4	Gaussian 80.3	Quadratic 82.6
Algorithm	SOFM 80.1	LVQ1 92.8	-
Parameter	**Mean performance (%) 14 × 14 SOFM**		
Decay	Linear 93.1	Exponential 94.6	-
Iterations	500 × (SOFM Size) 91.1	200 × (SOFM Size) 94.6	-
α_{min}	0.01, 93.4	0.001, 93.7	-
N_c tapper	Uniform 92.9	Gaussian 94.1	Quadratic 94.3
Algorithm	SOFM 93.6	LVQ1 92.9	-

Table 2. Mean performance of SOFM in acute heat stress under varying conditions after simulation.

Parameter	**Mean performance (%) 8 × 8 SOFM**		
Decay	Linear 84.4	Exponential 88.3	-
Iterations	500 × (SOFM Size) 88.1	200 × (SOFM Size) 87.8	-
α_{min}	0.01, 91.3	0.001, 92.45	-
N_c tapper	Uniform 90.73	Gaussian 92.3	Quadratic 93.2
Algorithm	SOFM 90.8	LVQ1 93.6	-
Parameter	**Mean performance (%) 10 × 10 SOFM**		
Decay	Linear 91.5	Exponential 91.7	-
Iterations	500 × (SOFM Size) 91.0	200 × (SOFM Size) 92.3	-
α_{min}	0.01, 92.4	0.001, 92.2	-
N_c tapper	Uniform 91.1	Gaussian 91.5	Quadratic 90.6
Algorithm	SOFM 91.4	LVQ1 92.8	-
Parameter	**Mean performance (%) 14 × 14 SOFM**		
Decay	Linear 94.1	Exponential 94.8	-
Iterations	500 × (SOFM Size) 93.1	200 × (SOFM Size) 94.2	-
α_{min}	0.01, 93.6	0.001, 92.1	-
N_c tapper	Uniform 91.8	Gaussian 92.1	Quadratic 94.3
Algorithm	SOFM 90.7	LVQ1 93.5	-

and 'co' respectively) as shown in **Figures 3–5**, are well demarcated showing topological ordering of the neurons.

It is important to note that topological ordering is observed, both in the plots of the weight vectors in the input space and in the assigning of labels during calibration. Labels form neat clusters with easily distinguishable boundaries. It is at the boundary that the confidence labels assigned to neurons are at the weakest, due to overlapping of the input distributions. Another important point is that the SOFM assigns equal number of neurons to each class, if the three classes were equally represented in the training set. When a particular class is under represented in the input set, it will also be under represented in the SOFM, because fewer weight vectors will be assigned to represent that class of input data. Results (**Figure 3**) also indicate that to some data set no decision has been taken by the output layer nodes and hence tie between the adjacent nodes seems to appear.

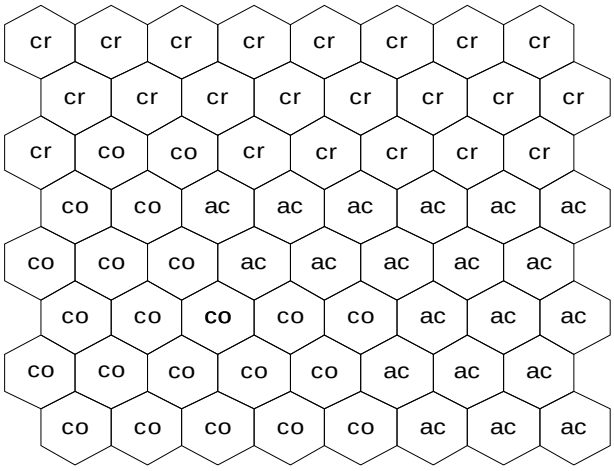


Figure 3. Topological ordering of the output neurons of 8 × 8 SOFM representing three patterns of heat stress-chronic, acute, and control (represented by class labels-cr, ac and co respectively).

