Differentiation of Glioblastomas and Metastases using ^{1}H-MRS spectral data

Félix Fernando González-Navarro[1], **Lluís A. Belanche-Muñoz**[2]
[1]Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
fgonzalez@lsi.upc.edu
[2]Dept. de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain
belanche@lsi.upc.edu

Abstract— *Hydrogen-1 magnetic resonance spectroscopy (^{1}H-MRS) allows noninvasive* in vivo *quantification of metabolite concentrations in brain tissue. In this work two of the most aggressive brain tumors are studied with the purpose of differentiating them. The challenging aspect in this task resides in that their radiological appearance is often similar, despite the fact that treatment of patients suffering these conditions is quite different. Efforts to differentiate between these two profiles are getting increasing attention, mainly because the consequences of performing an incorrect diagnosis. Due to the high dimensionality, initiatives oriented to reduce the description complexity become important. In this work we present a feature selection algorithm that generates relevant subsets of spectral frequencies. Experimental results deliver models that are both simple in terms of numbers of frequencies and show good generalization capabilities.*

Keywords: Feature Selection; Proton Magnetic Resonance Spectroscopy; Classification.

1. Introduction

Hydrogen-1 (^1H) magnetic resonance (MR) spectroscopy allows noninvasive *in vivo* quantification of metabolite concentrations in brain tissue. Nowadays, ^{1}H-MRS has proven its value as a powerful tool in the clinical assessment of several pathologies –as epilepsy, multiple sclerosis and several types of cancer, see e.g. [1], [2]. Its application in brain tumor oncologic diagnosis carries tremendous benefits to patients, relieving them from complicated surgical procedures and their collateral difficulties.

The use of systematic approaches based on ^{1}H-MRS data for the diagnosis and grading of adult brain tumors is subject of an extensive scientific research. One of these growing approaches takes as a backbone well-established machine learning techniques acting as model building methodlogies to discern and predict several classes of brain tumors. Previous work typically focuses on discriminating *normal* tissue from *high grade* malignant tumors, *low grade* malignant tumors, *meningiomas*, as well as other rare tumors or super-classes thereof [3], [4], [5]. In this work we go a step beyond by concentrating on the specific internal separation of the superclass formed by the high grade malignant tumors, differentiating from *Glioblastomas* and *Metastases*. Although their radiological appearance is often similar, the clinical management of patients suffering any of these conditions is quite different [6]. This circumstance may lead to a surgical biopsy for a definitive diagnosis, an undesirable situation in general. Therefore, the use of non-invasive techniques and the discriminative power of the derived information is ever encouraged in this setting.

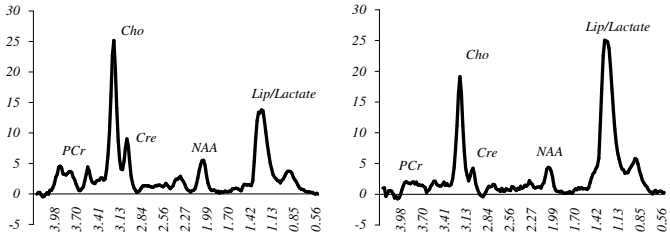

Fig. 1: Mean spectra for Glioblastomas (left) and Metastases (right).

2. High-grade gliomas and Metastases

Glioblastomas (GLI) are the most common and most aggressive type of primary brain tumor in humans and account for the 52% of all parenchymal brain tumor cases and 20% of all intracranial tumors. It affects mainly male patients aged 45 to 55, with a survival expectation of around 40 weeks after diagnosis and medical treatment [7], [8].

Brain Metastases (MET), which amounts about 35% of all cancer patients, are a major cause of death from cancer. Properly speaking, METs are not originated in the Central Nervous System (CNS). Their source is mostly based in lung cancer, breast cancer, melanoma, renal cancer and colon cancer. Mutated cells are spread out trough direct invasion to normal brain tissue or by alternate routes like the lymphatic and blood vessels circulating through the bloodstream. Patients with brain metastases at best reach a mean survival expectation, after whole brain radiotherapy, of a little more than seven months [9], [7].

