



UNIVERSITAT AUTÒNOMA DE BARCELONA

Grup d'Aplicacions Biomèdiques de la Ressonància Magnètica Nuclear.

**Comparativa de dues eines de *software* pel
desenvolupament de classificadors
d'Espectroscopia de Ressonància Magnètica
Nuclear de tumors cerebrals humans**

Silvia Ferrer Pueyo

16 de Juliol de 2010

Comparativa de dues eines de *software* pel desenvolupament de classificadors d'Espectroscopia de Ressonància Magnètica Nuclear de tumors cerebrals humans.

Memòria de pràctiques presentada per Silvia Ferrer Pueyo corresponent al treball d'investigació del Màster Oficial de Bioquímica, Biologia Molecular i Biomedicina de l'itinerari A.

Aquest treball ha estat realitzat a la Facultat de Biociències, al departament de Bioquímica i Biologia Molecular i al Grup d'Aplicacions Biomèdiques de la Ressonància Magnètica Nuclear (GABRMN <<http://gabrmn.uab.es/>>) de la Universitat Autònoma de Barcelona sota la supervisió de la Dra. M^a Margarida Julià Sapé i del Dr. Carles Arús Caraltó.

Cerdanyola del Vallès, 16 de juliol de 2010

Dra. M^a Margarida Julià Sapé

Dr. Carles Arús Caraltó

Silvia Ferrer Pueyo

ÍNDEX

| | |
|--|-----|
| ÍNDEX | III |
| ÍNDEX DE FIGURES I TAULES | V |
| ABREVIATURES..... | VI |
| 1. RESUM | 1 |
| 2. INTRODUCCIÓ | 2 |
| 2.1. Espectroscopia de Ressonància Magnètica..... | 2 |
| 2.1.1. Espectre de ressonància magnètica | 3 |
| 2.1.2. Paràmetres d'adquisició | 4 |
| 2.2. Metabòlits d'interès | 4 |
| 2.3. Tumors cerebrals humans..... | 6 |
| 2.3.1. Glioblastoma multiforme (GBM), grau IV..... | 7 |
| 2.3.2. Metàstasi (ME), grau IV | 7 |
| 2.3.3. Astrocitoma difús (A2), grau II..... | 7 |
| 2.3.4. Oligodendroglioma (OD), grau II | 8 |
| 2.3.5. Oligoastrocitoma (OA), grau II..... | 8 |
| 2.3.6. Meningioma (MN), grau I | 8 |
| 2.4. Anàlisi multivariant | 8 |
| 2.4.1. Selecció o reducció de variables..... | 8 |
| 2.4.2. Anàlisi discriminant lineal (LDA) | 9 |
| 2.4.3. Avaluació de classificadors. | 9 |
| 2.4.4. <i>Softwares</i> pel desenvolupament de classificadors..... | 10 |
| 3. OBJECTIUS..... | 11 |
| 4. MATERIALS I MÈTODES..... | 12 |
| 4.1. Origen de les dades utilitzades | 12 |
| 4.2. Adquisició de les dades | 12 |
| 4.3. Processat de les dades | 13 |
| 4.4. Dades utilitzades: training set i test set | 14 |

| | | |
|--------|--|----|
| 4.5. | Anàlisi de dades..... | 14 |
| 4.5.1. | SPSS | 15 |
| 4.5.2. | SC..... | 16 |
| 5. | RESULTATS | 17 |
| 5.1. | Variables significatives a TEC..... | 17 |
| 5.2. | Espectres promig..... | 17 |
| 5.3. | Resultats de classificació i avaluació amb l'SPSS..... | 20 |
| 5.4. | Resultats de classificació i avaluació amb l'SC | 21 |
| 6. | DISCUSSIÓ | 24 |
| 7. | CONCLUSIONS..... | 29 |
| 8. | BIBLIOGRAFIA..... | 30 |

ÍNDIX DE FIGURES I TAULES

| | |
|---|----|
| Figura 1. Transformació del senyal FID en el domini del temps a freqüències de ressonància, en un espectre d'un estàndard a alta resolució (in-vitro) on s'ha afegit TMS (Tetrametilsilà) [3]. | 3 |
| Figura 2. Percentatges de la incidència relativa dels tumors cerebrals primaris i del SNC als Estats Units del 2004-2006. Dades extretes del CBTRUS (Central Brain Tumour Registry of USA) [15]. | 6 |
| Figura 3. Esquema de l'anàlisi multivariant utilitzat en aquest treball ambdós softwares. | 15 |
| Figura 4. A dalt: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del training set. Abaix: Perfil espectroscòpic promig dels dos test set junts. En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN, u.a. fa referència a les unitats arbitràries. | 18 |
| Figura 5. A dalt: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del test set del CDP. Abaix: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del test set de l'IDI-Badalona. En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN; u.a. fa referència a les unitats arbitràries. | 19 |
| Figura 6. Encert de la classificació versus el nombre de variables seleccionades amb l'SPSS. | 20 |
| Figura 7. Diferencial de l'encert del training set versus el test set global i per separat fent servir l'SPSS. La línia continua correspon al test set global, la discontinua al CDP i la línia de punts a l'IDI-Badalona. | 21 |
| Figura 8. Encert de la classificació versus el nombre de variables seleccionades amb l'SC. | 22 |
| Figura 9. A l'esquerra: diferencial de l'encert del training set versus els test sets de 17 a 2 variables. La línia continua correspon al test set global, la discontinua al CDP i la línia de punts a l'IDI-Badalona. A la dreta: diferencial de l'encert del training set versus el conjunt test set de 100 a 2 variables. | 22 |
| Figura 10. A l'esquerra: Distribució a l'espai de classificació de les classes estudiades del training set amb 8 variables, el nombre òptim (90,3% d'encert). A la dreta: Distribució del test set amb 8 variables (84,4% d'encert). En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN. Les línies rectes representen les fronteres entre les classes definides pel classificador. | 23 |
| Figura 11. A l'esquerra: Distribució a l'espai de classificació de les classes estudiades del training set amb 100 variables (99,5% d'encert). A la dreta: Distribució del test set amb aquestes 100 variables (73,4% d'encert). En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN. Les línies rectes representen les fronteres entre les classes definides pel classificador. | 23 |
| Figura 12. Espectre i imatge de referència d'un pacient (et3321 en codi de la base de dades eTUMOUR) amb meningioma atípic de l'IDI-Badalona. La línia continua representa el perfil espectroscòpic d'aquest cas i en discontinu observem el perfil promig dels meningiomes del training set (n=58). | 25 |
| Figura 13. Espectre d'un tumor amb edema del CDP i la seva imatge de referència on està senyalat en groc la zona de l'edema i en vermell el vòxel per a l'enregistrament de la MRS. | 27 |
| Figura 14. A dalt: exemple de GBM atípic mal classificat pels classificadors desenvolupats pels dos softwares amb la imatge de referència del vòxel a TEC. Al mig: espectre d'un GBM típic i la imatge de referència. A sota: espectre d'un A2 típic amb la imatge del vòxel. | 28 |
| Figura 15. Espectre d'un OA que presentava una zona de senyal elevada entre els 2,0-2,5 ppm com els meningiomes i la seva imatge de referència del vòxel. | 28 |
| | |
| Taula 1. Desplaçament químic versus l'aigua a 4,75 ppm dels principals compostos que es poden detectar en diferents teixits mitjançant l'espectroscopia de ressonància magnètica "in vivo" segons es cita a la referència [6]. | 5 |
| Taula 2. Classificació WHO i característiques dels diferents graus dels tumors cerebrals humans. | 7 |
| Taula 3. Descripció dels training i test set utilitzats en aquest estudi. | 14 |
| Taula 4. Resum de les variables més rellevants a TEC seleccionades per l'SC i l'SPSS i el possible origen bioquímic de la variable seleccionada. | 17 |

ABREVIATURES

| | |
|--------------------|--|
| Ala | Alanina |
| A2 | Astrocitoma de grau II |
| BOO | <i>Bootstrapping</i> |
| Cho | Grup de metabòlits que contenen el grup trimetilamina, principalment: colina lliure, fosfocolina i glicerofosfocolina. |
| Cre | Creatina total |
| DMS | <i>Data Manipulation Software</i> |
| FID | <i>Free-induction decay</i> |
| GABRMN | Grup d'Aplicacions Biomèdiques de la Ressonància Magnètica Nuclear |
| GBM | Glioblastoma multiforme |
| Gln | Glutamina |
| Glu | Glutamat |
| Glx | Conjunt glutamat + glutamina |
| Gly | Glicina |
| ¹ H ERM | Espectroscopia de Ressonància Magnètica de protó |
| HLSVD | <i>Hankel-Lanczos singular value decomposition</i> |
| JAVA | Llenguatge de programació |
| Lac | Lactat |
| LDA | <i>Linear Discriminant Analysis</i> |
| LOO | <i>Leave One Out</i> |
| ml | Myo-inositol |
| MRI | <i>Magnetic Resonance Imaging</i> |
| MRS | <i>Magnetic Resonance Spectroscopy</i> |
| MV | Multi Voxel |
| NAA | N-Acetil Aspartat |
| OA | Oligoastrocitomes |
| OD | Oligodendroglomes |

| | |
|-------|--|
| ppm | Parts per milió (en RMN, escala adimensional de referenciació del desplaçament químic envers d'una referència) |
| PRESS | <i>Point Resolved Spectroscopy</i> |
| RMN | Ressonància Magnètica Nuclear |
| SC | Spectra Classifier 1.0 |
| SNC | Sistema Nerviós Central |
| SPSS | <i>Statistical Product and Service Solutions</i> |
| STEAM | <i>Stimulated Echo Acquisition Mode</i> |
| SV | Single Voxel |
| T | Tesla |
| TEC | Temps d'eco curt |
| TEL | Temps d'eco llarg |
| u.a. | Unitats arbitràries |
| WHO | World Health Organization |

1. RESUM

En aquest treball s'ha fet una avaluació comparativa dels resultats que es poden obtenir amb el *software* SpectraClassifier 1.0 (SC) desenvolupat al nostre grup de recerca, comparant-lo amb l'SPSS, un programa estadístic informàtic estàndard, en un problema de classificació de tumors cerebrals humans amb dades d'espectroscopia de ressonància magnètica de protó (¹H-ERM). El interès d'aquesta avaluació comparativa radica en la documentació dels resultats obtinguts amb els dos sistemes quan en la correcció dels resultats obtinguts, així com ponderar la versatilitat i usabilitat dels dos paquets de *software* per a una aplicació concreta d'interès al treball del GABRMN.

Per a aquest treball s'han utilitzat dades provinents de dos projecte europeus multicèntrics (INTERPRET i eTumour) en els quals vam participar. Les classes tumorals utilitzades (d'un total de 217 pacients) han sigut les majoritàries des del punt de vista epidemiològic: glioblastoma multiforme, metàstasi, astrocitomes de grau II, oligodendrogliomes de grau II, oligoastrocitomes de grau II i meningiomes de baix grau.

Amb les dades d'aquests pacients s'han dissenyat classificadors basats en l'anàlisi discriminant lineal (LDA), s'han avaluat amb diferents mètodes matemàtics i s'han testat amb dades independents.

Els resultats han estat satisfactoris, obtenint amb l'SC resultats més robusts amb dades independents respecte la classificació realitzada per l'SPSS.

2. INTRODUCCIÓ

2.1. *Espectroscopia de Ressonància Magnètica*

La ressonància magnètica és una tècnica no invasiva que permet l'estudi de sistemes vius, mitjançant l'obtenció d'imatges (MRI), amb la qual podem analitzar la morfologia, l'anatomia i la funcionalitat dels teixits. L'espectroscopia (MRS) és una de les variants de la tècnica que permet detectar certes molècules en dissolució i en l'ordre de concentració milimolar, i d'on podem obtenir perfils de metabòlits al teixit d'interès, com una metabolòmica *in vivo*.

El fenomen de la ressonància magnètica nuclear (RMN) es basa en l'existència de certs nuclis atòmics, com per exemple el del ^1H (protó), amb una abundància natural del 99%, i que té la propietat coneguda com *spin* (o rotació), que fa que aquest nucli "giri" sobre el seu propi eix en presència d'un determinat camp magnètic i adopti dues orientacions diferents.

