

Universitat Politècnica de Catalunya

Departament de Matemàtiques

Programa de Doctorat en Enginyeria Biomèdica

Caracterització morfològica de cèl·lules limfoides normals, reactives, anormals i blàstiques de sang perifèrica mitjançant processament digital d'imatges

escrita per

Laura Puigví Fernández

Tesi presentada per obtenir el títol de
"Doctora en Enginyeria Biomèdica"
per la Universitat Politècnica de Catalunya

Directors de la tesi

Dr. José Rodellar

Departament de Matemàtiques, Universitat Politècnica de Catalunya

Dra. Anna Merino

Unitat de Citologia, Laboratori Core, Hospital Clínic de Barcelona



Barcelona. Setembre, 2019

“Science is what we understand well enough to explain to a computer. Art is everything else we do.”

– Donald E. Knuth

Pròleg

És de justícia que comenci pels meus directors de tesi, el Dr. José Rodellar i la Dra. Anna Merino per haver-me donat l'oportunitat de poder realitzar la present Tesi Doctoral així com tot el seu suport i supervisió durant el desenvolupament de la recerca. És obvi que aquesta Tesi no hauria estat possible sense ells.

Agraeixo, igualment, el suport científic-econòmic rebut per part del grup de recerca CoDALab (Control, Modelització, Identificació i Aplicacions), especialment al doctor Santiago Alférez per la seva contribució en el processament digital d'imatges així com a la doctoranda Andrea Acevedo per ajudar en la creació de la base de dades i el suport econòmic de l'Universitat Politècnica de Catalunya a través d'un contracte predoctoral.

Estic molt agraïda a la Unitat de Citologia del Laboratori Core de l'Hospital Clínic de Barcelona per poder disposar de les seves instal·lacions per tal d'obtenir un nombre considerable d'imatges de cèl·lules així com l'accés a la informació clínica dels pacients inclosos en la recerca. Entre elles donar les gràcies especialment al José Luis Bedini.

També m'agradaria expressar el meu agraïment a tota la meva família, especialment a la meva mare Beatriz i al David, per animar-me sempre a seguir saltant els obstacles i treure les parts positives de les crítiques.

M'agradaria mostrar el meu agraïment a totes aquelles persones que durant el procés de recerca m'han donat suport constantment, especialment la Laura i a la recta final, la Marina.

Finalment, demano disculpes per endavant per no recordar a les línies anteriors a gent que realment també s'ho mereixen.

Abstract

The main objective of the present Doctoral Thesis is to obtain new quantitative features by means of digital image processing and machine learning for the differentiation of normal, reactive and malignant lymphoid cells of peripheral blood, contributing to an objective morphologic assessment.

The research addresses the following two issues:

- (1) Using machine learning, geometric, colour and texture descriptors are searched which have an explicit quantitative formulation and a reasonable qualitative interpretation in visual morphologic terms.
- (2) Considering cellular abnormalities established a priori, associated with specific diseases, the aim is to identify specific quantitative descriptors of morphological characteristics that cytologists recognize visually and usually express subjectively.

More than 200 patients and 16 different lymphoid cell groups have been included in the research. Using the CellaVision DM96, 12,000 images have been acquired and using the microscope Olympus BX43, 9,000. Almost 2,700 features including geometric, colour and texture (first and second-order statistics, granulometric, Wavelet and Gabor) have been analysed for lymphoid cell differentiation. Six colour spaces have been considered: RGB, CMYK, HSV, XYZ, Lab and Luv.

The 20 most efficient features for the differentiation between reactive lymphoid cells (infections) and neoplastic cells (abnormal lymphocytes in lymphoma or lymphoid blasts in acute leukaemia) have been analysed. The three most relevant descriptors for the recognition of the 12 lymphoid cell groups considered are geometric: 1) *nucleus/cytoplasm ratio*, 2) *nuclear perimeter* and 3) *cell diameter*. Most of the 20 descriptors show significant differences between pairs of abnormal lymphocytes which are difficult to recognize by morphology. Five colour and texture features are significant to discriminate reactive lymphocytes from abnormal lymphoid cells.