The radiological appearance of GLIs and METs is quite similar. Figure 1 depicts Mean Long-Echo Time spectral plots of both pathologies put side by side. Their similarity in metabolic levels, with the most prominent peaks in the Choline (Cho) zone, the Lipids-lactate region and Phosphocreatine (PCr), can be clearly seen. METs commonly show moderate to marked reduction of N-Acetyl aspartate (NAA, a major brain metabolite), a decreased Cr signal, and elevated Cho. Also, some metastases may also contain considerable concentrations of lipids [10].

As mentioned, efforts to overcome the difficulty of differentiating between these two profiles are scarce but are acquiring increasing attention. For instance, the authors in [6] studied short echo time spectral data from 23 glioblastomas and 24 metastases. An *ad-hoc* tool to analyze metabolites concentrations was used; they concluded that Lipid and Macromolecule signals can provide a significant discrimination. Further, [11] studied 43 patients having one of these two conditions. The method used was based in MR imaging and a kind of special signal processing called Dynamic Susceptibility Contrast Perfusion MR imaging; the study concludes positively about differentiating the two classes of tumor.

In the same line of research (MR imaging), several image processing techniques have been used to asses possible differences, which are compared with non-parametric statistical tests [12]. A logistic regression analysis with forward stepwise selection yields an AUC of 0.98. Unfortunately, in this work the presence of a validation scheme used to assess the true generalization ability of the proposed model is not fully specified. Thus, the possibility of an overfitted model cannot be discarded. Very interesting is recent work that extract features from 67 brain MRI data set and uses bagging (bootstrap aggregation with the majority vote rule) to classify known metastases, gliomas and meningiomas [13]. From five extracted features, an exhaustive search was done in order to find the optimal subset of features that feeds the classification stage. This proposal reaches as much as 97% of accuracy on discriminating Glioblastomas from Metastases. However, the best optimal subset has no medical interpretation due to the fact that the features were obtained from some purely mathematical transformations –e.g. symmetry of normal distributions, inverse difference moments and difference in variances.

More recent work [14] using MRI imaging analyzes a multi-centre ^{1}H-MRS brain tumor database [15]. A wrapper feature selection process uses direcly the performance of a single-layer perceptron (SLP), under the assumption that irrelevance affects negatively the predicted outputs of the SLP and thus can be used to detect irrelevant frequencies (seen as features of the model). A backward search strategy governs the Feature Selection (FS) process and stratified 5x5 cross-validation is carried out to choose among competing models (characterized by different sets of frequencies) and to estimate generalization ability. In this experimental setting, a remarkable 94.40% of accuracy (with 27 spectral frequencies) was reached.

Algorithm 1 Forward-Backward Spectral Selection

Input: $S = \{s_1, \ldots, s_n\}$: `Full feature set`;
C: `Class feature`
$J : 2^S \to \mathbb{R}$: `performance measure, to be maximized`

$BEST \leftarrow \{\operatorname{argmax}_{s_i \in S} J(\{s_i\})\}; j^{cur} \leftarrow J(BEST)$
$S \leftarrow S \setminus \{BEST\}$
repeat
  `***Forward Stage***`
  $s^{new} \leftarrow \operatorname{argmax}_{s_i \in S} J(BEST \cup \{s_i\})$
  $j^{new} \leftarrow J(BEST \cup \{s^{new}\})$
  if $j^{new} > j^{cur}$ **then**
    $BEST \leftarrow BEST \cup \{s^{new}\}$
    $j^{cur} \leftarrow j^{new}$
    $S \leftarrow S \setminus \{s^{new}\}$
  end if
  `***Backward Stage***`
  repeat
    $s^{new} \leftarrow \operatorname{argmax}_{s_i \in BEST} J(BEST \setminus \{s_i\})$
    $j^{new} \leftarrow J(BEST \setminus \{s^{new}\})$
    if $j^{new} \geq j^{cur}$ **then**
      $BEST \leftarrow BEST \setminus \{s^{new}\}$
      $j^{cur} \leftarrow j^{new}$
    end if
  until $BEST$ `does not change`
until $BEST$ `does not change`
Output: $BEST$ `: Best Spectral Subset`

3. An algorithm that selects subsets of spectral points

In this work we analyze this last problem, but tackling it with a simpler method. An interleaved Forward-Backward Feature Selection search is developed looking for the improvement in performance of several classification algorithms. The FS strategy needs no parameter optimization and its initialization responds to a simple deterministic criterion. Therefore, it yields an affordable solution in terms of difficulty of implementation, computational speed and interpretability by non-experts in machine learning methods. The algorithm is described in the listing **Algorithm 1**, and named FBSS: Forward-Backward Spectral Selection. Given a performance measure, to be maximized (e.g. the resampled evaluation of a classifier in a data sample), the algorithm adds and removes, in a step by step fashion, spectral points with the aim of improving current performance.