Quan s'aplica un **camp magnètic** a la mostra, els nuclis de ^1H adoptaran una de les dues orientacions possibles, determinades per la propietat de l'*spin*, les quals tenen energies lleugerament diferents, i el moment magnètic degut al *spin* nuclear gira al voltant de l'eix definit pel camp magnètic al qual està exposada la mostra; això es coneix com el moviment de precessió i està determinat segons l'equació de Larmor [1].

$$\nu = \gamma \cdot B / 2\pi$$

on ν es la freqüència de precessió, γ la constant giromagnètica i B el camp magnètic resultant del camp magnètic extern aplicat.

L'aplicació d'un **pols de radiofreqüència (RF)** excitarà els protons, provocant transicions entre nivells energètics, i mentre retornen a la situació inicial en absència d'excitació (procés de relaxació), produiran el senyal de ressonància detectable. Aquest temps de relaxació està descrit per les constants longitudinal (T_1) i transversal (T_2), dependents de la mostra.

Així, l'RMN estudia els nuclis que tenen *spin*, introduint-los dins d'un camp magnètic constant, per posteriorment, pertorbar-los amb l'ús de polsos de radiofreqüència (RF) de curta durada a la freqüència de precessió dels nuclis de la mostra [1].

Al deixar d'irradiar-la, els nuclis es relaxen tornant a la situació inicial d'equilibri. El senyal detectat va disminuint fins arribar a zero i queda recollit en el domini del temps per una bobina detectora de corrent; aquesta caiguda d'intensitat del petit corrent induït per la mostra degut a l'absència del pols de RF inicial es coneix amb el nom de *Free Induction Decay (FID)* [2].

La **Transformació de Fourier (TF)**, és la funció matemàtica més utilitzada i eficaç per analitzar aquests senyals de RMN, transformant-los del domini del temps al de les freqüències, que és on es poden observar els pics de ressonància dels protons de les diferents molècules observables per MRS.

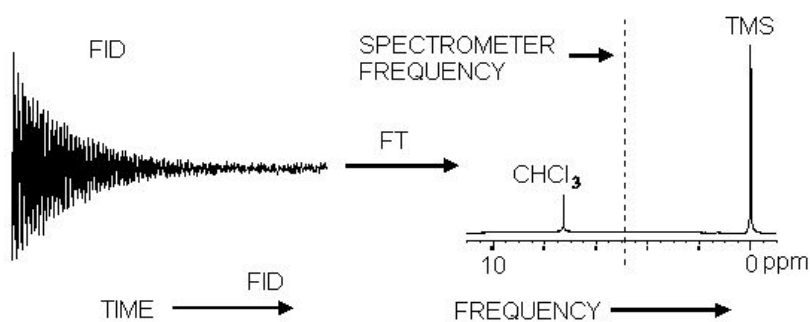


Figura 1. Transformació del senyal FID en el domini del temps a freqüències de ressonància, en un espectre d'un estàndard a alta resolució (*in-vitro*) on s'ha afegit TMS (Tetrametilsilà) [3].

El processament i l'anàlisi de les dades resultants tenen per objectiu final obtenir informació dels compostos presents en la mostra, i en alguns casos, determinar-ne la seva concentració. Així doncs, una vegada s'han adquirit les dades i s'han processat, obtenim un gràfic representació de la intensitat del senyal emès pel nucli de ^1H en funció de la seva freqüència de ressonància característica en ppm (també coneguda com desplaçament químic envers una referència freqüencial, interna o externa), que ve donada pel seu entorn molecular. Aquest gràfic es coneix amb el nom d'espectre de ressonància magnètica.

2.1.1. Espectre de ressonància magnètica

A partir d'un espectre de RMN podem:

- Identificar diferents senyals de grups químics funcionals de metabòlits per la seva freqüència de ressonància característica.
- Quantificar els metabòlits segons l'alçada o l'àrea, utilitzant una referència interna com podria ser l'aigua [4].

Així, podem detectar els compostos que originen el senyal de RMN i com canvia la seva concentració en situacions patològiques.

2.1.2. Paràmetres d'adquisició

Durant l'adquisició de les dades hi ha diferents paràmetres importants que podem controlar, els quals poden variar el patró espectral final i són; entre altres: el camp magnètic, el temps d'eco (TE), la seqüència de polsos de RF per l'adquisició de les dades i el volum de la mostra.

El camp magnètic estàndard que es fa servir als escàners clínics de ressonància magnètica que fan MRS té una intensitat de camp de 1,5 T, tot i que, actualment cada vegada són més freqüents els de 3T. TE és l'espai de temps que transcorre des d'un pols de RF fins a l'obtenció de l'eco al qual s'aplica la TF per donar l'espectre; se sol acceptar per convenció que el temps d'eco curt (TEC) és per intervals de temps menors que 40 ms, i el temps d'eco llarg (TEL) a partir de 40 ms. Respecte la seqüència, en espectroscopia clínica, podem parlar de l'*STEAM (Stimulated Echo Acquisition Mode)* i de la *PRESS (Point Resolved Spectroscopy)*, que són diferents tècniques per localitzar el volum que volem analitzar [5]. Finalment, el volum d'interès d'on s'adquireix el senyal podrà ser únic (*single voxel, SV*) o se'n podran utilitzar varis (*multi voxel, MV*).

2.2. Metabòlits d'interès

Una vegada obtenim l'espectre de ressonància magnètica de la mostra en estudi, cal identificar i interpretar-ne els resultats. Moltes de les ressonàncies observades tenen funcions metabòliques importants, per exemple, per la síntesi de membranes o la bioenergètica cel·lular.

Les ressonàncies principals que poden aparèixer en un espectre de RMN *in vivo* mesurades amb un escàner de 1,5T en diferents teixits, es descriuen a la Taula 1.

| (ppm) | Metabolitos |
|-----------|--|
| 0,8 - 1,1 | Leucina (Leu), isoleucina (Ile), valina (Val) |
| 0,8 - 2,5 | Àcids grassos (Lip) |
| 1,15 | Propilenglicol, etanol |
| 1,3 | Àcido Làctico (Lac) |
| 1,45 | Alanina (Ala) |
| 1,85 | Àcido acètic (Ac) |
| 2,02 | N-acetilaspártato (NAA), N-acetilaspártilglutamato (NAAG) |
| 2,1 - 2,5 | Àcido glutàmic (Glu), glutamina (Gln) |
| 2,25 | GABA |
| 2,6 | N-acetilaspártato, citrato (Cit) |
| 2,8 | Àcido aspàrtic (Asp) |
| 3,02 | Creatina (Cr), fosfocreatina (PCr) |
| 3,2 | Colina, etanolamina, fosforilcolina, fosforiletanolamina, Glicerofosforilcolina, glicerofosforiletanolamina (Cho), carnitina (Car) |
| 3,3 | Taurina (Tau), scyllo-inositol (sino) |
| 3,4 | Glucosa (Glc) |
| 3,55 | Myo-inositol (mIns), glicina (Gly) |
| 3,6 - 3,8 | Àcido glutàmic, glutamina |
| 3,5 - 4,0 | Arabitol, ribitol |
| 3,8 | Glucosa, manitol |
| 3,9 | Creatina, fosfocreatina |
| 5,3 - 5,7 | Àcids grassos |
| 7,0 | Carnosina |
| 7,3 | Fenilalanina (Phe) |
| 8,0 | Carnosina |

Taula 1. Desplaçament químic *versus* l'aigua a 4,75 ppm dels principals compostos que es poden detectar en diferents teixits mitjançant l'espectroscopia de ressonància magnètica "in vivo" segons es cita a la referència [6].

Les alteracions espectrals més destacables en els tumors cerebrals humans (veure també apartat 2.3) es descriuen a continuació:

- Disminució o absència de la ressonància de l'N-Acetil-Aspartat (NAA), que reflexa la pèrdua de neurones i axons i/o de la seva funcionalitat [7].
- El lactat (Lac) [8] que normalment s'associa a l'existència d'una alta activitat proliferativa tumoral, i també es troba en líquids quàstics.
- L'Alanina (Ala) se sol observar als meningiomes (tumor a les meninges) [9] i pot aparèixer en altres tipus tumorals tot i que en menor proporció.
- Els lípids es relacionen amb l'existència de necrosi [10], pel que solen trobar-se en tumors d'alt grau, tals com metastasis i glioblastomes, però també en altres tipus. La seva presència s'utilitza com a marcador de necrosi [11].

- La creatina i la fosfocreatina (Cre) normalment estan disminuïdes, cosa que suggereix l'existència d'unes necessitats d'esmoreïment energètic inferiors a les del parènquima cerebral normal.
- La colina i compostos similars com la fosfocolina i la glicerofosfocolina (Cho). El seu augment se sol explicar per l'augment de la proliferació cel·lular [6]. El quocient entre Cre i Cho és utilitzat freqüentment per a comparar entre diferents patologies i casos clínics com a marcador de malignitat [12].

2.3. Tumors cerebrals humans

Els tumors del Sistema Nerviós Central (SNC) en adults, tenen una incidència d'aproximadament el 2% del total dels tipus tumorals. Tot i així, tenen un gran índex de mortalitat i generació d'incapacitat [13].

L'Organització Mundial de la Salut va establir un sistema de classificació dels diferents tumors del SNC en funció del seu grau de malignitat (Taula 2). Aquest consisteix en subdividir els tumors en quatre categories basant-se en criteris histopatològics i de supervivència dels pacients [14].

Per a la realització d'aquest treball es van seleccionar els sis tumors amb més prevalència, i que es descriuen a continuació:

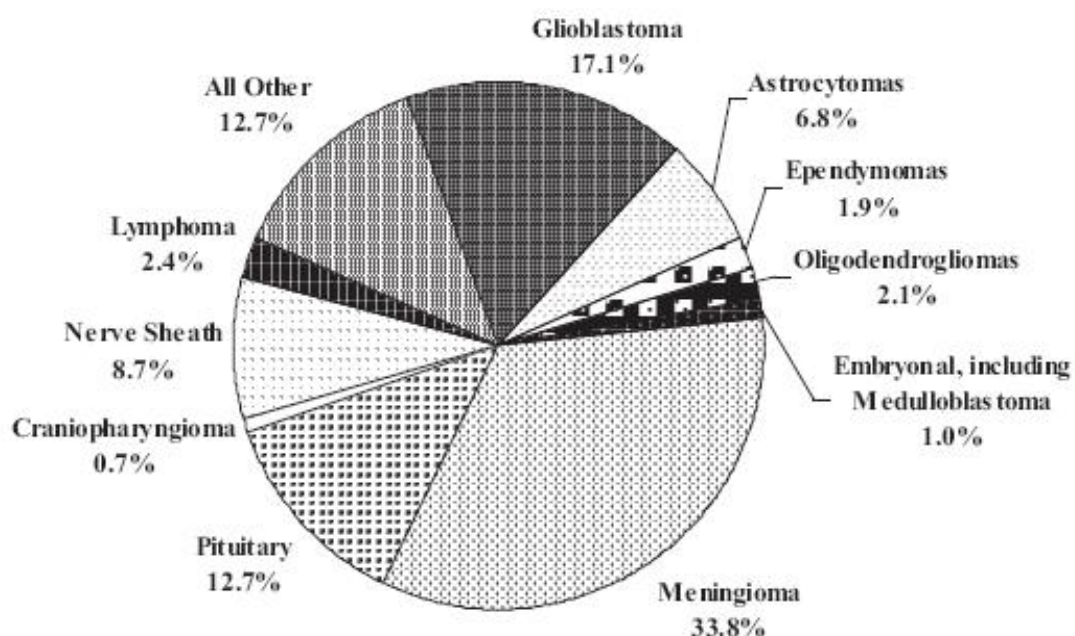


Figura 2. Percentatges de la incidència relativa dels tumors cerebrals primaris i del SNC als Estats Units del 2004-2006. Dades extretes del CBTRUS (Central Brain Tumour Registry of USA) [15].

| Classificació WHO | Característiques tumorals |
|-------------------|--|
| Grau I | Baix potencial proliferatiu i cèl·lules diferenciades. Possibilitat de cura després de la resecció quirúrgica |
| Grau II | Capacitat infiltrativa, baixa activitat mitòtica però potencial per a ser recurrents. Alguns tumors solen progressar cap a graus més alts. |
| Grau III | Alta activitat mitòtica, capacitat infiltrativa manifestada i anaplàsia. Evidència histològica de malignitat. |
| Grau IV | Alta activitat mitòtica, propensió a la necrosi i associats a la progressió ràpida pre i post-operatòria. |

Taula 2. Classificació WHO i característiques dels diferents graus dels tumors cerebrals humans.