Quantitative descriptors have been identified for the detection of specific cell morphologic abnormalities of certain lymphoid neoplasms, which have shown good specificity and sensitivity using the two different image acquisition systems. Regarding nuclear abnormalities, the detection of mature and condensed chromatin seen in chronic lymphatic leukaemia cells has been achieved by the *correlation of the cyan of the nucleus*, and the cerebriform chromatin characteristic of Sézary cells, by means of the *standard deviation of the granulometric curve of the cyan component of the nucleus*.

Regarding the cytoplasm, *hairiness* descriptor has been able to detect cytoplasmic villi present in villous lymphocytes in hairy cell leukaemia and splenic marginal zone lymphoma. The *skewness of the histogram of the u component of the cytoplasm* has shown to be useful for detecting azurophilic cytoplasmic granules seen in abnormal lymphocytes in T-cell large granular lymphocytic leukaemia.

The results of this Doctoral Thesis provide objectivity in the morphologic assessment of normal, reactive and neoplastic lymphoid cells. Obtaining quantitative descriptors for abnormal lymphoid cells, which are specific to certain lymphoid neoplasms with peripheral blood expression, could facilitate their detection. Haematological analysers based on digital image analysis could benefit from the use of quantitative descriptors, such as those described herein, in order to discriminate between reactive and neoplastic lymphoid cells.

Keywords: *Computer-based image analysis; Blood morphology; Hematopathology.*

Resum

L'objectiu general de la present Tesi és obtenir nous descriptors quantitius mitjançant processament digital d'imatges i aprenentatge automàtic per a la diferenciació de cèl·lules limfoides normals, reactives i malignes de sang perifèrica, contribuint a una anàlisi objectiva de la citologia sanguínia.

La recerca s'ha enfocat des de dues perspectives:

(1) Partint de l'aprenentatge automàtic, s'han buscat descriptors geomètrics, de color i de textura que tinguin una formulació quantitativa explícita i una interpretació qualitativa raonable en termes morfològics visuals.

(2) Partint d'anormalitats cel·lulars establertes a priori, associades a malalties específiques, l'objectiu és identificar descriptors quantitius específics de característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius.

S'han inclòs més de 200 pacients i 16 grups cel·lulars limfoides. Amb el sistema CellaVision DM96 s'han adquirit 12.000 imatges i amb el microscopi Olympus BX43, 9.000. Per a la diferenciació cel·lular s'han analitzat 2.700 descriptors geomètrics, de color i de textura (estadístics de primer i segon ordre, granulomètrics, Wavelet i Gabor). S'han considerat sis espais de color (RGB, CMGN, HSV, XYZ, Lab i Luv).

S'han analitzat els 20 descriptors geomètrics, de color i de textura més eficients per a la diferenciació entre cèl·lules limfoides reactives (infeccions) i neoplàsiques (anormals als limfomes o blasts limfoides a les leucèmies agudes). Els descriptors més rellevants per al reconeixement dels 12 grups cel·lulars limfoides són geomètrics: 1) relació nucli/citoplasma, 2) perímetre del nucli i 3) diàmetre de la cèl·lula. La majoria dels 20 descriptors mostren diferències significatives entre parelles de limfòcits anormals de difícil reconeixement per morfologia. Cinc descriptors de color i textura són significatius per discriminar els limfòcits reactius dels anormals.

S'han identificat descriptors quantitativs per a la detecció d'anomalies morfològiques específiques de cèl·lules de determinades neoplàsies limfoides, que han mostrat una bona especificitat i sensibilitat amb els dos sistemes diferents d'adquisició d'imatges. En relació a anormalitats nuclears, la detecció de la cromatina madura i condensada de les cèl·lules de la leucèmia limfàtica crònica ha estat possible mitjançant la *correlació del cian del nucli*, i de la cromatina cerebriforme característica de les cèl·lules de Sézary mitjançant la *desviació estàndard de la corba granulomètrica del component cian del nucli*.

Pel que fa al citoplasma, s'han detectat les prolongacions citoplasmàtiques dels limfòcits de la tricoleucèmia i del limfoma de la zona marginal esplènica mitjançant el descriptor *hairiness*. *L'asimetria de l'histograma del component u del citoplasma* ha demostrat ser útil per detectar els grànuls azuròfils dels limfòcits grans granulars.