Specifically, in every iteration of the outer loop, one feature is added to the current best solution $BEST$, as long as this step improves on current performance j^{cur}. Then a variable number of feature removal steps is carried out, inasmuch the same condition of improved performance is met. This scheme is oriented to favour solutions with low numbers of features. The outer iteration also ends when no further improvement is observed. This strategy bears some

resemblances with a floating search algorithm in its forward version [16]. However, it has a far lower computational cost given that discarded features are not considered again for another inclusion round. Note also that current subset performance is not compared specifically against the best performance achieved for the *same* size of the current subset (as floating methods do).

4. Experimental settings

The examined ^{1}H-MRS data is drawn from a database belonging to the *International Network for Pattern Recognition of Tumors Using Magnetic Resonance* (INTERPRET), an European research project aimed to develop systematic tools to enable radiologists and other clinicians without special knowledge or expertise to diagnose and grade brain tumors routinely using magnetic resonance spectroscopy [15].

An essential variable in the acquisition of ^{1}H-MRS spectra is the choice of echo time. With short echo times (around 20 milliseconds), larger numbers of metabolites are detected (myoinositol, glutamate, glutamine), but it is more likely that peak superimposition will occur, causing difficulty in spectroscopic curve interpretation. By using long echo times (>135 milliseconds), most metabolites in the brain are lost (except that of choline, creatine, N-acetyl aspartate and lactate), but with better definition of peaks, thereby facilitating graphical analysis [10].

The specific data set used is constructed by single voxel ^{1}H-MR spectra acquired *in vivo* from brain tumor patients in two configurations: Long Echo Time (PRESS 135 to 144 ms), named LET, and Short Echo Time (PRESS 30 to 32 ms), named SET. Brain pathologies that conform both configurations are distributed as following:

- LET: 195 cases which include 55 meningiomas, 78 glioblastomas, 31 metastases, 20 astrocytomas grade II, 6 oligoastrocytomas grade II and 5 oligodendrogliomas grade II.
- SET: 217 cases with: 58 meningiomas, 86 glioblastomas, 38 metastases, 22 astrocytomas grade II, 6 oligoastrocytomas grade II, and 7 (SET) oligodendrogliomas grade II.

Glioblastomas (GLI) and Metastases (MET) were extracted in this study using both LET and SET data sets. Moreover, a third configuration (named LSET) was prepared in order to explore the discriminative power of the *merged* LET and SET data, formed by the 195 common observations of the LET and SET data sets.

Three well-known linear classifiers were used as a criterion function J: Fisher's Linear Discriminant Classifier (LDC), the k-nearest-neighbors technique with Euclidean metric (kNN) and hyper-parameter $k \in \{1 \dots 15\}$; and the Support Vector Machine with linear kernel (lSVM)[1] and C regularization constant as the only hyper-parameter (with $C = 2^c$, with c running from -7 to 7). In all cases, the FBSS algorithm described above is used to maximize performance of several classifiers estimated by means of averaged values of 10x10 cross-validation (ten times 10-fold cross-validation or 10x10cv for short).

[1]For the experiments, we use a MATLAB implementation; specifically, for the SVMs we use the MATLAB interface to LIBSVM [17].

Table 1: LET 10x10 cross-validation classification results (average and standard error). α indicates the value of the hyper-parameter (if any): k for kNN and C for the lSVM; the *ppm* column expresses metabolites (in ppm) of the BEST spectral subset obtained for the LET data set.