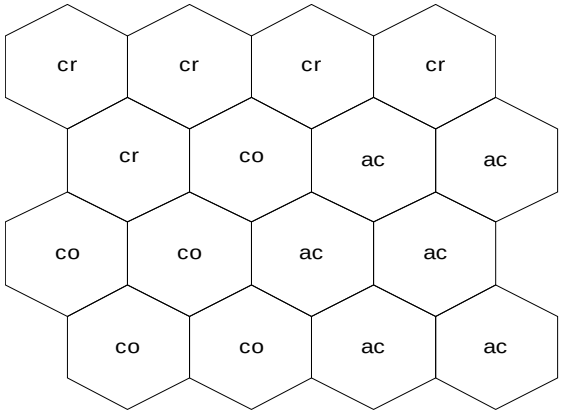


Figure 4. Topological ordering of the output neurons of 4 × 4 SOFM representing three patterns of heat stress-chronic, acute, and control (represented by class labels-cr, ac and co respectively).

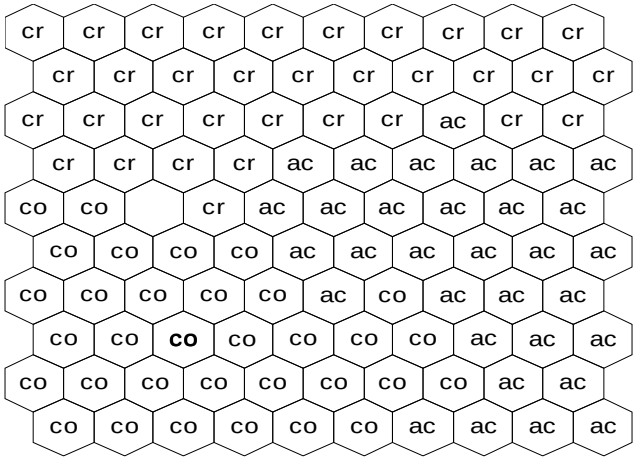


Figure 5. Topological ordering of the output neurons of 10 × 10 SOFM representing three patterns of heat stress-chronic, acute, and control (represented by class labels-cr, ac and co respectively).

Depending on the individual performance of several SOFM networks, 10 × 10 SOFM was selected and trained for recognizing three clusters. In heat stress detection, using 10 × 10 SOFM, average selectivity table (**Table 3**) confirms the presence of three separate patterns-chronic, acute and control groups. For acute and control patterns, Gaussian taper function finds advantage over the other two functions. Again training of SOFM followed by fine-tuning with LVQ increases the performance. This percentage increase in selectivity with uniform taper function is maximum for chronic and its control group (4.01%) and minimum for acute group (1.29%) whereas, with Gaussian it is almost identical (chronic-2.57%, acute-2.03%, control-2.33%). Quadratic taper function gives rise to an increase of 2.41% for chronic, 1.96% for acute and 2.91% for control patterns. The study of average sensitivity of 10 × 10 SOFM (**Table 4**) ascertains the appearance of three distinct patterns in which all three taper functions were found to offer identical results. However, small variation in the average sensitivity can be realized for LVQ.

4. DISCUSSION

The changes in body temperature of the present rat model of heat stress due of acute or chronic heat stress is the sterotype phenomena of heat stress and thus confirms the stressful events along with EEG variations in these experiments. Special features provided by unsupervised networks of labeling large quantities of data and training them in a self-organized manner have been successfully utilized in the present work. Since a number of training and testing parameters play important roles in the clustering problems, the effect of the various parameters on the training of a SOFM has been studied by performing simulation in Matlab. To achieve better classification result, each parameter was varied one by one during simulation and has been tabulated. Performance of different SOFM was finally shown in terms of selectivity and sensitivity. SOFM and LVQ were primarily used for detecting three different patterns-chronic, acute, and control. EEG data of control subjects of chronic group and acute group have been mixed together for training, as variation with respect to stressed subjects were not found to be much. Each simulation was performed on a different input data set. After SOFM training was complete, LVQ was used to calibrate SOFM.

Results obtained from the mean changes in Euclidian distance between weight vectors during training of SOFM of varying size and varying parameters indicated the time of commencement of ordering phase and fine tuning phase. The importance of appropriate selection of taper functions and decay functions has also been well studied by comparative analysis. After the LVQ was applied, all nodes in the output layer were found to be so tuned that three classes of heat stress become apparently separated. Observations suggest that the reason behind the largest change encountered in performance of many simulations, might have been due to the use of different neighborhood taper functions. Overall, performance was found to be better for quadratic taper over the other two tapers. In some simulations, it was witnessed that the increase in performance due to LVQ was negligible, which might have occurred owing to the fact that the input clusters were already reasonably well defined and with a minimal overlap.