Els resultats de la present Tesi proporcionen objectivitat en l'avaluació morfològica de cèl·lules limfoides normals, reactives i neoplàsiques. L'obtenció de descriptors quantitativs per cèl·lules limfoides anormals específiques de determinades neoplàsies limfoides amb expressió a sang perifèrica podria facilitar la seva detecció. Els analitzadors hematològics basats en l'anàlisi digital d'imatges es podrien beneficiar de l'ús de descriptors quantitativs, com els descrits, per tal de discriminar entre cèl·lules limfoides reactives i neoplàsiques.

Paraules clau: Anàlisi d'imatges assistit per ordinador; Morfologia sanguínia; Hematopatologia.

Taula de continguts

Llistat de taules.....	17
Llistat d'il·lustracions.....	19
Llistat d'abreviatures	25
1 Motivació i objectius	29
1.1 Introducció	29
1.2 Motivació	31
1.3 Estat de l'art	34
1.3.1 Descriptors geomètrics	34
1.3.2 Descriptors de color i de textura.....	35
1.3.3 Reconeixement de cèl·lules limfoides anormals	36
1.4 Objectius	38
1.5 Organització de la Tesi i contribucions.....	40
2 Fonaments biomèdics del reconeixement de cèl·lules sanguínies.....	43
2.1 Introducció	43
2.2 Hematopoesi.....	44
2.3 Anàlisi morfològica de la sang perifèrica	46
2.3.1 Frotis de sang perifèrica	46
2.3.2 Fixació.....	47
2.3.3 Tinció amb May Grünwald-Giemsa.....	47
2.3.4 Observació al microscopi	48
2.3.5 Cèl·lules benignes de la sang perifèrica	49
2.3.6 Rellevància del frotis de sang perifèrica en el diagnòstic de les neoplàsies limfoides.....	52
2.4 Mètodes d'anàlisi cel·lular automatitzada	53
2.5 Sistemes automatitzats de morfologia digital	57
2.6 Neoplàsies limfoides i classificació	62
2.6.1 Leucèmia limfoblàstica aguda.....	64
2.6.2 Leucèmia limfàtica crònica	64
2.6.3 Leucèmia prolimfocítica B	65
2.6.4 Tricoleucèmia.....	65
2.6.5 Limfoma de la zona marginal esplènica.....	66
2.6.6 Limfoma fol·licular	66
2.6.7 Limfoma de cèl·lules del mantell.....	66
2.6.8 Mieloma múltiple amb cèl·lules plasmàtiques circulant a sang perifèrica.....	66
2.6.9 Leucèmia prolimfocítica T	67
2.6.10 Síndrome de Sézary.....	67
2.6.11 Leucèmia de limfòcits grans granulars tipus T.....	67
3 Metodologia de recerca	69

3.1	Introducció	69
3.2	Adquisició i emmagatzematge de les imatges.....	71
3.2.1	Grups de cèl·lules d'interès	71
3.2.2	Sistemes d'adquisició de les imatges digitals.....	73
3.2.3	Creació de la base de dades	78
3.3	Algoritme de segmentació.....	80
3.3.1	Preprocessament.....	80
3.3.2	Segmentació	82
3.3.3	Postprocessament	84
3.4	Extracció de descriptors	85
3.4.1	Descriptors geomètrics	86
3.4.1.1	Àrea	87
3.4.1.2	Rati nucli/citoplasma	87
3.4.1.3	Rati nucli/cèl·lula	87
3.4.1.4	Perímetre	87
3.4.1.5	Hairiness.....	88
3.4.1.6	Diàmetre equivalent	89
3.4.1.7	Circularitat.....	89
3.4.1.8	Àrea convexa.....	89
3.4.1.9	Excentricitat cònica	90
3.4.1.10	Elongació.....	90
3.4.1.11	Solidesa	90
3.4.1.12	Bounding box	90
3.4.1.13	Extensió.....	90
3.4.1.14	Convex hull	91
3.4.2	Descriptors de color.....	91
3.4.2.1	Espais de color.....	93
3.4.2.2	Descriptors de primer ordre basats en l'histograma	98
3.4.3	Descriptors de textura.....	101
3.4.3.1	El concepte de textura	101
3.4.3.2	Descriptors de segon ordre basats en la matriu de co-ocurrència.....	102
3.4.3.3	Descriptors granulomètrics.....	108
3.4.3.4	Altres descriptors.....	112
3.5	Classificació.....	113
3.6	Anàlisi estadística dels descriptors.....	115
3.6.1	Conceptes bàsics.....	115
3.6.2	Estadística descriptiva	116
3.6.2.1	Variables quantitatives	116
3.6.2.2	Variables qualitatives	118
3.6.3	Contrast d'hipòtesi	119

