Targeting brain gliomas energy metabolism for classification purposes

M.G. Kounelakis, M.E. Zervakis  
Dept. of Electronic & Computers Engineering  
Technical University of Crete  
Chania, Crete, Greece  
mkoune@gmail.com

Dept. of Electrical & Computer Engineering  
University of Akron,  
Akron, USA

G.J. Postma, L.M.C. Buydens  
Dept. of Analytical Chemistry  
Radboud University  
Nijmegen, the Netherlands

X. Kotsiakis  
Dept. of Neurosurgery  
General Hospital of Chania  
Chania, Crete, Greece

Abstract—The aim of this study is to reveal the discriminative potential of energy related metabolites in brain gliomas classification. The proposed analysis considers two aspects, the statistical and biological verification of metabolites’ effects. In particular, Magnetic Resonance Spectroscopic Imaging (MRSI) is first employed for the statistical evaluation of metabolites. Five of the identified significant metabolites, namely glucose, pyruvate, lactate, alanine and lipids, are involved in the energy production process, necessary for the survival of the cell. In the second stage of analysis, we consider specific metabolic pathways like glycolysis, lactate fermentation, citric acid cycle and lipogenesis in order to evaluate the role of these metabolites in the energy production process.

The results of the proposed double process have shown that these five metabolites are capable to discriminate the three types of gliomas provided. Low grade glioma (GR2) can be discriminated from intermediate grade ones (GR3) with accuracy of 86%. Intermediate grade gliomas versus high grade ones (GR4) with accuracy of 98% and finally low grade versus high grade with accuracy of 100%. Accuracy was extensively evaluated with the area under the ROC (AUROC) global metric.

Keywords—Energy metabolism; brain gliomas; magnetic resonance spectroscopy

I. INTRODUCTION

Energy production is a vital metabolic function of a human cell. Under normal conditions the human brain cells acquire their energy through predefined metabolic pathway called glycolysis. Glycolysis (a sugar splitting process), as shown in Fig. 1, involves a series of biochemical reactions in which glucose is broken down to pyruvate with the release of usable energy in the form of ATP (adenosine triphosphate) molecules. Under aerobic conditions (aerobic glycolysis), the brain derives almost all of its energy from the aerobic oxidation of glucose [1-3]. Pyruvate metabolite which is the end product of this process enters the mitochondria of neurons and glia and the citric acid cycle, also called TCA cycle begins were 36 ATP molecules of energy are finally produced. Only about 13% of glycolytic pyruvate is converted to lactate under normal conditions [3].

In contrast to normal brain cells the malignant rapidly-growing glioma cells present very high glycolytic rates [4]. There are two common reasons for this fact. The classical explanation is that there is poor blood supply to tumors causing local depletion of oxygen. The other explanation stems from the well known hypothesis of Otto Warburg, who claimed that most cancer cells predominantly produce energy by glycolysis followed by lactic acid (lactate) fermentation in the cytosol, rather than by oxidation of pyruvate in mitochondria like most normal cells [5]. This occurs even if oxygen is plentiful. Warburg postulated that this change in metabolism is the fundamental cause of cancer [6], a claim now known as the Warburg effect. This effect may simply be a consequence of damage to the mitochondria in cancer cells, or an adaptation to low-oxygen environments within tumors, or a result of cancer genes shutting down the mitochondria because they are involved in the cell's apoptosis program, which would otherwise kill cancerous cells.

The Warburg effect may also be associated with cell proliferation. Since glycolysis provides most of the building blocks required for cell proliferation, it has been proposed that cancer cells (and normal proliferating cells) may need to activate glycolysis despite the presence of oxygen in order to proliferate [7].

When oxygen is depleted, as for instance in hypoxic-necrotic tumorous tissues of gliomas, the dominant glycolytic product in many tissues is lactate and the process is known as anaerobic glycolysis.

Thus, lactate metabolite is a sensitive indicator of anaerobic glycolysis and reduced cellular oxygenation in living tissues. Along similar lines, alanine, in conjunction with lactate, increases in tissues during hypoxia; made by transamination of pyruvate to prevent further increases in lactate [8].
Furthermore, lipids and fatty acid syntheses are increased in gliomas [9].

Therefore, reliable estimates of the levels of glucose, pyruvate, lactate, alanine and lipid substrates are of special interest for the clinical management of brain gliomas patients.

The main goal of this study is to reveal the significance of these metabolites in the discrimination of different glioma types, thus providing a way to further improve the diagnostic accuracy in today's clinical practices. This is done through a double process that evaluates the significance of metabolites in both a statistical and a biological perspective. The statistical evaluation involves the derivation of significant metabolites from a dataset of MRSI exams. Besides, the biological framework of evaluation involves verification of the results through existing studies, as to enhance their validity and importance.

II. MATERIALS AND METHODS

The metabolic pathways that compose the energy production process are included in the well known Human Metabolome Database (HMDB) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) available in the internet. A thorough examination of the biochemical properties of glucose, pyruvate, lactate, alanine and lipids showed that these metabolites are vital for the energy production, as described in the Introduction section above.