2.3.1. Glioblastoma multiforme (GBM), grau IV

Com podem observar a la Figura 2, els glioblastomes multiformes (GBM) són els tumors glial primaris amb més incidència i els que tenen el pitjor pronòstic. Progressen a partir dels astrocitomes anaplàsics o s'originen *de novo*. Es caracteritzen per una alta capacitat proliferativa amb tendència a la invasivitat i amb la presència de pleomorfisme cel·lular i necrosi. L'esperança de vida dels malalts és aproximadament un any [16, 17].

2.3.2. Metàstasi (ME), grau IV

Els tumor metastàsics (ME) s'originen fora del SNC però arriben al cervell ja sigui per invasió directa de teixits adjacents, o per transport limfàtic. L'origen és divers, essent el tumor primari del tracte respiratori el més comú (50% de les metàstasi cerebrals). Espectroscòpicament és difícil diferenciar les metàstasis dels GBM, per això se solen estudiar dins del mateix grup [10, 18-20].

2.3.3. Astrocitoma difús (A2), grau II

Estan caracteritzats per un alt grau de diferenciació cel·lular, un creixement lent i certa capacitat infiltrativa. Igual que els GBM i les metàstasis tenen origen astrocític, i tot i que són tumors de baix grau, tenen la tendència intrínseca de progressar cap a graus més alts.

2.3.4. Oligodendroglioma (OD), grau II

Són tumors glials molt ben diferenciats, localitzats principalment als hemisferis cerebrals i procedents de cèl·lules precursors de l'oligodendroglioma. Segons les estadístiques mostrades a la figura 1, els oligodendrogliomes constitueixen el 2,1% del total dels tumors cerebrals.

2.3.5. Oligoastrocitoma (OA), grau II

Els oligoastrocitomes tenen característiques intermèdies entre dos tipus tumorals: els oligodendrogliomes i els astrocitomes difusos de grau II. Se solen formar als hemisferis cerebrals i es caracteritzen per tenir una baixa activitat mitòtica i per l'absència de necrosi; també poden presentar microcalcificacions.

2.3.6. Meningioma (MN), grau I

Són els tumors primaris amb millor pronòstic. Es formen a les meninges i tenen força similitud amb les cèl·lules aracnoidees [21]. Es consideren tumors benignes per la baixa capacitat proliferativa i infiltrativa, i l'alta taxa de recuperació dels malalts després de la cirurgia de resecció.

2.4. Anàlisi multivariant

L'anàlisi multivariant és una disciplina estadística utilitzada per determinar el paper de varis factors de la mostra. Dintre de l'anàlisi multivariant, i més concretament dins de l'anàlisi de classificació, existeixen diverses tècniques, unes dedicades a la selecció de variables i altres utilitzades per a la classificació pròpiament dita, com l'anàlisi discriminant lineal (LDA). En aquest últim es centrarà aquest treball. Aquest anàlisi de classificació seria molt laboriós si no tinguéssim a la disposició eines informàtiques tals com l'SPSS [22] o l'SC [23], entre altres.

2.4.1. Selecció o reducció de variables.

La selecció de variables és una tècnica que discrimina entre les característiques més rellevants d'un conjunt de vectors de dades i que són capaces de crear un model de classificació robust dels possibles subgrups o classes del conjunt d'aquestes dades. Existeixen diferents tècniques de selecció de variables, com la *forward feature selection* i la *backward feature selection* [24]. Per seleccionar les variables d'interès, s'utilitzen criteris matemàtics, entre els que podem destacar amb respecte al present treball:

- Criteri de Lambda de Wilks [25], utilitzat per l'SPSS.

- Criteri de *correlation-based feature subset selection* [26], utilitzat per l'SC.

2.4.2. Anàlisi discriminant lineal (LDA)

L'LDA consisteix a trobar funcions discriminants que són combinacions lineals de les variables seleccionades als nostres espectres. Aquestes funcions discriminants han d'agrupar els casos en classes de manera que la variació intra-classe sigui mínima i la inter-classe sigui màxima. A més, aquest mètode és molt interessant perquè permet predir a quin grup pertany un nou cas no utilitzat prèviament per desenvolupar les funcions discriminants lineals [24].

Per a començar un LDA cal crear la matriu de dades amb les quals entrenarem el sistema (*training set*). Per avaluar la robustesa de les nostres funcions discriminants, es recomana utilitzar una matriu independent (*test set*) amb la qual testarem la qualitat del classificador obtingut.

2.4.3. Avaluació de classificadors.

Una vegada s'obté el resultat de l'LDA, caldrà avaluar la robustesa de les fórmules per assegurar-nos de que es tracta d'un bon classificador i per validar-los inicialment sense utilitzar el *test set*. Existeixen diferents mètodes d'avaluació, els quals s'aplicaran a la matriu utilitzada per l'entrenament del sistema, al conjunt de *training*. Per ordre de menys a més representativitat:

- *Fold cross-validation* o validació creuada: consisteix en fer "n" particions de les dades en subgrups, testant la funció discriminant obtinguda en cada subgrup, de 1 a n. [27].
- *Leave One Out* (LOO): és una variant del *fold cross-validation* que divideix la mostra en tants grups com casos tinguem, testant el classificador amb un cas cada vegada.
- *Bootstrapping* (BOO): actua igual que el LOO, però fent un remostratge amb repetició, de manera que es pot repetir un nombre determinat i gran de vegades (per exemple 1000 cops). És el mètode més representatiu que normalment s'ajusta més a la realitat.

2.4.4. Softwares pel desenvolupament de classificadors

2.4.4.1. SPSS

L'SPSS [22] és un programa d'aplicacions estadístiques molt utilitzat en ciències socials i en treballs d'investigació de mercat. Té la capacitat de permetre l'ús de bases de dades de gran mida, a més de permetre anàlisis multivariants i de classificació, entre d'altres tècniques.

2.4.4.2. Spectra Classifier (SC)

L'Spectra Classifier 1.0 [28] és un programa en JAVA desenvolupat al GABRMN, que ens facilita l'ús de les eines informàtiques per dissenyar i implementar els classificadors basats en l'espectroscopia de RMN *in-vivo* de *Single Voxel* (SV) i *Multi Voxel* (MV). El seu principal objectiu és permetre als usuaris menys experimentats en l'anàlisi multivariant obtenir un anàlisi automatitzat de reconeixement de patró de les seves dades.

3. OBJECTIUS

L'Objectiu principal d'aquest treball ha estat la avaluació comparativa d'un *software*, l'SC [28], creat pel Grup d'Aplicacions Biomèdiques de la Ressonància Magnètica Nuclear (GABRMN), comparant els resultats obtinguts amb un programa estadístic estàndard, l'SPSS 17.0, amb un conjunt de dades d'MRS SV procedents de dos estudis multicèntrics en pacients afectats per tumors cerebrals.

4. MATERIALS I MÈTODES

4.1. Origen de les dades utilitzades

Les dades utilitzades en aquest treball procedeixen en primer lloc de pacients participants al Projecte Europeu INTERPRET: “International Network for Pattern Recognition of Tumour using Magnetic Resonance” (EU-IST-1999-10310) [10, 29]; i en segon lloc del Projecte Europeu eTUMOUR: “Web Accessible MR Decision Support System for Brain Tumour Diagnosis and Prognosis, Incorporating *in vivo* and *ex vivo* Genomic and Metabolomic Data” (FP6-2002-LIFESCHIHEALTH 503094) [19]

Els centres participants en el projecte INTERPRET i d'on s'han adquirit aquestes dades van ser l'IDI-Bellvitge (Institut de Diagnòstic per la Imatge de l'Hospital Universitari de Bellvitge, a l'Hospitalet de Llobregat, Barcelona), l'IDI-Badalona (IDI de l'Hospital Germans Trias i Pujol, Badalona), el Centre de Diagnòstic Pedralbes a Esplugues de Llobregat, Barcelona (CDP-CETIR), St George's University of London, UK (SGUL); Universitair Medisch Centrum Nijmegen, The Netherlands (UMCN); Fundación para la Lucha contra las Enfermedades Neurológicas de la Infancia, Buenos Aires, Argentina (FLENI) i de la Uniwersytet Medyczny w Lodz, Lodz, Poland (MUL). Les dades procedents del projecte eTUMOUR utilitzades en aquest treball es van obtenir a l'IDI-Badalona (IDI de l'Hospital Germans Trias i Pujol, Badalona) i al Centre de Diagnòstic Pedralbes a Esplugues de Llobregat, Barcelona (CDP-CETIR).

Les dades de tots els pacients després de signar el consentiment informat van passar a formar part de les bases de dades dels projectes [30, 31] de les quals es van extreure els seus espectres SV de ressonància magnètica i el diagnòstic histopatològic.

Degut a la migració que es va efectuar posteriorment a l'establiment inicial de la base de dades d'INTERPRET [31], es va realitzar un control de qualitat, tant de les dades clíniques com de les espectrals seguint el protocol publicat [32], per així garantir la qualitat, traçabilitat i precisió de les dades presents en la base de dades finalment utilitzada.

4.2. Adquisició de les dades

Tant durant el projecte INTERPRET com en el projecte eTUMOUR es van definir protocols d'adquisició de les dades per garantir els resultats obtinguts als diferents hospital i poder comparar-los. A l'article de Tate, *et al* [10] estan descrits els protocols d'INTERPRET que posteriorment es varen fer servir pel projecte eTUMOUR. Els centres van

utilitzar, entre d'altres, el protocol d'espectroscopia de volum únic (single voxel o SV) a un camp magnètic de 1,5T [29]. Es van obtenir llavors espectres a TEC (20-32 ms) amb supressió de l'aigua, si era disponible.

Per adquirir les dades de SV es van utilitzar dues seqüències diferents: STEAM, amb TEC de 20 ms i PRESS amb TEC de 30 i 32 ms. En estudis anteriors es va demostrar que tot i variar lleugerament el patró espectral, es podien combinar perfectament les dades pel desenvolupament de classificadors matemàtics [10, 33].

Es va establir que la mida del vòxel havia de ser el més gran possible i representativa del tumor, a més d'evitar la contaminació amb teixit normal, necrosi o zones cístiques adjacents [10], resultant en uns volums d'entre 1,7 i 8 cm³.

4.3. *Processat de les dades*

Per estandarditzar el procediment de transformació del senyal en el domini del temps o en el de les freqüències, es va utilitzar la DMS (*Data Manipulation Software*) [34], un programa informàtic que automatitza el processament dels espectres adquirits en els diferents escàners de RMN tot i els seus diferents formats i així evitar els errors de manipulació manual dels usuaris.

Els passos que segueix aquest *software* són els següents:

1. Ajust de la fase dividint el senyal de l'espectre de metabòlits amb el de l'espectre sense supressió de l'aigua, segons el mètode de Klose [35].
2. Ajust de la freqüència del pic de l'aigua a 4,75 ppm (referència interna de desplaçament químic).
3. Filtrat de l'aigua residual amb l'algoritme HLSVD [10], entre els 4,3 i 5,1 ppm amb 10 lorentzianes.
4. Apoditzat de 1Hz, per disminuir el soroll i suavitzar el perfil dels pics.
5. Posteriorment, es van convertir mitjançant la DMS [34], a format .art cobrint el interval de 7,10 ppm a -2,7 ppm amb 512 punts, amb el propòsit d'obtenir vectors espectrals que continguin només la informació d'interès. A més, respecte al nombre de punts, la DMS realitza un procés de normalització a *unitlength*.

$$x' = \frac{x}{\sqrt{\sum x^2}}$$

on x és l'alçada de cada punt d'un espectre abans de normalitzar i x' després de normalitzar.

6. Els espectres obtinguts en tots els casos contenien 512 variables, corresponents a les alçades dels punts espectrals concrets, en el rang [7,1; -2,7] ppm. El rang de dades analitzat va ser només de 4,22 ppm a 0,59 ppm i les gràfiques d'espectres promig es van construir utilitzant l' EXCEL (*Microsoft Office 2007*).

4.4. Dades utilitzades: training set i test set

El *training set* estava constituït per una matriu disponible en el sistema de suport a la decisió INTERPRET versió 3.0.2 [36] i estava format per 217 espectres de pacients a temps d'eco curt (TEC), veure taula 3.