In the present work, an attempt has been made to classify stressful conditions by means of changes in EEG signals induced by high environmental temperatures. The review of literature suggests that no work has been reported that classifies heat stressed conditions from normal candidates with help of SOFM and LVQ. The ANN provides reliable information about the stressed and normal artifact-free EEG power spectra. However, in practical applications, EEG artifacts can influence the sensitivity of the network. EEG pattern recognition by ANN and clinical skill, are, however, not mutually exclusive but even reinforce each other, and it is believed that a human clinician must remain a necessary component of computerized diagnostic procedures to ensure a significant, high level of diagnostic validity [24].

Table 3. Average selectivity of 10 × 10 SOFM for heat stress detection.

	Uniform taper		
	Chronic	**Acute**	**Control**
SOFM	88.42	88.91	87.86
LVQ1	92.43	90.2	91.64
	Gaussian taper		
SOFM	87.62	89.50	88.02
LVQ1	90.19	91.53	90.35
	Quadratic taper		
SOFM	87.09	88.87	87.09
LVQ1	89.23	90.83	90

Table 4. Average sensitivity of 10 × 10 SOFM for heat stress detection.

	Uniform taper		
	Chronic	**Acute**	**Control**
SOFM	88.30	87.54	87.81
LVQ1	91.16	90.70	89.80
	Gaussian taper		
SOFM	87.64	88.70	87.64
LVQ1	89.97	90.65	90.23
	Quadratic taper		
SOFM	86.89	88.58	86.98
LVQ1	89.25	91.06	89.95

Studies indicate that, usually only 1-5% of the EEG record is in clinical interest [25]; neural networks can become useful for the on-line classification of EEG waves. Since, exposure to high environmental heat has significant effects on brain signal, SOFM with LVQ can be used efficiently to identify the changes in brain signals, occurred due the stressful events and can also be used further to develop an automated detection system for psychophysiological analysis. Although, in the present study, on-line classification was not carried out, it may be possible with the help of fast computers and specific software. Furthermore, EEG technicians can easily be trained for the manual selection of already detected events whereas; recognition of abnormal patterns in the background of the ongoing EEG requires substantial experience.

REFERENCES

[1] Sinha, R.K. (2007) Study of changes in some pathophysiological stress arkers in different age groups of an animal model of acute and chronic heat stress. *Iranian Biomedical Journal*, **11(2)**, 101-111.

[2] Claude, R., Cristian, G. and Limoge, A. (1998) Review of neural network application in sleep research. *Journal of Neuroscience methods*, **79(2)**, 187-193.

[3] Morstyn, R., Duffy, F.H. and McCarley, R.W. (1983) Altered topography of EEG spectral content in schizophrenia. *Electroencephalography and Clinical Neurophysiology*, **56(4)**, 263-271.

[4] Lin, S.L., Tsai, Y.J. and Liou, C.Y. (1993) Conscious mental tasks and their EEG signal. *Medical and Biological Engineering and Computing*, **31(4)**, 421-426.

[5] Sarbadhikari, S.N., Dey, S. and Ray, A.K. (1996) Chronic exercise alters EEG power spectra in an animal model of depression. *Indian Journal of Physiology and Pharmacology*, **40(1)**, 47-57.

[6] Cesarelli, M., Clemente, F. and Bracale, M. (1990) A flexible FFT algorithm for processing biomedical signals using a PC. *Journal of Biomedical Engineering*, **12(6)**, 527-530.

[7] Sinha, R.K. (2003) Artificial Neural Network detects changes in electro-encephalogram power spectrum of different sleep-wakes in an animal model of heat stress. *Medical and Biological Engineering and Computing*, **41(5)**, 595-600.

[8] Veselis, R.A., Reinsel, R., Sommer, S. and Carlon, G. (1991) Use of neural network analysis to classify electroencephalographic patterns against depth of midazolam sedation in intensive care unit patients. *Journal of Clinical Monitoring*, **7(3)**, 259-267.