Classifier	Accuracy	α	ppm
LDC	86.21 ± 0.34	–	L3.79 L3.22 L3.03 L2.96 L1.32 L3.07
NN	88.59 ± 0.26	2	L2.75 L2.52 L2.08 L1.80 L1.49 L0.79
lSVM	92.43 ± 0.39	$2^{-0.5}$	L3.83 L3.74 L3.39 L3.22 L3.15 L2.58 L2.43 L2.08 L1.91 L1.65 L1.34

Table 2: SET 10x10 cross-validation classification results (average and standard error). α indicates the value of the hyper-parameter (if any): k for kNN and C for the lSVM; the *ppm* column expresses metabolites (in ppm) of the BEST spectral subset obtained for the SET data set.

Classifier	Accuracy	α	ppm
LDC	85.22 ± 0.25	–	S4.23 S3.51 S2.31 S1.46 S1.36 S0.64
NN	86.48 ± 0.65	4	S4.19 S2.63 S2.18 S1.86
lSVM	94.71 ± 0.22	2	S4.23 S3.79 S3.38 S3.26 S3.09 S2.98 S2.71 S2.60 S2.46 S2.35 S2.25 S1.97 S1.36

Table 3: LSET 10x10 cross-validation classification results (average and standard error). α indicates the value of the hyper-parameter (if any): k for kNN and C for the lSVM; the *ppm* column expresses metabolites (in ppm) of the BEST spectral subset obtained for the combined LSET data set.

Classifier	Accuracy	α	ppm
LDC	92.24 ± 0.23	–	L4.25 L2.37 L2.33 L2.08 L1.57 L1.30 L1.25 S3.96 S3.24 S2.39 S1.72 S1.57
NN	88.32 ± 0.54	1	S2.69 S2.67 S2.63 S1.49 S1.48
lSVM	92.92 ± 0.21	2	L4.04 L3.96 L2.94 L2.63 L2.33 L1.25 S4.10 S3.28 S2.37 S2.35 S1.40

5. Experimental results and discussion

Tables 1, 2 and 3 contain the experimental results on LET, SET and LSET data, respectively. It is observed that the lSVM classifier offers the best overall performance, specially on the SET data set, with near 95% average 10x10cv accuracy and only 13 spectral frequencies. Similar results (around 92% of 10x10cv accuracy) are delivered by LET-lSVM, LSET-LDC and LSET-lSVM with almost the same number

of spectral points (11 or 12). A non-parametric Wilcoxon signed-rank test is used for the (null) hypothesis that the median of the differences between the errors obtained by SET-lSVM against any of the combinations LET-lSVM and LSET-lSVM is zero. This hypothesis has to be rejected at the 95% level of confidence, thus showing a significant difference in medians. The obtained *p*-values are shown in Table 4 (right part). Figure 2 presents the mean spectra of the two classes (Glioblastomas and Metastases) for the SET data and the spectral points of the best solution: the combined SET-lSVM model. It is observed that most of the selected points fall into the regions that show a marked difference between signals. Table 4 (left part) gives the 10x10cv confusion matrix for the SET-lSVM model.

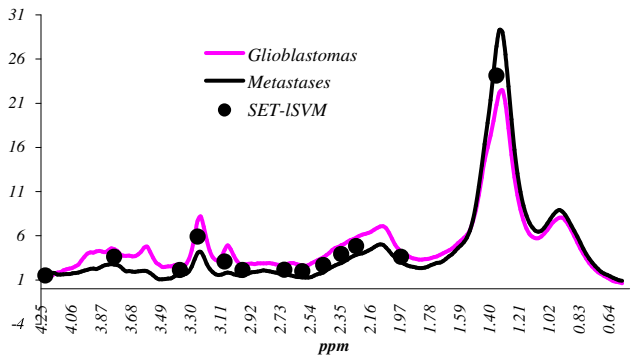

Fig. 2: The subset of spectral frequencies used by the SET-lSVM model that yields best performance –see Table 2– as positioned in the full metabolic spectrum.

Table 4: Left: Averaged 10x10cv accuracies expressed as a confusion matrix for the SET-lSVM model. True class falls vertically. Right: Wilcoxon signed rank test *p*-values comparing LET-lSVM and LSET-lSVM models against SET-lSVM.