El *test set* estava constituït pels casos TEC de pacients de l'àrea de Barcelona recollits durant el projecte eTumour [19] i que eren totalment independents als casos utilitzats per l'entrenament. Una vegada optimitzat el classificador amb el *training set*, calia corroborar amb el *test set* si aquell classificador era idoni amb altres casos independents.

Per a començar, es varen fer agrupacions dels diferents tumors en classes d'interès tal i com s'havia fet en d'altres treballs anteriors que havien utilitzat la matriu de dades INTERPRET [10, 18-20, 33]. Classe 1: Glioblastoma multiforme (GBM) i metàstasis (ME); classe 2: Astrocitomes grau II (A2), Oligodendrogliomes (OD) i Oligoastrocitomes (OA); i classe 3: Meningiomes (MN).

| | TEC | GBM i ME | A2, OD, OA | MN | TOTAL |
|-----------------|---------------------|----------|------------|----|-------|
| | Training set | 124 | 35 | 58 | 217 |
| Test set | CDP | 20 | 15 | 2 | 37 |
| | IDI-Badalona | 20 | 5 | 2 | 27 |

Taula 3. Descripció dels *training i test set* utilitzats en aquest estudi.

4.5. Anàlisi de dades

A la Figura 3 es mostra l'esquema de l'anàlisi multivariant que es va dur a terme mitjançant els dos *softwares*, l'SPSS i l'SC.

Per l'anàlisi de les dades es va utilitzar l'LDA en ambdós *softwares*, podent així comparar les diferències alhora de trobar el millor classificador. El nombre de funcions discriminants que es podien obtenir estava limitat al nombre de classes, en el nostre cas en teníem 3, per tant, vàrem obtenir dues funcions discriminats.

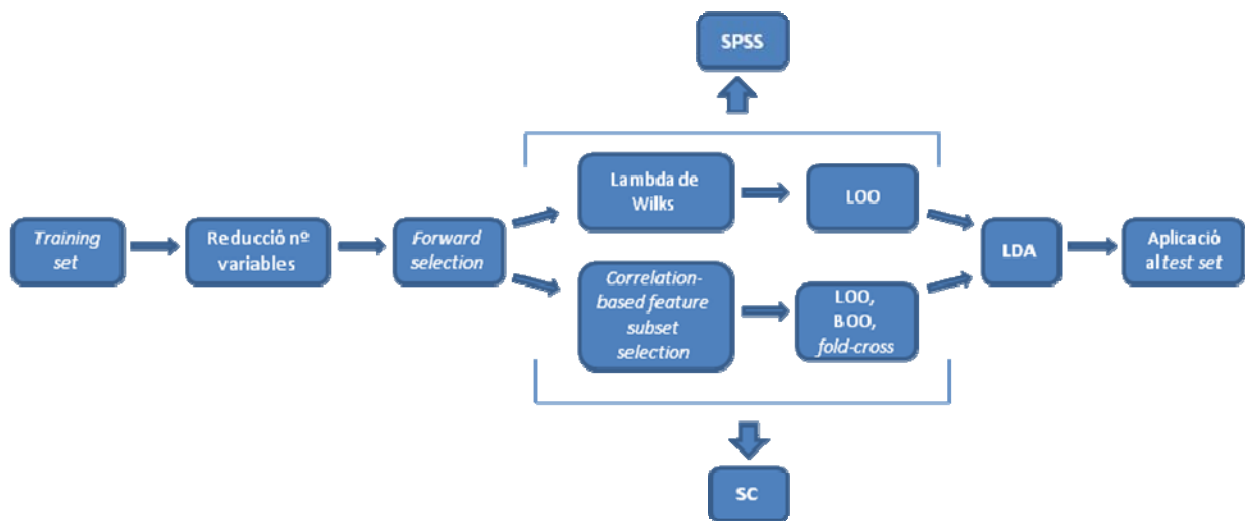


Figura 3. Esquema de l'anàlisi multivariant utilitzat en aquest treball ambdós *softwares*.

En primer lloc, es van desenvolupar classificadors per la nostra mostra amb l'SPSS (3 classes) amb el nombre de variables que el programa seleccionava automàticament sense intervenció de l'usuari. A continuació es va assignar el mateix nombre de variables a l'SC per veure els resultats de manera comparativa.

L'estratègia va consistir a continuació en disminuir el nombre de variables utilitzades per desenvolupar el classificador per ambdós *softwares* i estudiar les diferències obtingudes en l'encert dels classificadors al conjunt d'entrenament i del *test set* independent. Amb l'SC també vàrem augmentar el nombre de característiques en algun cas concret.

Finalment, es va estudiar com variava el percentatge d'encert respecte el nombre de variables utilitzades per a desenvolupar el classificador i el diferencial de l'encert *training set* i *test set*; és a dir, la diferència que hi havia entre l'encert obtingut amb les dades d'entrenament i l'encert obtingut amb els casos *test*.

4.5.1. SPSS

La selecció de variables es va fer amb la *forward selection* i utilitzant el criteri Lambda de Wilks [25] en el rang d'interès (4,22 a 0,59 ppm). Per a la classificació, es va triar l'opció de no ponderar el resultat per la grandària relativa dels diferents grups al conjunt d'entrenament fent servir la matriu de covariància intragrupal. I finalment, l'avaluació del grau d'encert del classificador es va dur a terme amb el mètode *Leave One Out* (LOO) per a comprovar l'eficàcia dels resultats sobre el grup *training set*. Podem veure l'esquema del procés a la Figura 3.

El procés de selecció de variables era automàtic, de manera que no es podia ajustar res. A continuació es va repetir l'anàlisi baixant el nombre de variables consecutivament fins a un mínim de dues.

4.5.2. SC

Com es pot observar a l'esquema de la Figura 3, amb l'SC la reducció de variables es va fer mitjançant *forward selection* però utilitzant un criteri diferent a l'SPSS, el conegut com *correlation-based feature subset selection* [26], ja que era el criteri implementat a l'SC. La classificació es va realitzar amb l'LDA en el rang espectral d'interès abans mencionat i per a l'avaluació es van poder triar entre tres mètodes que ens ofería el programa: LOO, BOO i *fold-crossvalidation*.

Partint del nombre de variables que ens havia seleccionat l'SPSS, se li va demanar a l'SC que seleccionés el mateix nombre, i finalment es va seguir el mateix procediment però baixant el nombre de variables fins a a dues.

5. RESULTATS

5.1. Variables significatives a TEC

Les variables més significatives seleccionades per l'SC i l'SPSS respectivament les veiem a la Taula 4.

| VARIABLES MÉS RELLEVANTS A TEC SEGONS L'SPSS I L'SC | | | | |
|---|------|-------------------------------------|------|-------------------------------------|
| | SPSS | possible metabòlit | SC | possible metabòlit |
| 1ra | 1,24 | lípid, lactat | 1,23 | lípid, lactat |
| 2ona | 2,30 | glutamat, glutamina, macromolècules | 2,19 | glutamat, glutamina, macromolècules |
| 3ra | 3,02 | creatina | 3,03 | creatina |
| 4a | 3,75 | glutamat, glutamina, alanina | 2,29 | glutamat, glutamina, macromolècules |
| 5a | 3,58 | myo-inositol, glicina | 3,63 | myo-inositol, glutamat, glutamina |
| 6a | 3,98 | creatina | 2,40 | glutamat, glutamina, macromolècules |
| 7a | - | - | 1,34 | lactat |
| 8a | - | - | 3,80 | glucosa |

Taula 4. Resum de les variables més rellevants a TEC seleccionades per l'SC i l'SPSS i el possible origen bioquímic de la variable seleccionada.

Per ordre de rellevància, els dos *softwares* seleccionaven els lípid-lactat, el conjunt glutamat-glutamina (Glx) més les macromolècules i la creatina com a senyals més rellevants per a classificar les mostres. A continuació l'SPSS escollia una altra ressonància del conjunt Glx, seguida pel ml-glicina i per un altre pic de creatina; en canvi, l'SC seleccionava un altre pic de la regió del Glx i les macromolècules, dos pics de Glx, el lactat i finalment la possible contribució del senyal de glucosa.

5.2. Espectres promig

Els espectres promig obtinguts dels casos *training set* i dels *test set* global de les tres classes d'interès utilitzats en aquest treball es mostren a la Figura 4.

A la Figura 5 en canvi es mostren els espectres promig de les dades *test* de cada centre per separat, del CDP i de l'IDI-Badalona respectivament.

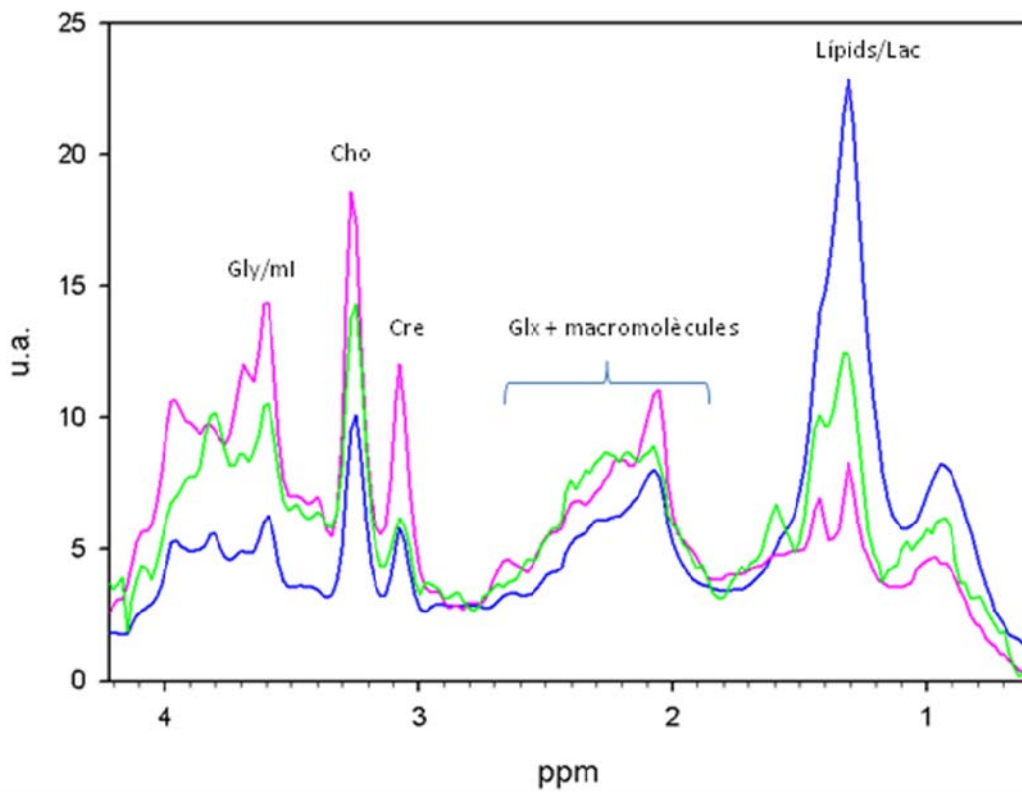
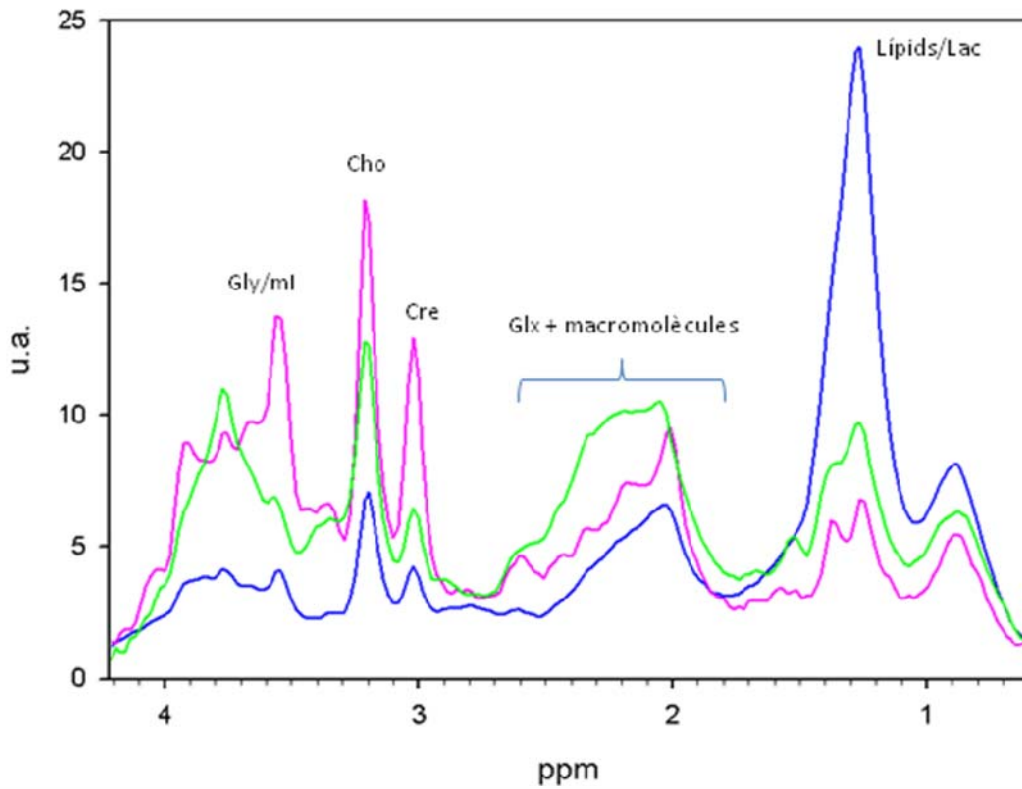


Figura 4. A dalt: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del training set. Abaix: Perfil espectroscòpic promig dels dos test set junts. En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN, u.a. fa referència a les unitats arbitràries.