3.6.4	Normalitat.....	120
3.6.5	Homoscedasticitat	121
3.6.6	Comparació de dues mostres	122
3.6.7	Comparació múltiple	123
3.6.8	Correlació.....	124
3.6.9	Avaluació de l'eficiència diagnòstica d'un test.....	125
4	Descriptors quantitius per a la diferenciació morfològica de cèl·lules limfoïdes anormals a partir d'imatges de sang perifèrica adquirides amb el sistema DM96	129
4.1	Introducció	131
4.2	Material i mètodes.....	133
4.2.1	Obtenció dels pacients i les mostres.....	133
4.2.2	Adquisició de les imatges digitals	135
4.2.3	Processament digital de les imatges	136
4.2.4	Extracció de descriptors	136
4.2.4.1	Descriptors geomètrics	136
4.2.4.2	Descriptors de color i textura	136
4.2.5	Anàlisi estadística dels descriptors.....	140
4.3	Resultats.....	141
4.3.1	Descriptors geomètrics	144
4.3.2	Descriptors de color i textura	148
4.4	Discussió.....	164
4.5	Conclusions.....	167
5	Descriptors citològics quantitius per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, les cèl·lules de Sézary, els limfòcits vellosos i els limfòcits grans granulars mitjançant l'anàlisi d'imatges utilitzant el sistema DM96.....	169
5.1	Introducció.....	171
5.2	Material i mètodes.....	173
5.2.1	Obtenció dels descriptors citològics quantitius i les corresponents escales numèriques	173
5.2.2	Avaluació dels descriptors quantitius per a la classificació de frotis	178
5.3	Resultats.....	182
5.3.1	Determinació dels descriptors citològics quantitius i les corresponents escales numèriques	182
5.3.2	Avaluació dels descriptors quantitius per a la classificació de frotis	192
5.4	Discussió.....	197
5.5	Conclusions.....	201
6	Descriptors quantitius per a la diferenciació morfològica de cèl·lules limfoïdes anormals a partir d'imatges de sang perifèrica adquirides amb un microscopi convencional (sistema BX43)	203
6.1	Introducció	205
6.2	Material i mètodes.....	206
6.2.1	Obtenció dels pacients i les mostres.....	206
6.2.2	Adquisició de les imatges digitals	208
6.2.3	Processament de les imatges digitals.....	208