[9] Doering, A., Galicki, M., Witte, H. and Krajeca, V. (1995) Structure optimization of neural networks with A*-algorithm application in EEG pattern analysis. *Medical Information for Patients*, **8(1)**, 814-817.

[10] Witte, H., Doering, A., Galicki, M., Dorschel, J., Krajeca, V. and Eiselt, M. (1995) Application of optimized pattern recognition units in EEG analysis: Common optimization of preprocessing and weights of neural networks as well as structure optimization. *Medical Information for Patients*, **8(1)**, 833-837.

[11] Sinha, R.K., Aggarwal, Y. and Das, B.N. (2007) Backpropagation artificial neural network classifier to detect changes in heart sound due to mitral valve regurgitation. *Journal of Medical Systems*, **31(3)**, 205-209.

[12] Jansen, B.H. (1990) Artificial neural nets for K-Complex detection. *IEEE Engineering in Medicine and Biology Magazine*, **9(3)**, 50-52.

[13] Bankman, I.N., Sigillito, V.G., Wise, R.A. and Smith, P.L. (1992) Feature-based detection of the K-complex wave in the human electroencephalogram using neural networks. *IEEE Transactions on Biomedical Engineering*, **39(12)**, 1305-1310.

[14] Wu, F.Y., Slater, J.D., Honing, L.S. and Ramsay, R.E. (1993) A neural network design for event-related potential diagnosis. *Computers in Biology and Medicine*, **23(33)**, 251-264.

[15] Gupta, L., Molfese, D.L. and Tammana, R. (1995) An artificial neural network approach to ERP classification. *Brain and Cognition*, **27(3)**, 311-330.

[16] Bruha, I. and Madhvan, G.P. (1989) Need for a knowledge based subsystem in evoked potential neural net recognition system. *Proceedings of the* 11th *Annual International Conference on IEEE-EMBS*, **6**, 2042-2043.

[17] Holdaway, R.M., White, M.W. and Marmarou, A. (1990) Classification of somatosensory evoked potentials recorded from patients with severe head injuries. *IEEE Engineering in Medicine and Biology Magazine*, **9(3)**, 43-49.

[18] James, C.J., Jones, R.D., Bones, P.J. and Carrol, G.J. (1996) The self-organizing feature map in the detection of epileptiform transients in the EEG. *Proceedings of the* 18th *international Conference of the IEEE Engineering in Medicine and Biology Society*, Amsterdam, 1996, 913-914.

[19] Adeli, H., Zhou, Z. and Dadmehr, N. (2003) Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods*, **123(1)**, 69-87.

[20] Kohonen, T. (1990) The self-organizing map. *Proceedings of the IEEE*, **73**, 1464-1480.

[21] Kohonen, T. (1989) Self organization and associative memory. 3rd Edition, Springer-Verlag, Berlin.

[22] Kohonen, T. (1988) Learning vector quantization. *Neural Networks*, **1**, 303.

[23] Dubois, M., Sato, S., Lees, D.E., Bull, J.M., Smith, R., White, B.G., Moore, H. and Macnamara, T. (1980) Electroencephalographic changes during whole body hyperthermia in humans. *Electroencephalography and Clinical Neurophysiology*, **50(5-6)**, 486-495.

[24] Sarbadhikari, S.N. (1995) A neural network confirms that physical exercise reverses EEG changes in depressed rats. *Medical Engineering & Physics*, **17(8)**, 579-582.

[25] Jandó, G., Siegel, R.M., Horvàth, Z. and Buzsàki, G. (1993) Pattern recognition of the electroencephalogram by artificial neural networks. *Electroencephalography and Clinical Neurophysiology*, **86(2)**, 100-109.

J. Biomedical Science and Engineering, 2010, 3, 538-542 **JBiSE**
doi:10.4236/jbise.2010.35075 Published Online May 2010 (http://www.SciRP.org/journal/jbise/).

Medical equipments high precise detection technology basing on morphology-harris operator

Yang-Yang Mei, Hai-Ming Xie, Lu Han, Shi-Jun Guo

School of Medical Instrument and Food Engineering, USST, Shanghai, China.
Email: myy19860103@126.com

Received 5 January 2010; revised 8 February 2010; accepted 4 March 2010.