True Class	SET-lSVM GLI	SET-lSVM MET
GLI	7.26 ± 0.07	0.34 ± 0.06
MET	0.21 ± 0.04	2.19 ± 0.05

	SET-lSVM
LET-lSVM	0.002
LSET-lSVM	0.006

6. Biological interpretation

The metabolic definition of the subset of spectral points depicted in Figure 1 is as follows [18]:

- S4.23 *Threonine*: A large neutral amino acid.
- S3.79 *Alanine*: A nonessential amino acid that has been observed in increased levels in meningiomas.
- S3.38 *Scyllo-inositol*: An isomer associated in high levels with Alzheimer's disease.
- S3.26 *Myo-inositol*: Its function is not enough understood, although it is believed to be a requirement in cell growth. Altered levels have been linked with Alzheimer's disease, hepatic encephalopathy and brain injury.
- S3.09: *Phenylalanine*: An aromatic amino acid that presents elevated readings in phenylketonuria, an abnormal phenylalanine metabolization.
- S2.98-2.60 *Glutathione-Cysteine, Glutatione-Glutamate*: An anti-oxidant essential for maintaining normal red-cell structure. Altered levels have been reported in Parkinsons's disease and other neurodegenerative diseases.
- S2.71-1.97 *NAA-Aspartate, NAA-Acetyl*: The NAA is a free amino acid whose function is poorly understood, but is is commonly believed to provide a marker of neuronal density. Among the NAA grpous, the N-Aspartate ($^3CH_2$) group and the N-Acetyl ($^2CH_3$) group were selected.
- S2.46-2.35 Glutamate-Glutamine: Glutamate is an excitatory neurotransmitter, which plays a role in mitochondrial metabolism. Glutamine plays a role in detoxification and regulation of neurotransmitter activities. These two metabolites resonate closely together.
- 2.25 *Valine*: An essential amino acid necessary for protein synthesis.
- S1.36 *Lipids/Lactate* peak: seen in condition of necrosis.

7. Conclusions

Despite being a non-invasive technique that provides rich information about the biochemistry of brain tissue, ^{1}H-MRS is still not fully consolidated as a standard method for clinical diagnosis. To become a recurrent tool, ^{1}H-MRS algorithms or machine learning proposals must be robust and highly reliable. Working towards this direction, the solutions reported in this paper give a drastic reduction in dimensionality with competitive performance in subsets of spectral points from a ^{1}H-MRS data set of brain tumors. Two classes of brain tumors were explored, Glioblastomas and Metastases which present a certain degree of similarity in their radiological spectrum, making their differentiation a difficult undertaking. In light of the promising obtained results using FBSS under the described experimental settings and considerations, some conclusions can be pointed out:

- Feature selection appears to be a solid step in dimensionality reduction in this addressed problem.
- With less than 10% of spectral points (13 out of 195, for the obtained SET-lSVM model), near 95% of accuracy can be achieved discriminating Glioblastomas from Metastases.
- Of the three data sets, given by two echo times and their fusion, the SET data set delivers better performance, confirming previous findings [5].
- The resulting subset of selected spectral frequencies has a medical interpretation in terms of known metabolites

–see [18]–, which may become helpful to radiologists as an aid in medical diagnosis.

Future research in this path will include studying more than two classes (tumour subtypes) and exploring the potential of other low-complexity classifiers.

Acknowledgments

Authors gratefully acknowledge the former INTERPRET partners (INTERPRET, EU-IST-1999-10310) and, from 1st January 2003, Generalitat de Catalunya (grants CIRIT SGR2001-194, XT 2002-48 and XT 2004-51); data providers: Dr. C. Majós (IDI), Dr.A. Moreno-Torres (CDP), Dr. F.A. Howe and Prof. J. Griffiths (SGUL), Prof. A. Heerschap (RU), Dr. W. Gajewicz (MUL) and Dr. J. Calvar (FLENI); data curators: Dr. A.P. Candiota, Ms. T. Delgado, Ms. J. Martín, Mr. I. Olier, Mr. A. Pérez and Prof. Carles Arús (all from GABRMN-UAB). Authors also acknowledge funding for CICyT TIN2009-13895-C02-01; the Mexican CONACyT and Baja California University and thank the anonymous reviewers for their helpful suggestions.