6.2.4	Anàlisi estadística dels descriptors	208
6.3	Resultats	209
6.3.1	Reformulació del descriptor <i>hairiness</i>	212
6.4	Discussió	215
6.5	Conclusions	217
7	Descriptors quantitativs citològics per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, les cèl·lules de Sézary, els limfòcits vellosos i els limfòcits grans granulars mitjançant l'anàlisi d'imatges utilitzant el sistema BX43.....	219
7.1	Introducció	221
7.2	Material i mètodes.....	222
7.2.1	Càlculs dels descriptors quantitativs citològics i les corresponents escales numèriques ..	222
7.2.2	Avaluació dels descriptors quantitativs per a la classificació de frotis	225
7.3	Resultats	228
7.3.1	Correlació dels descriptors i les corresponents escales numèriques.....	228
7.3.2	Avaluació dels descriptors quantitativs per a la classificació de frotis	233
7.4	Discussió	238
7.5	Conclusions	240
8	Conclusions i contribucions de la present Tesi Doctoral. Perspectives de futur	241
8.1	Conclusions	242
8.1.1	Conclusions derivades de la primera publicació	242
8.1.2	Conclusions derivades de la segona publicació.....	243
8.1.3	Conclusions derivades de l'estudi comparatiu dels dos sistemes d'adquisició d'imatges	244
8.1.4	Consideracions finals	245
8.2	Contribucions de la tesi.....	246
8.2.1	Articles	246
8.2.2	Conferències i comunicacions	246
8.2.3	Premis.....	248
8.3	Perspectives de futur	249
Referències	251
Glossari	259
9	Annexos.....	262
9.1	Annex A	263
9.1.1	<i>Relació nucli/citoplasma</i>	266
9.1.2	<i>Perímetre nuclear</i>	268
9.1.3	<i>Diàmetre equivalent de la cèl·lula</i>	269
9.1.4	<i>Curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula</i>	270
9.1.5	<i>Asimetria de l'histograma del blau de la cèl·lula</i>	271
9.1.6	<i>Information measure of correlation 1 del magenta de la cèl·lula</i>	272
9.1.7	<i>Desviació estàndard de l'histograma del component blau de la cèl·lula</i>	273
9.1.8	<i>Entropia de l'histograma del component magenta de la regió de la cèl·lula</i>	274

9.1.9	<i>Asimetria de l'histograma del component verd de la cèl·lula</i>	275
9.1.10	<i>Mitjana de la corba pseudo-granulomètrica del negre del citoplasma</i>	276
9.1.11	<i>Mitjana de l'histograma del component blau del nucli</i>	277
9.1.12	<i>Cluster shade de la saturació de la cèl·lula</i>	278
9.1.13	<i>Mitjana de l'histograma del magenta del citoplasma</i>	279
9.1.14	<i>Homogeneïtat del component magenta de la cèl·lula</i>	280
9.1.15	<i>Curtosi de l'histograma de la saturació de la cèl·lula</i>	281
9.1.16	<i>Information measure of correlation 1 del cian del nucli</i>	282
9.1.17	<i>Hairiness</i>	283
9.1.18	<i>Desviació estàndard de l'histograma de la saturació de la cèl·lula</i>	284
9.1.19	<i>Circularitat nuclear</i>	285
9.1.20	<i>Homogeneïtat del negre del citoplasma</i>	286
9.2	Annex B	287