![Figure 1. The energy production process in a human cell](image)

TABLE I. ANALYSIS OF THE DATASET

<table>
<thead>
<tr>
<th>Tissue type</th>
<th>No of subjects</th>
<th>No of voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR2</td>
<td>10</td>
<td>176</td>
</tr>
<tr>
<td>GR3</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td>GR4</td>
<td>7</td>
<td>70</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>303</td>
</tr>
</tbody>
</table>

GR2: glioma grade 2, GR3: glioma grade 3 and GR4: glioma grade 4

In order then to identify the significance of these metabolites in the discrimination of different types of gliomas, we estimated their peak area levels in a real dataset which consists of short echo magnetic resonance spectroscopy imaging (MRSI) data from 21 glioma patients, as presented in Table I.

The two-dimensional MRSI data was collected by the Radboud University and contains 303 pre-processed proton-MRSI (1H-MRSI) volume elements (voxels) corresponding to 303 spectra. Each patient case had passed strict quality control and validation procedures, including consensus histopathologic determination.

The peak areas, Fig. 2, were obtained by peak integration [10, 11] of glucose (at 3.44 ppm), pyruvate (at 2.37 ppm), alanine (at 1.48 ppm), lactate (at 1.33 ppm) and lipids (the sum of the resonances at 0.90 and 1.30 ppm).

The discriminative potential of these five metabolites was evaluated through the 3 different binary classification schemes (GR2 vs GR3, GR3 vs GR4 and GR2 vs GR4), using Support Vector Machines (SVMs) classifier [12] with a Radial Basis Function (RBF) kernel. Kernel-based methods can perform well in processing when high dimensional and heterogeneous data are under scrutiny. The parameters of RBF kernels (C, γ) were tuned in each binary classification in such a way that the smallest possible number of support vectors would be retained for training purposes, as to avoid overtraining.

The SVM solution is based only on those data samples that are at the margin of the decision boundary, called support vectors (lie on the dashed lines in Fig. 3). In essence, SVM attempts to find the best separating hyperplane to distinguish between the two classes of interest, positive (+1) and negative (−1). This is done by maximizing the distance $\frac{2}{\|w\|}$ between the two parallel lines $(w \times x) + b = 1$ and $(w \times x) + b = -1$, which form the margin of separation of the two classes as shown in Fig. 3.

The separating hyperplane passes through the middle of this margin with equation $(w \times x) + b = 0$. The decision function, (1) then, is a function of the form:

$$f(x) = \text{sgn}(w \cdot x + b)$$

(1)

where $w$ represents the direction vector of the hyperplane. The sign of the value returned by (1) indicates the predicted class associated with example $x$, while $|f(x)|$ indicates the confidence level of the resulting decision.
The evaluation of classifier performance was assessed through a 10-fold Cross Validation (CV) method and the classification accuracy was measured in terms of AUROC values. In 10-fold CV, 100 stratified random splits of the initial dataset containing 303 voxels were created. For each binary classification scheme the train sets contained 90% of the voxels of each of the two classes and the remaining 10% of these two classes were contained in the test sets.

The statistical significance of the selected metabolites was also measured through Independent-Samples t-test. The values of the means of each one of the five metabolites were estimated in each class and the p-values were computed. A mean’s difference was considered significant (S) when its p-value was less than 0.05, highly significant (HS) for less that 0.001 and non-significant (NS) when greater than 0.05.

The software package used for the construction of SVM classifier was Matlab ver. 7.9 and for the statistical analyses the SPSS ver. 17.

### III. RESULTS

After the application of the SVM classifier with an RBF kernel, the results obtained are shown in Table II. In this Table the three binary schemes selected for classification and the corresponding number of voxels per scheme, are shown.

Following the strategy described in the Materials and Methods section the statistical t-test has been applied to estimate possible significant differences between the mean values of the five metabolites in each type of glioma dataset. The results of this procedure are presented in Table III.

A graphical representation of Table III is shown in Fig. 4. The colored bars clearly present the mean values differences of glucose, pyruvate, alanine, lactate and lipids in each type of gliomas.

**TABLE II. BINARY CLASSIFICATION SCHEMES**

<table>
<thead>
<tr>
<th>Binary schemes</th>
<th>Voxels / scheme</th>
<th>AUROC measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR2 vs GR3</td>
<td>176 vs 57</td>
<td>86 %</td>
</tr>
<tr>
<td>GR3 vs GR4</td>
<td>57 vs 70</td>
<td>98 %</td>
</tr>
<tr>
<td>GR2 vs GR4</td>
<td>176 vs 70</td>
<td>100 %</td>
</tr>
</tbody>
</table>

**TABLE III. INDEPENDENT SAMPLES T-TEST RESULTS**

<table>
<thead>
<tr>
<th>Metabolites</th>
<th>GR2 vs GR3</th>
<th>GR3 vs GR4</th>
<th>GR2 vs GR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>NS</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>Pyruvate</td>
<td>S</td>
<td>NS</td>
<td>S</td>
</tr>
<tr>
<td>Lactate</td>
<td>HS</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>Alanine</td>
<td>HS</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>Lipids</td>
<td>HS</td>
<td>HS</td>
<td>HS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>&gt; 0.05</td>
</tr>
</tbody>
</table>