ABSTRACT

Medical equipments related to life safety of human, it is important to detect by a high precise method. Image mosaic which based on Harris corner operator is a commonly used method in this area; Harris operator has low calculation burden, it is simple and stable, so it is more effective comparing with other feature point extracted operators. But in this algorithm, corner points can only be detected in a single-scale, there may be losing information of corner points, causing corner point location offset, extracting false corner points because of noise. In order to solve this question, the acquired images should be processed by dilation and erosion operation firstly, then do image mosaic. Results show that image noise can be eliminated effectively after those morphological processes, as well as the false positive noise generated by image glitch. The success rate of image mosaic and detection accuracy can be greatly improved through the Morphology-Harris operator. Measurement of precision instruments which based on this new method will improve the measurement accuracy, and the research in this area will promote the further development of machine vision technology.

Keywords: Image Mosaic; Harris Operator; Feature Point Extraction; High Precise Detection; Dilation and Erosion

1. INTRODUCTION

With the improvement of living standards, people are concerned about the health problems more and more, especially in recent years, surgical accidents have been happening continuously, which make the physician-patient relationship nervous. In this situation, the controversy focused on medical equipment whether it is qualified or not is increasing, so it is important to detect medical equipment in a high precise and steady way. Traditional detection methods are inefficiency, the accuracy is often affected by Man-made factors. Revolutionary changes of precision equipment detection have happened after the rapid development of computer. Through computer vision, we can make non-contact detection for medical instruments which not only improves the detection accuracy but also reduces the impact of Man-made factors. However, a number of common machine vision detection equipments which are restricted by the optical field of view can not capture large view field and wide angle images, and if we pursuit large field of view simply, the image resolution will be reduced, the detection accuracy will be reduced too. In this case, some requirements of detection accuracy of medical devices have not been fulfilled. To resolve this problem, this paper presents a method of using image mosaic which based on Morphology-Harris operator to obtain images of the medical instruments, enhancing image resolution and measuring accuracy.

2. OVERVIEW OF IMAGE MOSAIC

Image mosaic is a technology which is used to solve the problem of small view field that cannot make a big picture. It uses computer to match several images automatically to merge a large panoramic image, at the same time because of the mosaic of one image from two images, the image resolution has been greatly improved, so it has a wide using range in modern life. In foreign country, the panoramic image mosaic technology which based on the movement was brought by Richard Szeliskj in 1996, it is the representative and becomes a classic image mosaic algorithm [1], Richard Szeliskj became the founders of the field of image mosaic, he proposed theory has become a classic theoretical system, even today there are still many people study his theory of splicing. On this basis, Shmuel Peleg, Benny Rousso and others brought an automatic adaptation image mosaic model in 2000, it will select adaptive model of splicing in accordance with the different camera movements [2]. Then M. Brown published an article entitled Recongising Panoramas on the General Assembly of ICCV in 2003 [3], and image mosaic using of SIFT algorithm [4]

was first brought in this article, the algorithm will be automatic completed fully and the results are effective. At the same time, the development of image mosaic is growing rapidly in our country too, algorithms which were proposed by Xiao-Rui Wang, Zu-Xun Zhang *et al.* greatly promoted the development of image mosaic technology.

Image mosaic technology was early used in photographic mapping to integrate a large number of satellite images. In this application, the accuracy of splicing does not ask for much. But, in detection of precision instruments, especially for precision medical equipments, it is the key point; detection accuracy is affected directly by it. The image mosaic model is shown in **Figure 1**. Image registration is the key of image mosaic; it is closely related to the success of it. At present, the method of image registration, based on feature point, is used commonly and widely. It has several advantages such as low calculation burden, high sensitive and little affected by noise [5].

3. HARRIS CORNER OPERATION

In field of image mosaic, points that change obviously are called feature points. Extraction of feature points is the basis of image registration. There is one sort of feature points called corner points, which are changing intensely of gray value or the intersection point of the outline in image, they reflect the important information of image, so we can get the information that is important and ignore information that is secondary by corner points. Harris corner detector is the most commonly used operator in image mosaic.