References

[1] H. Bruhn, J. Frahm, M. Gyngell, K. Merboldt, W. Hanicke, R. Sauter, and C. Hamburger, "Noninvasive differentiation of tumors with use of localized h-1 mr spectroscopy in vivo: initial experience in patients with cerebral tumors," *Radiology*, vol. 172, pp. 541–548, 1989.

[2] J. Hansen, W. Backes, K. Nicolay, and M. Kooy, "^{1}h-mr spectroscopy of the brain: Absolute quantification of metabolites," *Radiology*, vol. 246, no. 2, pp. 318–332, August 2006.

[3] A. Tate, S. Crabb, J. Griffiths, S. Howells, R. Mazucco, L. Rodrigues, and D. Watson, "Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra," *NMR in Biomedicine*, vol. 19, pp. 411–434, 2006.

[4] J. Garcia, S. torjada, C. Vidal, M. Julià-Sapé, J. Luts, S. V. Huffel, C. Arús, and M. Robles, "On the use of long te and short te sv mr. spectroscopy to improve the automatic brain. tumor diagnosis." In ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/ida/reports/07-55.pdf, Tech. Rep., 2007.

[5] A. Vellido, E. Romero, F. González-Navarro, L. Belanche-Munoz, M. Julià-Sapé, and C. Arús, "Outlier exploration and diagnostic classification of a multi-centre ^{1}h-mrs brain tumour database," *Neurocumputing*, vol. 13-15, pp. 3085–3097, Elsevier 2009.

[6] S. Kirstie, M. Murphy, P. Wilkins, A. Bell, J. Griffiths, and F. Howe, "Differentiation of metastases from high-grade gliomas using short echo time ^{1}h spectroscopy," *Journal of Magnetic Resonance Imaging*, vol. 20, no. 2, pp. 187–192, 2004.

[7] W. H. Organization, *WHO Classification of Tumours of the Central Nervous System*, I. A. for Research on Cancer, Ed., 2007.

[8] G. Weiss, *Clinical Oncology*. Appleton and Lange, 1998.

[9] F. Ali-Osman, *Brain tumors*. Humana Press, 2005.

[10] M. Castillo, L. Kwock, and S. Mukherji, "Clinical applications of proton mr spectroscopy," *AJNR*, vol. 17, pp. 1–15, January 1996.

[11] S. Cha, S. Cha, J. Lupo, M. Chen, K. Lamborn, M. McDermott, M. Berger, S. Nelson, and W. Dillon, "Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion mr imaging," *American Journal of Neuroradiology*, vol. 28, pp. 1078–1084, 2007.

[12] S. Wang, S. Kima, S. Chawlaa, R. Wolfa, W. Zhanga, D. O'Rourkeb, K. Judyb, E. Melhema, and H. Poptania, "Differentiation between glioblastomas and solitary brain metastases using diffusion tensor imaging," *Neuroimage*, vol. 44, pp. 653–660, 2009.

[13] P. Georgiadis, , D. Cavouras, I. Kalatzis, D. Glotsos, E. Athanasiadis, S. Kostopoulos, K. Sifaki, M. Malamasand, G. Nikiforidis, and E. Solomou, "Enhancing the discrimitation accuracy between metastases, gliomas and meningiomas of brain mri by volumetric tectural features and ensemble pattern recognition methods," *Magnetic Resonance Imagen*, vol. 27, pp. 120–130, 2009.

[14] E. Romero, A. Vellido, M. Julià-Sapé, and C. Arús, "Discriminating glioblastomas from metastases in a sv ^{1}h-mrs brain tumor database," in *European Society for Magnetic Resonance in Medicine and Biology (ESMRMB Congress 2009)*, 2009.

[15] INTERPRET, "International network for pattern recognition of tumours using magnetic resonance project," 2002, http://azizu.uab.es/INTERPRET.

[16] P. Pudil, J. Novovicová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 11, pp. 1119–1125, 1994.

[17] C. Chang and C. Lin, "Libsvm : a library for support vector machines," 2002, in http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

[18] V. Govindaraju, K. Young, and A. Maudsley, "Proton nmr chemical shifts and coupling constants for brain metabolites," *NMR in Biomedicine*, vol. 13, pp. 129–153, 2000.