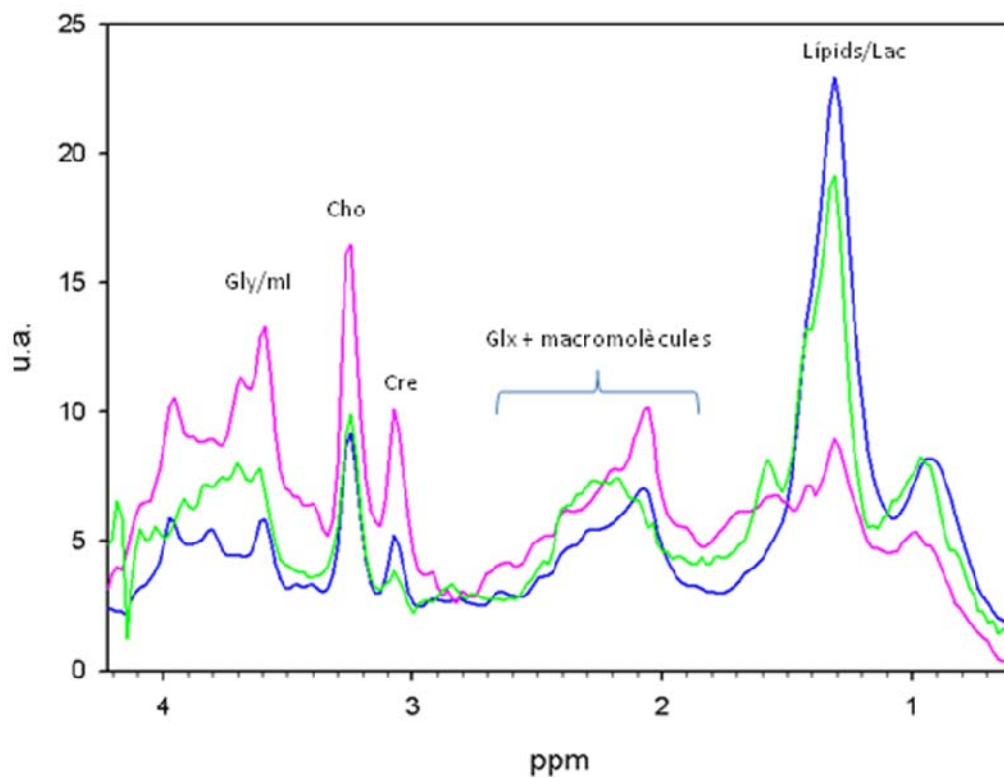
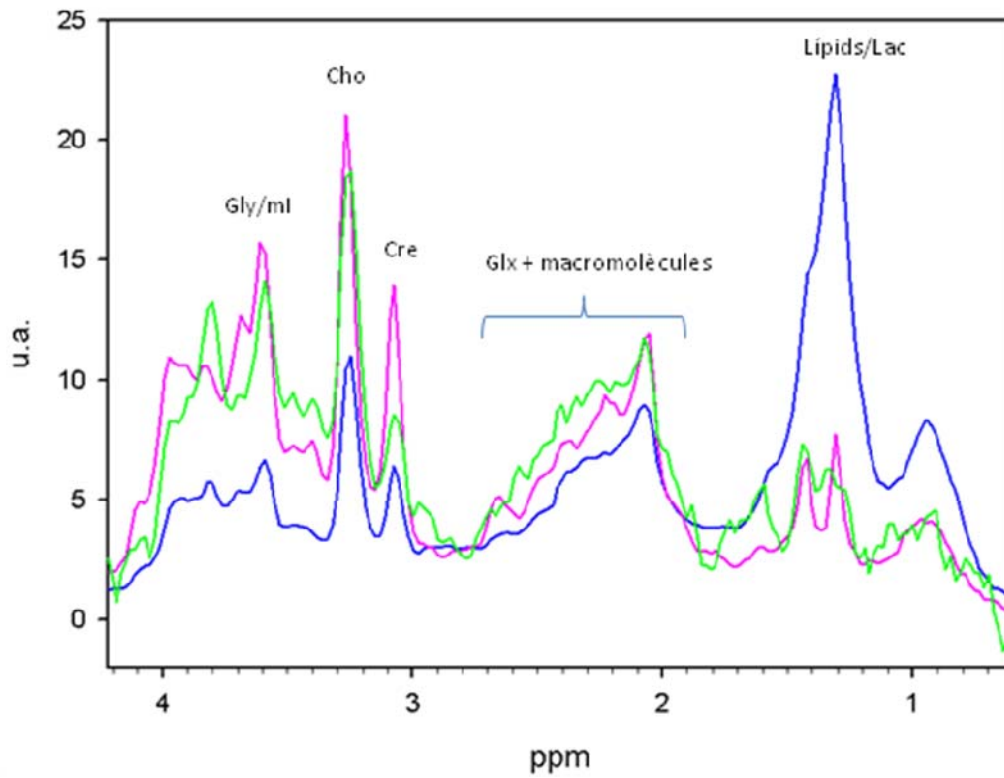


Figura 5. A dalt: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del test set del CDP. Abaix: Perfil espectroscòpic promig de les tres classes estudiades en el rang d'interès del test set de l'IDI-Badalona. En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN; u.a. fa referència a les unitats arbitràries.

Com es pot observar a figura 4, el perfil espectroscòpic del *training set* i del *test set* global era molt semblant, però quan ens centràvem en els centres individualment (Figura 5) és quan observàvem certes diferències. El més destacable va ser el perfil espectroscòpic dels meningiomes, on s'observava que les dades de l'IDI-Badalona diferien molt del patró esperat. En primer lloc, destacava un senyal de lípids, marcadors de necrosi, i amb una intensitat pròxima a la dels GBM i ME; mentre que els pics de Cho i Cre es trobaven significativament disminuïts, indicant malignitat.

5.3. Resultats de classificació i avaluació amb l'SPSS

L'SPSS va seleccionar un total de 17 variables a partir de les quals vàrem anar repetint la classificació treient-ne una cada vegada, per ordre de rellevància segons el criteri de la Lambda de Wilks. A la Figura 6 podem observar l'encert de la classificació obtingut amb l'SPSS en funció del nombre de variables utilitzat pel conjunt de *training* i de *test*.

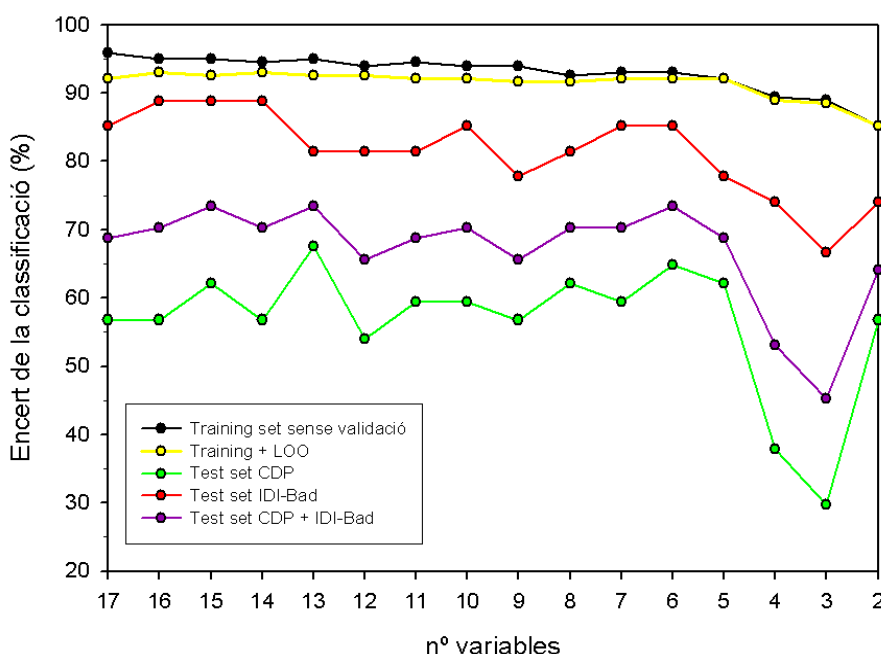


Figura 6. Encert de la classificació *versus* el nombre de variables seleccionades amb l'SPSS.

La classificació duta a terme amb l'SPSS i amb les dades de *training* va donar bons resultats, tan amb el nombre de variables que va escollir per defecte el *software* com quan es van anar baixant fins a dos; l'èxit de classificació es mantenia en el rang del 85 al 92% una vegada avaluat el sistema amb el mètode LOO. Una vegada es va passar a testar el classificador amb les dades independents, es va observar una baixada significativa de l'encert obtingut, amb uns resultats al voltant del 70% amb els dos grups de *test* junts o inclús del 45% amb només tres variables (30% en *test set* de l'IDI-Badalona).

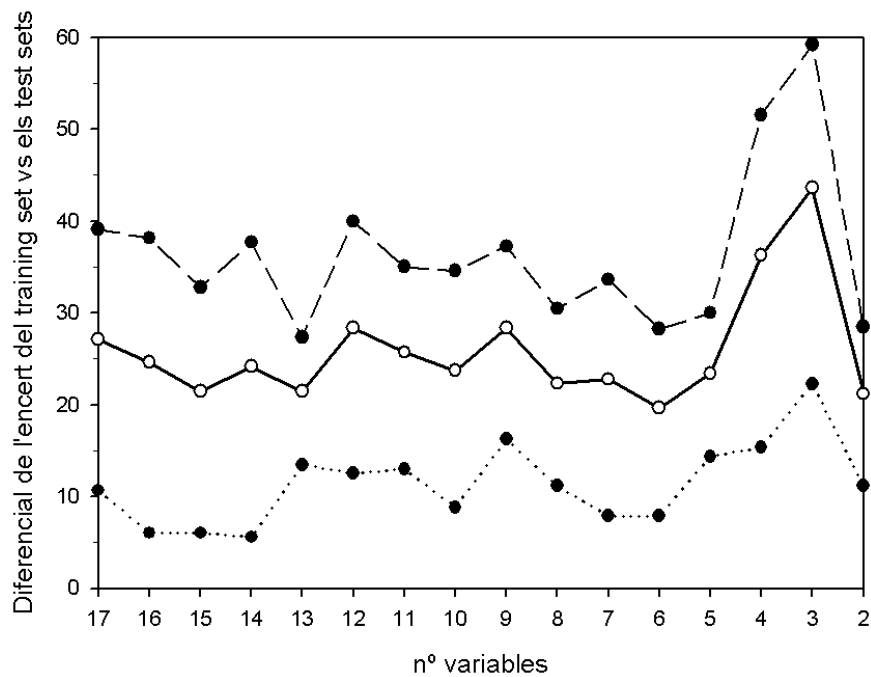


Figura 7. Diferencial de l'encert del *training set* versus el *test set* global i per separat fent servir l'SPSS. La línia contínua correspon al *test set* global, la discontinua al CDP i la línia de punts a l'IDI-Badalona.