3.1. Principle of Harris Corner Operator

Harris detector was brought forward by C. Harris and M. J. Stephens in 1988 and based on the feature points of image [6]. While this feature detector is usually called corner detector, it is not just the corners selecting, but rather any image location having large gradients in all directions at a predetermined scale [7]. Signals are processed through auto-correlation function and the result is a matrix *M* of auto-correlation function. Eigenvalue of matrix *M* is first-order curvature of the correlation function, and the point could be viewed as a corner if two curvature values are high. Only the first-order difference of the gray image is used in Harris, so the operation is simple and is widely used in practice. The principle is to calculate the gray value changes during the window moving along any direction; it is supposed that the small window of center pixel point (x, y) has moved u along the direction of X, as well as V along the direction of Y, then analytical expression about gray value changes will be given out by Harris, as **Eq.1**:

$$E(u,v)=\sum_{x,y} w(x,y)\left[I(x+u),(y+v)-I(x,y)\right]^2 \quad (1)$$

In the function above, $w(x,y)$ stands for the window function; $\left[I(x+u),(y+v)-I(x,y)\right]^2$ stands for the gradient of gray image; usually, $w(x,y)$ stands for gauss filter. For a small displacement, **Eq.1** can be replaced by **Eq.2**, as follows:

$$E(u,v)\cong[u,v]M\begin{bmatrix}u\\v\end{bmatrix} \quad (2)$$

M can be expressed as follows:

$$M=\sum w(x,y)\begin{bmatrix}I_x^2 & I_xI_y\\ I_xI_y & I_y^2\end{bmatrix} \quad (3)$$

In **Eq.3**, I_x and I_y are the first derivatives of each image channel. E is very similar to the local auto-correlation function and M describes the shape of the origin of auto-correlation function. λ_1, λ_2 are assumed to be two eigenvalues of M, and they will proportion with the local auto-correlation function, constituting a non-variable rotation to M [8]. Through judgment λ_1 and λ_2 the region which changed slowly could be determined, as well as corner and edge.

The maximum value of local region can be defined by Harris feature point:

$$R=D(M)-kT^2(M) \quad (4)$$

$T(M)$ denotes the matrix trace of M, and $T(M) = \lambda_1+\lambda_2$; $D(M)$ stands for the determinant of matrix M, and $D(M)=\lambda_1\lambda_2$. The value of k is usually recommended of 0.02 - 0.06.

With the theory of Harris corner detection, several steps could be concluded as follows:

1) Filter every pixel of the image by horizontal and vertical difference operator, through this, I_x and I_y can be obtained [9].

2) Deal with the four elements of matrix M by gauss filter, a new M will be obtained. Gaussian function is showed below:

$$Gauss=\exp\left(-\frac{(x^2+y^2)}{2\partial^2}\right) \quad (5)$$

3) Calculate the interesting value of each pixel by

Figure 1. Image mosaic model.

function **Eq.4**.

4) Set the threshold and do non-maximum suppression by it.

The corner points, detected by Harris detection are shown in **Figure 2**.

3.2. Disadvantage of Harris Operator and Ameliorate Method

Harris operator is a widely used feature point extraction operator and has several advantages, but it has some disadvantages too. For example, in this algorithm, corner points can only be detected in a single-scale, there may be losing information of corner points, causing corner point location offset, extracting false corner points because of noise, inaccurate localization to T and diagonal T and so on.

For tiny medical devices, measurement accuracy is extremely demanding, but because of light, as well as the lens distortion, the image edges will produce some distortion. If the device edge which near the image edge has a glitch, the glitch will have a shape distortion in the two images for mosaic, causing corner point location offset, thus affecting the accuracy of image mosaic. For which we first using morphological dilation and erosion operation to preprocess the image, eliminating the false positive noise generated by image glitch, thus eliminating the inaccurate positioning of the corner points caused by noise, as well as the effect of the corner point location offset, greatly improved the splicing accuracy and detection accuracy.

4. EXPERIMENT RESULTS

4.1. Dilation and Erosion Operation

According to the illumination of 3.2, first, we process the image through dilation and erosion operation. The result is shown in **Figure 3**, the left part is the former image, we can see the glitch clearly in the small circle marked on the image, and right half part is image after processing, the glitch is completely disposed.