Figure 4. The mean peak-area values of the 5 metabolites measured in the 3 classes provided.
Examining the classification results in Table II we can clearly observe that the highest AUROC accuracy values of 98% and 100% are derived from GR3 vs GR4 and GR2 vs GR4 binary schemes, respectively. On the other hand, in GR2 vs GR3 scheme a smaller AUROC value of 86% is obtained. The latter is possibly due to the fact that class GR2 in the initial dataset depicts great heterogeneity between the patients. More specifically, 90 out of 176 voxels of this class are diffuse astrocytoma while the rest 86 voxels are a mixture of low grade gliomas (oligodendrogliomas). This has a negative impact to the classifier learning process.

The high accuracy values of the second and third binary schemes (GR3 vs GR4 and GR2 vs GR4) can be also justified from the statistical analysis results presented in Table III and Fig. 4. As we can observe, the glucose metabolite plays a very significant role in the discrimination of intermediate from high and low from high grade gliomas. This does not happen, however, in the first binary scheme where this metabolite has not a significant contribution in the discrimination of low from intermediate grade gliomas since its mean values do not differ among the two classes.

The significance of these metabolites is further verified in a second stage considering their biological effects as reported in other relevant studies. Glucose forms the initial metabolic substrate that triggers the glycolysis pathway, as explained in the Introduction. In low and intermediate grade tumors, where the tumorous cells are proliferating with a slower rate, the energy demands remain low. This implies that the glycolysis process in not an urgent process at this level. This is not the case in high grade tumors, as in GR4. At that stage of the disease, the energy needs exceed the predefined limits that cause extremely high glycolytic rates i.e. high glucose consumption [4].

Now, when oxygen is present in a tumorous area, the high glucose consumption forces the generation of pyruvate in the mitochondria and so there is no need for lactate production. But when highly hypoxic-necrotic areas exist, like in GR4, glucose is converted to lactate (anaerobic glycolysis) which is an energy rich molecule. In addition, given some oxygen, lactate can also be converted to pyruvate and so enter the mitochondria to generate a bucket load of ATP but also lipids, as shown in Fig. 1. So lactate with oxygen is a potent combination for ATP generation. Furthermore, alanine can be produced during hypoxia by transamination to pyruvate from another amino acid [8].

As observed in Fig. 4 pyruvate levels vary in GR2 vs GR3 and GR2 vs GR4. This is not the case though when cancer cells rapidly proliferate upgrading thus from GR3 to GR4. This is expected since pyruvate is immediately converted to either lactate after the glycolytic process (Warburg effect) through the fermentation process or lipids through the lipogenesis process. Therefore, lactate lipids and alanine metabolic levels increase rapidly as the malignancy increases. The highly statistical differences in their mean values also prove this tendency. Another important observation in Fig. 4 is the rapid increase of both lactate and lipids in GR4 compared to their levels in GR3 and GR2. This fact underlines the lack of oxygen in this tumorous tissue [13-14].

Comparing to similar studies [15-16] we can observe that high classification rates were achieved in this study too, but in contrast to these studies the biological significance of the metabolic markers was further evaluated through their energy metabolic pathway.

Based on the above-mentioned observations and the classification accuracies obtained using these five metabolites, we conclude with high confidence to the fact that these metabolites possess a significant role in the discrimination of glioma types and grades. In essence, by examining their biochemical profile we can determine the energy flux within the cancerous cells, which can be supportive in today’s clinical practice.

IV. CONCLUSIONS

The study of bioenergetics in gliomas is a very promising field for clinical and biological management of complex brain tumors since the energy production and consumption processes underlie the cancer cells behavior. Glycolysis pathways (both aerobic and anaerobic), lactate fermentation, citric acid cycle and lipogenesis are crucial tasks for the proper function of any cell, since they are responsible for the production and consumption of the energy resources. This fact motivated us to examine these pathways in order to reveal the biochemical profile of several metabolites and furthermore test their discriminative power in a real database of glioma patients.

On the other hand, $^1$H or proton Magnetic Resonance Spectroscopy Imaging (1H-MRSI) can provide information on the metabolic profile of tissue and is increasingly being used as a noninvasive method for classifying brain lesions. Earlier MRS studies have shown clear differences between the $^1$H spectra of brain tumors and normal brain tissue [17].

Taking advantage of the biochemical activities of glucose, pyruvate, lactate, alanine and mobile lipids and also their peak area levels derived from the analysis of $^1$H-MRSI spectra, we managed to show that these five metabolites are highly significant for grade discrimination, as well as for the classification of new glioma patients.

ACKNOWLEDGMENT

The authors acknowledge Professor A. Heerschap from University Medical Center of Nijmegen, department of Radiology for his contribution in this study.

REFERENCES