Tal i com observem a la Figura 7, calculant el diferencial del percentatge d'encerts *training versus test set* global, es va obtenir una mitjana de 22-28 % de diferència i un 19,7 % en el millor dels casos utilitzant les 6 variables més rellevants per l'SPSS alhora de classificar. Respecte les dades del CDP, es van obtenir pitjors resultats amb grans diferències amb el *training set*. En canvi, el *test set* de l'IDI-Badalona es va classificar millor, obtenint un baix diferencial comparant amb el conjunt *training*.

5.4. Resultats de classificació i avaluació amb l'SC

L'encert de classificació respecte el nombre de variables obtingut amb l'SC es mostra a la Figura 8.

Els resultats de classificació amb l'SC van ser globalment millors que amb l'SPSS. En primer lloc, el classificador amb el grup de *training* es movia en el rang del 80-93 % d'encerts depenent del nombre de variables utilitzades una vegada validat el sistema; un percentatge lleugerament inferior al obtingut amb l'SPSS. Però, quan es va testar aquest classificador amb les noves dades del conjunt de *test* independent, es van aconseguir percentatges semblants als que s'obtenien amb el grup de *training*, al voltant del 80-84%, a diferència de l'SPSS, que no superava el 70%.

El diferencial de l'encert de classificació *training versus test* es mostra a la Figura 9:

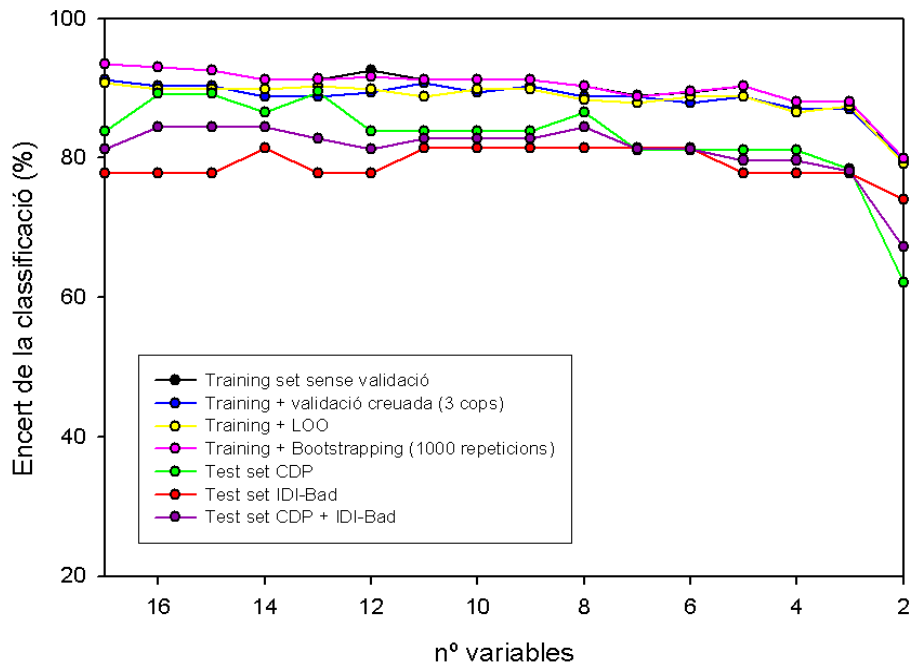


Figura 8. Encert de la classificació *versus* el nombre de variables seleccionades amb l'SC.

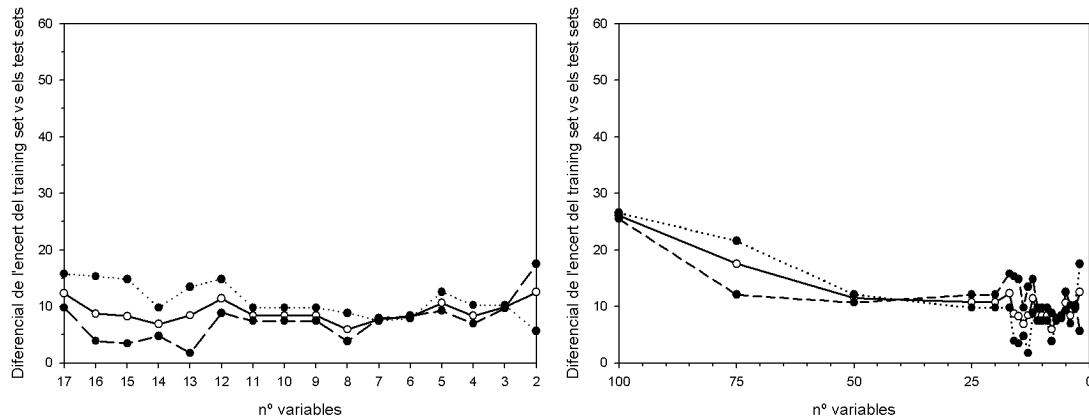


Figura 9. A l'esquerra: diferencial de l'encert del *training set* versus els *test sets* de 17 a 2 variables. La línia contínua correspon al *test set* global, la discontinua al CDP i la línia de punts a l'IDI-Badalonà. A la dreta: diferencial de l'encert del *training set* versus el conjunt *test set* de 100 a 2 variables.

Els percentatges del diferencial amb l'SC anaven del 6 al 12 %, resultats molt més propers entre el *training* i el *test*, obtenint el mínim del 5.9 % amb 8 variables en aquest cas.

A continuació es va estudiar el cas extrem utilitzant arbitràriament, 100 variables per a classificar amb l'SC, i així valoràvem el comportament del sistema en aquelles circumstàncies extremes afavoridores d'un sobreentrenament potencial del classificador. Es va observar que un nombre excessiu de característiques empitjorava clarament l'encert del classificador conseqüència d'aquest possible sobreentrenament; és a dir, tot i arribant al 99,5% de classificació amb el *training set* (resultat no mostrat), amb el *test set*, els percentatges d'encerts baixaven fortament per sota del 75% (Figura 9, dreta).

Així doncs, parlem de classificador òptim quan el diferencial *training-test* és mínim i amb el menor nombre de variables possible. Amb l'SPSS trobem l'òptim a les 6 variables amb un diferencial del 19.7%, mentre que amb l'SC el trobem amb 8 variables però un diferencial més petit, del 5.9%.

A la figura 10 i 11 es poden observar la distribució dels casos d'entrenament i els de *test* independents en l'espai latent definit per les dues equacions de discriminació del classificador desenvolupat, quan l'SC utilitzava el nombre òptim de variables per a desenvolupar aquest classificador.

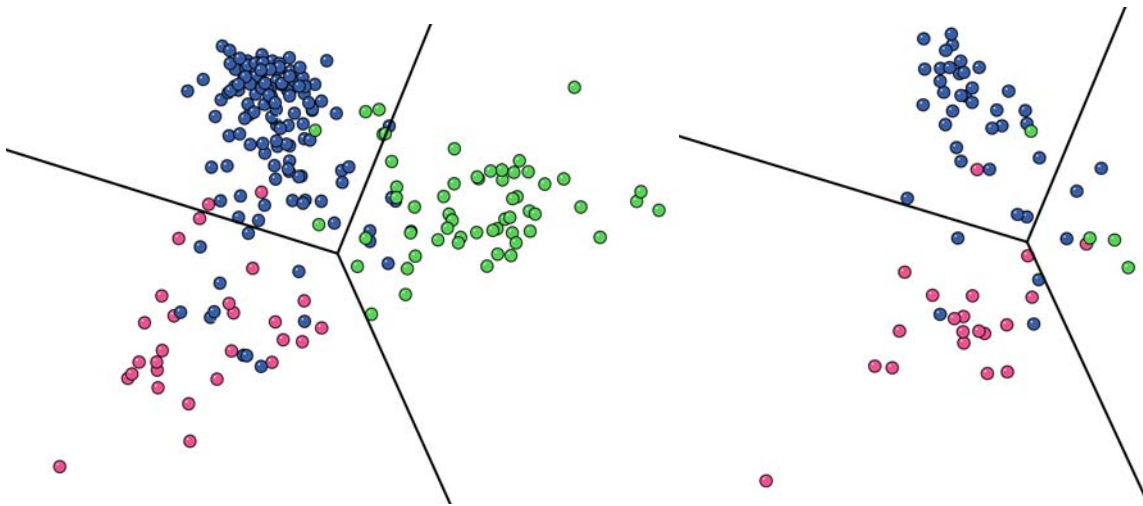


Figura 10. A l'esquerra: Distribució a l'espai de classificació de les classes estudiades del *training set* amb 8 variables, el nombre òptim (90,3% d'encert). A la dreta: Distribució del *test set* amb 8 variables (84,4% d'encert). En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN. Les línies rectes representen les fronteres entre les classes definides pel classificador.

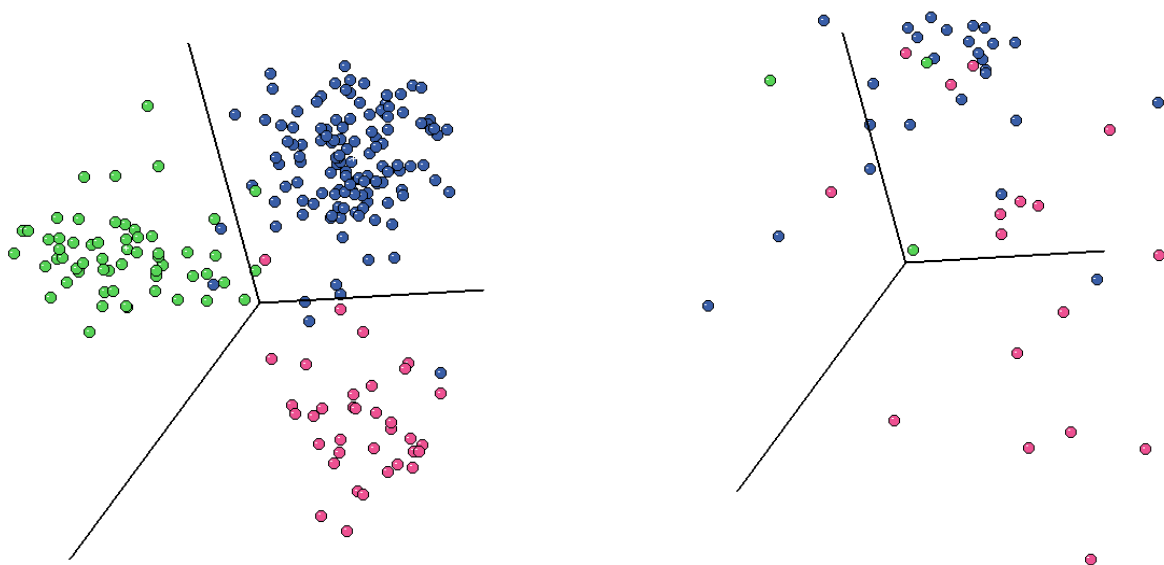


Figura 11. A l'esquerra: Distribució a l'espai de classificació de les classes estudiades del *training set* amb 100 variables (99,5% d'encert). A la dreta: Distribució del *test set* amb aquestes 100 variables (73,4% d'encert). En blau els GBM + ME, en rosa els A2 + OD + OA; i en verd els MN. Les línies rectes representen les fronteres entre les classes definides pel classificador.

El *training set* quedava ben classificat amb un 90,3% d'encert i d'igual manera que passava amb el *test set*, amb un 84,4%. En canvi, a la Figura 11 observem com el *training* quedava perfectament classificat amb 100 variables (99,5% d'encert), però quan introduïem les dades independents podíem observar com el sistema només en classificava correctament un 73,4%.

6. DISCUSSIÓ

En aquest treball s'ha avaluat l'ús de l'SC com a sistema de desenvolupament de classificadors d'espectres de ressonància magnètica nuclear de SV de pacients afectats amb tumors cerebrals, comparant els resultats amb l'ús de l'SPSS.

Les diferències obtingudes entre els dos *softwares* podrien explicar-se principalment pels criteris de selecció de variables que utilitza cada un (Figura 3). L'SPSS ha utilitzat el criteri de la Lambda de Wilks mentre que l'SC ha utilitzat el *correlation-based feature subset selection* [26], criteri potencialment molt més efectiu per seleccionar variables significatives en mostres de MRS de tumor cerebral com es pot veure en aquest treball.