4.2. Image Mosaic

In this part, we use tablet punch to demonstrate the mosaic course. The size of tablet punch determines the weight and density of tablets, the precise of punch is important. First of all, two images of tablet punch using a large field camera can be acquired, then do dilation and erosion operation, the result can be shown in **Figure 4**. Be attention that the two images must have some overlapped parts, and this is the basis of image mosaic. Then we can extract the feature points of each image by Harris operator; do image registration by the normalized cross-correlation and image fusion. The result is shown in **Figure 5**.

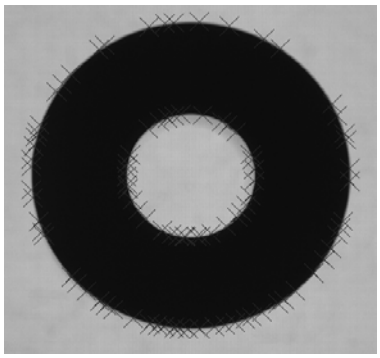


Figure 2. Corner points detected by Harris.

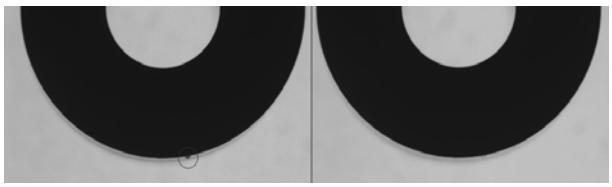


Figure 3. Dilation and erosion processing.

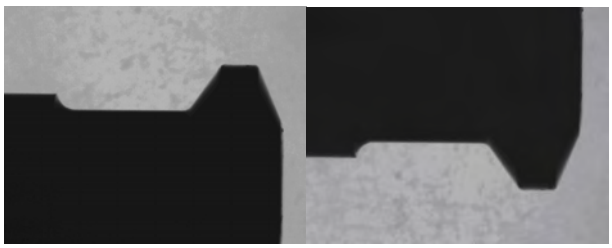


Figure 4. Images waiting for mosaic.

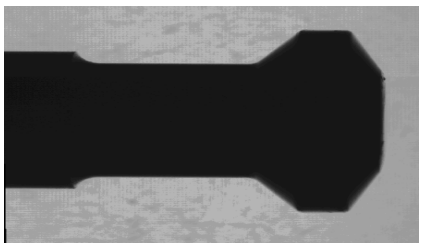


Figure 5. Image after mosaic.

4.3. Test and Contrast

After image dilation and erosion processing, mosaic images through registration and fusion, then the width of tablet punch is measured. Here, we test the width in three situations: single image, mosaic image without preprocess, mosaic image by Morphology-Harris Operation. The result images are shown in **Figure 6**, the left is single image, middle is non-process image result, the right one is mosaic image through Morphology-Harris operator. We found the accuracy of the width in the right

one is the highest, the left one is the lowest. Size is measured by a method which based on sub-pixel measurement, the pixel value is not an integer number in this way and the measurement result has a higher precise; the value at top left corner of image is every pixel value which is measured after the camera is calibrated using the 30 mm standard block .

In order to compare the result conveniently, we can see the **Table 1** below. In this table, five values can be measured through moving the high of camera in every situation. Absolute difference between single image size and standard size is 0.567 mm; the value between non-preprocess mosaic image and standard size is 0.430 mm; and the value between Mosaic image of morphological processes and standard size is 0.080 mm. Comparing the results, it is obvious to see that the measurement accuracy of mosaic image after morphological processes is much higher than the single image and Mosaic image of non-preprocess. Through this comparison, the important of dilation and erosion processing of image mosaic in the precision instrument measurement is proved, and the detection accuracy will be improved through image mosaic.

5. CONCLUSIONS

Medical equipments related to life safety of human, accurate measurement is particularly important. In this paper, we use the morphology-Harris operator preprocessing images, eliminate the noise caused by lens distortion and light uneven, extract feature point accurately, position accuracy is higher than the untreated image. Through dilation and erosion processes, the accuracy of the mosaic image is improved, simultaneity the cost of additional hardware needed is reduced too. The experiment results has proved that, accuracy of medical device testing which based on morphology-Harris operator of image mosaic fully meets the requirements. This is an efficient, high-precision medical device detection technology, suitable for promotion.

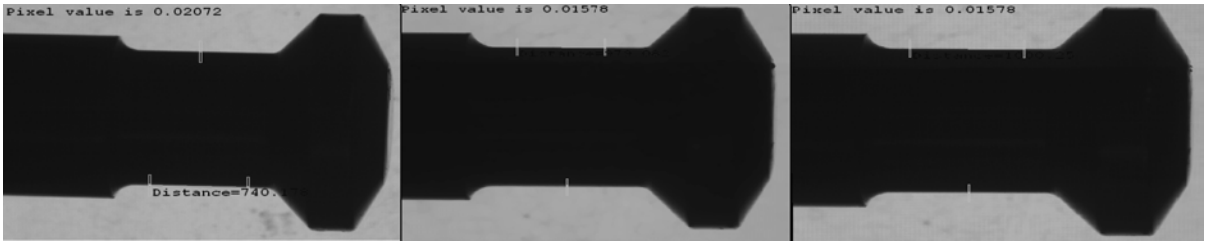


Figure 6. Width of tablet punch image.

Table 1. Result compared.

	pixel value	actual value of every pixel	actual value of width	average value of width
Single image	740.178	0.02072 mm	15.336 mm	15.313 mm
	765.842	0.02003 mm	15.340 mm	
	690.223	0.02222 mm	15.337 mm	
	801.132	0.01899 mm	15.210 mm	
	787.045	0.01949 mm	15.341 mm	
Mosaic image of non-preprocess	979.082	0.01578 mm	15.450 mm	15.450 mm
	901.110	0.01715 mm	15.453 mm	
	1008.037	0.01532 mm	15.437 mm	
	880.029	0.01756 mm	15.450 mm	
	947.501	0.01632 mm	15.461 mm	
Mosaic image of morphological processes	1000.250	0.01578 mm	15.784 mm	15.800 mm
	921.283	0.01715 mm	15.800 mm	
	1031.527	0.01532 mm	15.803 mm	
	899.544	0.01756 mm	15.796 mm	
	969.301	0.01632 mm	15.819 mm	
Standard size	-----	-----	15.880 mm	15.880 mm

6. ACKNOWLEDGEMENTS

Here, I wish to thank my tutor Professor Haiming Xie, under his careful guidance, I completed the paper and related experiments, thanks my lab colleagues Lu Han, Shijun Guo, and under their help I finished my dissertation fast and accurately.

REFERENCES

[1] Szeliski, R. (1996) Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*, **16(2)**, 22-30.

[2] Peleg, S. and Rousso, B. (2000) Mosaicing on adaptive manifolds. *IEEE Transactions on PAMI*, **22(10)**, 1144-1154.

[3] Brown, M. and Lowe, D.G. (2003) Recognising panoramas. *Proceedings of IEEE International Conference on Computer Vision*, **2(1)**, 1218-1225.

[4] Mikolajczyk, K. and Schmid, C. (2004) Scale & Affine invariant interest point detectors. *International Journal of Computer Vision*, **60(1)**, 63-86.

[5] Xie, D.H. and Zhan, Z.Q. (2003) Improving Harris corner detector. *Journal of Geomatics*, **28(2)**, 22-23.

[6] Harris, C. and Stephens, M.J. (1988) A combined corner and edge detector. *Proceedings Fourth Alvey Vision Conference*, Manchester, 1988, 147-151.

[7] Qian, W., Fu, Z.Z. and Liu, Q.Q. (2008) Voting-strategy-based approach to image registration. *Opto-Electronic Engineering*, **35(10)**, 86-91.

[8] Zhao, W.J., Gong, S.R. and Liu, C.P. (2008) Adaptive Harris corner detection algorithm. *Computer Engineering*, **34(10)**, 212-214.

[9] Zhang, J.F. and Xing, X. (2008) Static image registration research based on Harris corner. *Science and technology information*, **20**, 9-10.

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, 2010, 3.
Published Online May 2010 in SciRes (http://www.scirp.org/journal/jbise).

TABLE OF CONTENTS

Volume 3, Number 5, May 2010