Els metabòlits, possibles contribuents principals a les variables més rellevants, que ens seleccionaven els *softwares* coincidien alhora de triar les tres primeres, corresponents al senyal de lípids-lactat ($\approx 1,30$), Glx + macromolècules ($\approx 2,00-2,50$) i la Cre ($\approx 3,03$). A partir d'aquí el criteri utilitzat per cada un fa variar lleugerament l'ordre de selecció de la resta de metabòlits, ja que l'SPSS segueix amb un pic de ml-Glx ($\approx 3,60-3,80$), el ml-glicina ($\approx 3,55$) i finalment un altra pic de creatina ($\approx 3,93$); mentre que l'SC selecciona un nou senyal a la zona del Glx-Ala i altres compostos ($\approx 3,60-3,80$), un pic de Glx i macromolècules-lípids ($\approx 2,10-2,50$), el lactat ($\approx 1,33$) i el senyal de glucosa ($\approx 3,80$) com a últim metabòlit. També és important destacar que a l'SPSS només li calien 6 variables per obtenir el mínim d'encert òptim, a diferència de l'SC que en necessita 8, però tot i així, els percentatges d'èxit són significativament més alts amb l'SC, demostrant així la seva validesa pel desenvolupament de classificadors per MRS de tumors cerebrals.

Respecte els *test sets* hi ha varis punts interessants a comentar. En primer lloc, com s'observa a la Figura 5, els dos meningiomes de l'IDI-Badalona presentaven un perfil espectroscòpic promig que no corresponia a l'estàndard del tipus tumoral, així que ens vàrem centrar en l'estudi dels casos particulars d'aquests meningiomes. Tal i com es descriu a l'article de Vellido, *et al.* [37] i de García-Gómez *et al.* [19], poden aparèixer aquest tipus de conflictes amb els test sets utilitzats per testar classificadors. Així doncs, com va definir Vellido *et al.* [37], parlem de l'aparició de *artifact-related*

outliers o senyals artefactuals conseqüència de problemes en l'adquisició de les dades, *distinct outliers* referits a casos atípics no representatius del grup tumoral al qual pertanyen, o finalment de *class outliers* corresponents a casos clarament diagnosticats a un tipus tumoral concret per l'anatomia patològica, però que presenten un perfil espectroscòpic similar a altres tipus.

En el nostre cas, es va veure que el perfil espectroscòpic dels meningiomes de l'IDI-Badalona era el promig de només dos casos, els quals es tractaven, segons l'anatomia patològica, d'un meningioma atípic i un meningioma transicional [30]. L'espectre del primer diferia del patró esperat, alterant així el perfil promig del tipus tumoral; es podria doncs categoritzar com a *class outlier* [19, 37]. A la Figura 12 es mostra l'espectre d'aquest meningioma atípic *versus* el patró estàndard dels meningiomes i la imatge del vòxel d'on es va obtenir el perfil espectral.

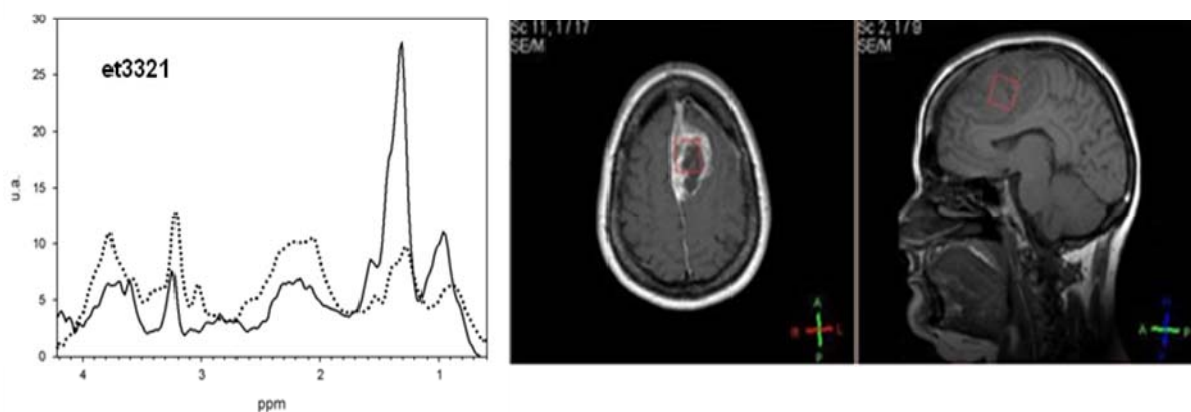


Figura 12. Espectre i imatge de referència d'un pacient (et3321 en codi de la base de dades eTUMOUR) amb meningioma atípic de l'IDI-Badalona. La línia continua representa el perfil espectroscòpic d'aquest cas i en discontinu observem el perfil promig dels meningiomes del *training set* (n=58).

L'espectre d'aquest cas no corresponia al promig dels meningiomes estàndards (veure Figures 4 i 5). Radiològicament, el cas estava orientat cap a tres diagnòstics diferents (meningioma, hemangiopericitoma i metàstasi); en canvi, histològicament els quatre patòlegs que van revisar la preparació de la biòpsia del cas van coincidir en que es tractava d'un meningioma atípic (grau II) tot i la necrosi observada. Això es podia justificar perquè el tumor era heterogeni, tal i com podem observar a la Figura 12, i perquè el vòxel estava situat sobre la zona amb necrosi. Per tant, la biòpsia va determinar clarament que la mostra estudiada es tractava d'un tumor de grau II, però possiblement es tractava en global d'un grau III o IV.

En segon lloc, es va observar que els *test sets* es comportaven de manera diferent segons el *software* utilitzat per desenvolupar el classificador a partir del conjunt d'entrenament. Alhora de testar el classificador, vàrem observar que

amb l'SPSS el percentatge d'encerts dels casos CDP amb el nombre òptim de variables va ser del 65%, mentre que el de l'IDI-Badalona era del 85% (veure Figura 6). En canvi, el mateix procediment amb l'SC, obteníem un 87% amb els casos del CDP i un 82% amb els de l'IDI-Badalona (Figura 8).

Per això, es van analitzar els casos mal classificats de manera consistent amb el *software* SPSS, és a dir, es va buscar quins eren els casos que encara que variéssim el nombre de variables utilitzades per a la classificació sempre es classificaven malament. Aquests casos es van analitzar un per un a la base de dades eTumour, i es va observar diferents fenòmens:

Al conjunt de dades de CDP, format per 37 casos, hi havia 15 casos mal classificats de manera consistent. Un cop analitzats, es va veure que en un 27% d'ells, els vòxels s'havien posicionat sobre àrees amb edema (Figura 13), en un altre 27% es tractava de tumors d'alt grau (GBM) que no tenien senyal de lípids (Figura 14), fenomen que ha estat descrit prèviament [10]. En un altre 40% dels casos (majoritàriament de patologia A2, OD, OA), els espectres presentaven un senyal a la regió de macromolècules (2,0-2,5 ppm) que era el pic més important de tot l'espectre (Figura 15), i en aquesta situació els casos eren classificats com a meningiomes.

Al conjunt de test de 27 casos d'IDI-Badalona, només n'hi havia 4 de mal classificats de manera consistent, i un 50% d'aquests corresponia a glioblastomes sense lípids.

A continuació es va analitzar com havia classificat l'SC aquests casos consistentment mal classificats en SPSS. Es va veure que l'SC classificava bé 10 casos d'un total de 15 dels que havien sigut mal classificats amb SPSS, i això corresponia als casos que presentaven el gran pic, que més aviat podia correspondre a la contribució de diferents substàncies donada la forma de "meseta" a la regió dels 2,0-2,5 ppm.

Com es pot observar a la Taula 4, l'SPSS no pren tantes variables en la regió dels 2,0-2,5 ppm; mentre que l'SC sí selecciona variables més rellevants en aquesta zona. Així doncs, els casos que presentaven un gran pic en aquesta regió (l'11% del *test set* del CDP), quedaven mal classificats per l'SPSS i per això es van trobar percentatges d'encerts més baixos amb SPSS i també això explicaria possiblement el diferent comportament dels *test set* amb els classificadors obtinguts en els dos programes. Al ser la proporció dels diferents tipus de raons per una mala classificació diferent entre els dos *test set*, depenent de quines avantatges o desavantatges amb respecte a aquestes causes proveïssin els classificadors produïts per un i altre programa el comportament seria diferent.

En resum, que una estratègia acurada de selecció de variables és vital per optimitzar el grau d'encert del classificador obtingut.

El primer tipus de casos que no es classificaven correctament trobats en aquest estudi van ser els tumors que presentaven teixit edematós (acumulació de líquid anormal) adjacent a la massa tumoral. El seu perfil espectroscòpic es caracteritzava per l'aparició d'un gran pic de Lac, marcador de metabolisme anaeròbic, i de Cho, indicant una alta proliferació cel·lular. Aquests casos es solien classificar malament amb l'SPSS, i quedaven millor discriminats amb l'SC.

A la Figura 13 veiem un exemple de tumor amb edema i el seu espectre.

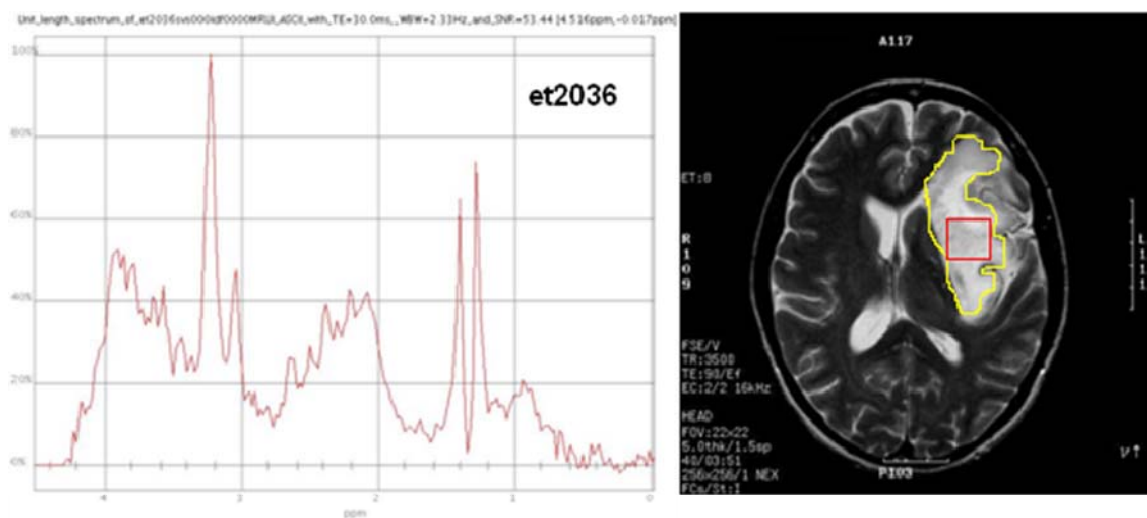


Figura 13. Espectre d'un tumor amb edema del CDP i la seva imatge de referència on està senyalat en groc la zona de l'edema i en vermell el vòxel per a l'enregistrament de la MRS.

Un altre exemple de *test set* mal classificat seria el GBM del cas "et2273", el qual presentava un gran pic de glicina-mi i un baix senyal de lípids; espectre típic d'un tumor astrocític de grau més baix. Podem comparar el seu patró amb el típic d'un GBM i d'un A2 com es mostra a la Figura 14. Aquest tipus de casos van quedar mal classificats tan per l'SPSS com per l'SC, considerant-se per tant, possibles *class outliers*.

Finalment, a la Figura 15 podem observar un exemple d'espectre que corresponia a un oligoastrocitoma però que quedava classificat com un meningioma degut al pic a la regió 2,0-2,5 ppm amb l'SPSS. Aquest mateix cas, en canvi, es classificava correctament per l'SC.

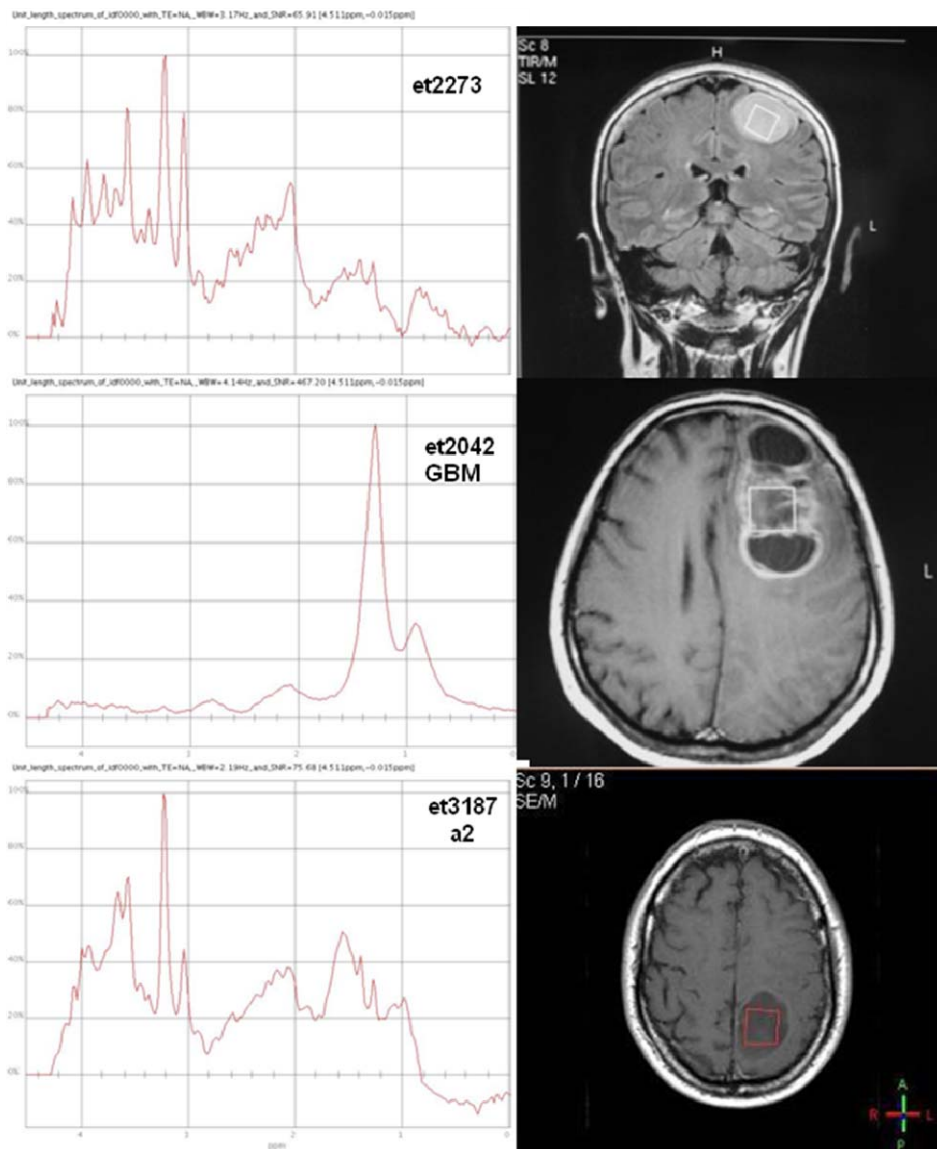


Figura 14. A dalt: exemple de GBM atípic mal classificat pels classificadors desenvolupats pels dos *softwares* amb la imatge de referència del vòxel a TEC. Al mig: espectre d'un GBM típic i la imatge de referència. A sota: espectre d'un A2 típic amb la imatge del vòxel.

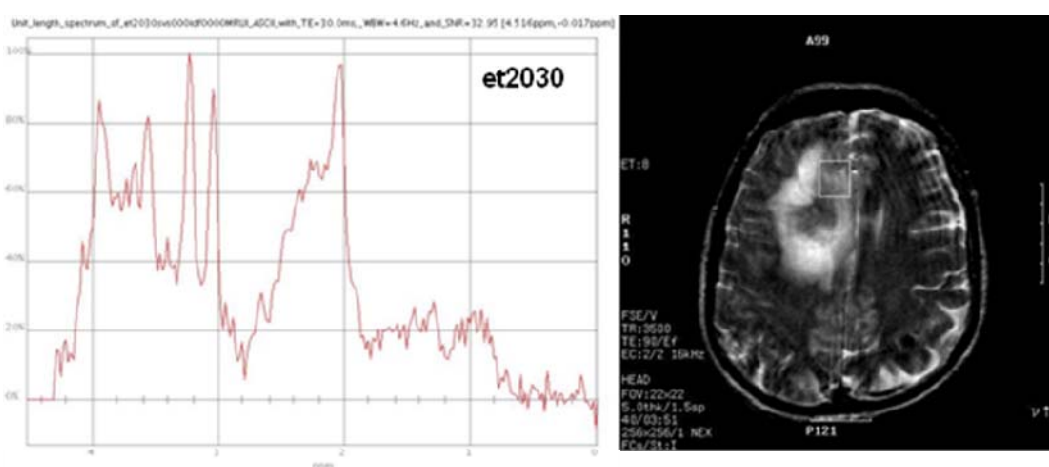


Figura 15. Espectre d'un OA que presentava una zona de senyal elevada entre els 2,0-2,5 ppm com els meningiomes i la seva imatge de referència del vòxel.

7. CONCLUSIONS

1. L'SC ens ha facilitat el desenvolupament de classificadors d'espectres SV ^1H ERM de tumors cerebrals humans basats en l'LDA, mitjançant una interfície senzilla, obtenint anàlisis semi-automatitzats de reconeixement de patró de diferents tipus tumorals.
2. Comparant els resultats de classificació obtinguts amb l'SPSS i l'SC s'ha comprovat que les eines matemàtiques de selecció de variables del vector de dades espectrals que incorpora l'SC són molt més robustes alhora de discriminar pacients amb tumors cerebrals explorats amb SV ^1H MRS.
3. S'ha comprovat que després d'un exercici de classificació, per molt automatitzat que sigui, sempre cal tornar a les fonts de les dades per revisar possibles inconsistències, artefactes o presència d'*outliers*.
4. Podem concloure afirmant que en les nostres mans l'SC és un programa robust pel desenvolupament de classificadors de SV ^1H MRS *in-vivo* de tumors cerebrals humans.

8. BIBLIOGRAFIA

1. Gadian, D.G., *Nuclear magnetic resonance and its applications to living systems*. 1982: Oxford Science Publications.
2. Westbrook, C., *MRI at a Glance*, ed. B. Science. 2007.
3. <http://hi-tec.tripod.com/OFFLINE/nmr.pdf>.
4. Simões, R.V., et al., *Perturbation of mouse glioma MRS pattern by induced acute hyperglycemia*. *NMR in Biomedicine*, 2008. **21**(3): p. 251-264.
5. Moonen CT, v.K.M., van Zijl PC, Cohen J, Gillen J, Daly P, Wolf G., *Comparison of single-shot localization methods (STEAM and PRESS) for in vivo proton NMR spectroscopy*. *NMR Biomed.* , 1989. **2**(5-6): p. 201-8.
6. Gili, J., *Apuntes del curso de doctorado de la UAB sobre "Introducción Biofísica a la Resonancia Magnética Aplicada a la Clínica"*. 2008.
7. Isabella M. Burtscher, M., *Proton Magnetic Resonance Spectroscopy in Brain Tumours: Clinical Applications*, in *Department of Diagnostic Radiology*. 1999, Lund University Hospital: Lund, Sweden.
8. De Graaf, R.A., *In vivo NMR Spectroscopy: Principles and Techniques*, ed. Wiley. 1998.
9. Majós, C., et al., *Utility of proton MR spectroscopy in the diagnosis of radiologically atypical intracranial meningiomas*. *Neuroradiology*, 2003. **45**(3): p. 129-136.
10. Tate, A., et al., *Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra*. *NMR in Biomedicine*, 2006. **19**(4): p. 411 - 434.
11. Rémy, C., et al., *Evidence That Mobile Lipids Detected in Rat Brain Glioma by 1H Nuclear Magnetic Resonance Correspond to Lipid Droplets*. *Cancer Research*, 1997. **57**(3): p. 407-414.
12. Gill SS, T.D., Van Bruggen N, Gadian DG, Peden CJ, Bell JD, Cox IJ, Menon DK, Iles RA, Bryant DJ, et al., *Proton MR spectroscopy of intracranial tumours: in vivo and in vitro studies*. *J Comput Assist Tomogr*. 1990. **14**(4): p. 497-504.
13. Bernstein, M.B., M.S., *Neuro-oncology: The Essentials*. 2nd edition ed. 2008: Thieme.
14. David N. Louis, H.O., Otmar D. Wiestler, Webster K. Cavenee, ed. *WHO Classification of Tumours of the Central Nervous System*. 4th ed. 2007, International Agency for Research on Cancer (IARC): Lyon, France.
15. <http://www.cbtrus.org/reports/reports.html>.
16. Graziella Filippini, C.F., Amerigo Boiardi, Giovanni Broggi, Maria G. Bruzzone, Dario Caldiroli, Rita Farina, Mariangela Farinotti, Laura Fariselli, Gaetano Finocchiaro, Sergio Giombini, Bianca Pollo, Mario Savoiano, Carlo L. Solero, and Maria G. Valsecchi, *Prognostic factors for survival in 676 consecutive patients with newly diagnosed primary glioblastoma*. *Neuro-Oncology*, 2008. **10**(1): p. 78-87.
17. Tugcu B, P.L., Gunaldi O, Tanriverdi O, Akdemir H., *Efficacy of clinical prognostic factors on survival in patients with glioblastoma*. *Turk Neurosurg.*, 2010. **20**(2): p. 117-25.
18. Lukas, L., et al., *Brain tumor classification based on long echo proton MRS signals*. *Artificial Intelligence in Medicine*, 2004. **31**(1): p. 73-89.
19. García-Gómez, J., et al., *Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy*. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2009. **22**(1): p. 5-18.
20. Devos, A., et al., *Classification of brain tumours using short echo time 1H MR spectra*. *Journal of Magnetic Resonance*, 2004. **170**(1): p. 164 - 175.
21. Ng HK, T.C., Lo ST., *Meningiomas and arachnoid cells: an immunohistochemical study of epithelial markers*. *Pathology*, 1987. **19**(3): p. 253-7.
22. <http://www.spss.com/es/>.
23. Ortega-Martorell, S., et al., *SpectraClassifier 1.0: a user friendly, automated MRS-based classifier-development system*. *BMC Bioinformatics*, 2010. **11**(1): p. 106. <http://gabrmn.uab.es/sc>.
24. Lozares, C.L.P., *Anàlisi multivariable de dades estadístiques*, ed. U.A. Barcelona. 2000.

25. Álvarez Cáceres, R., *Estadística multivariante y no paramétrica con SPSS: aplicación a las ciencias de la salud*, ed. D.d. Santos. 1995.
26. MA., H., *Correlation-based feature subset selection for machine learning*. 1999, University of Waikato: Hamilton, New Zealand.
27. Duda, R.O., Hart PE, Stork, DG., *Pattern Classification*, ed. Second. 2001, New York: John Wiley & Sons.
28. Ortega-Martorell, S., et al., *SpectraClassifier 1.0: a user friendly, automated MRS-based classifier-development system*. BMC Bioinformatics, 2010. **11**: p. 106.
29. <http://gabrmn.uab.es/interpret/index.html>.
30. <http://lithium.uab.es:8145/eTumourFrozen/>. Base de dades projecte eTumour.
31. <http://azizu.uab.es:8120/iFrozenDB>. Base de dades projecte INTERPRET.
32. Julià-Sapé, M., et al., *A Multi-Centre, Web-Accessible and Quality Control-Checked Database of in vivo MR Spectra of Brain Tumour Patients*. Magnetic Resonance Materials in Physics, Biology and Medicine, 2006. **19**(1): p. 22-33.
33. Tate, A.R., et al., *Automated classification of short echo time in In Vivo ¹H brain tumor spectra: A multicenter study*. Magnetic Resonance in Medicine, 2003. **49**(1): p. 29-36.
34. <http://azizu.uab.es/dms>.
35. Klose, U., *In vivo proton spectroscopy in presence of eddy currents*. Magn Reson Med., 1990. **14**(1): p. 26-30.
36. <http://azizu.uab.es/dss>.
37. Alfredo, V., et al., *Outlier exploration and diagnostic classification of a multi-centre 1H-MRS brain tumour database*. Neurocomput., 2009. **72**(13-15): p. 3085-3097.