Llistat de taules

Taula 4.1: Distribució dels diferents grups de cèl·lules limfoides, el nombre de pacients i imatges incloses a l'estudi.	134
Taula 4.2: Conjunt de 2.676 descriptors que s'han extret en la metodologia descrita en aquest capítol.	137
Taula 4.3: Els 20 descriptors més rellevants seleccionats en base a maximitzar la rellevància i minimitzar la redundància. Per a cada descriptor s'inclou la posició, el component i l'espai de color en el cas dels descriptors de color i textura, la regió d'interès i la classe de descriptor.	141
Taula 4.4: <i>p</i> valors obtinguts mitjançant el test de Kruskal-Wallis-Dunn i l'ajust de Bonferroni. Comparació múltiple dels valors quantitatius dels 20 descriptors entre sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil. Un <i>p</i> valor inferior a 0,05 s'ha considerat significatiu.	142
Taula 4.5: <i>p</i> valors dels descriptors de color i textura que mostren significació estadística quan es comparen els seus valors de les cèl·lules limfoides reactives amb cadascun dels nous subconjunts de limfòcits anormals.	143
Taula 4.6: Valors quantitatius dels descriptors geomètrics més rellevants. Els valors que es mostren són les medianes (rang interquartílic).	145
Taula 4.7: Correlacions de la relació nucli/citoplasma amb quatre descriptors (5è, 7è, 8è i 9è). A la taula es mostren els coeficients <i>r</i> i el rang dels <i>p</i> valors, els qual s'han obtingut mitjançant la correlació de Pearson. Un <i>p</i> valor inferior a 0,05 s'ha considerat estadísticament significatiu.	154
Taula 5.1: Tipus de cèl·lula, nombre (<i>n</i>) de pacients i imatges utilitzades en el primer conjunt d'imatges.	174
Taula 5.2: Tipus de cèl·lula, nombre (<i>n</i>) de pacients i imatges utilitzades en el segon conjunt d'imatges.	177
Taula 5.3: Diagnòstic dels pacients utilitzats a l'avaluació de l'eficiència diagnòstica dels descriptors (segona etapa de l'estudi) d'acord amb la classificació de l'OMS 2016, immunofenotip de les cèl·lules limfoides circulants a sang perifèrica i nombre de pacients (<i>P</i>) inclosos.	179
Taula 5.4: Nombre total d'imatges corresponents a cada diagnòstic i descriptor del tercer conjunt d'imatges.	180
Taula 5.5: Regió d'interès, característica citològica qualitativa, categoria, descriptor quantitatiu, nombre d'imatges, valors llindars mostren la mediana i el rang interquartílic (RIQ) i els <i>p</i> valors obtinguts de la comparació per parelles utilitzant el test de Kruskal-Wallis post-hoc i Dunn aplicant l'ajust de Bonferroni. Una $p < 0.05$ s'ha considerat estadísticament significativa. DE, desviació estàndard.	182
Taula 5.6: Valors quantitatius dels quatre descriptors que han permès la quantificació de les corresponents característiques citològiques qualitatives per als diferents grups cel·lulars del conjunt 2 d'imatges així com la regió d'interès a partir de la qual s'ha calculat cada descriptor. Els valors mostren la mediana (rang interquartílic).	183
Taula 5.7: Valors de tall (<i>cut-off</i>), àrea sota la corba (AUC), sensibilitat (<i>Sn</i>) i especificitat (<i>Sp</i>) obtinguts per a cada descriptor mitjançant l'anàlisi ROC.	192
Taula 5.8: Nombre total d'imatges corresponents a cada diagnòstic, mediana, rang interquartílic (RIQ) i imatges positives per a cada descriptor quantitatiu específic obtingudes a l'avaluació dels frotis.	192
Taula 6.1: Distribució dels diferents grups de cèl·lules limfoides, el nombre de pacients i imatges incloses a l'estudi.	207
Taula 6.2: <i>p</i> valors obtinguts mitjançant el test de Kruskal-Wallis-Dunn i l'ajust de Bonferroni. Comparació múltiple dels 20 descriptors obtinguts mitjançant el sistema DM96 i el BX43 per a sis parelles de subconjunts de cèl·lules limfoides anormals. Un $p < 0,05$ s'ha considerat significatiu. NS, no significació.	210
Taula 6.3: <i>p</i> valors dels descriptors geomètrics i de color i textura que mostren significació estadística quan es comparen els seus valors de les cèl·lules limfoides reactives amb cadascun dels nous subconjunts de limfòcits anormals.	211

Taula 7.1: Tipus de cèl·lula, nombre (n) de pacients i imatges utilitzades a l'estudi.....	223
Taula 7.2: Descriptors quantitius proposats per a la quantificació de quatre característiques citològiques.	224
Taula 7.3: Nombre total d'imatges adquirides amb el microscopi BX43 corresponents a cada diagnòstic i descriptor.....	226
Taula 7.4: Resultats de la correlació de Pearson dels tres descriptors entre el sistema DM96 i el BX43. Es mostra el coeficient R de Pearson, la recta d'ajust, l'interval de confiança del 95 % i el p valor.....	228
Taula 7.5: Valors dels quatre descriptors que han permès la quantificació de les corresponents característiques citològiques qualitatives en el conjunt d'imatges adquirides amb el sistema BX43. Els valors mostren la mediana i el rang interquartílic entre parèntesi.	229
Taula 7.6: Recta d'ajust, cut-offs d'ambdós sistemes d'adquisició d'imatges i valor lliard del sistema DM96 utilitzats en l'avaluació de l'eficiència diagnòstica dels tres descriptors.....	233
Taula 7.7: Nombre total d'imatges corresponents a cada diagnòstic, mediana, rang interquartílic (RIQ) i imatges positives per a cada descriptor obtingudes a l'avaluació de la classificació de frotis.....	233

Taula 9.1: Resultats de la correlació de Pearson dels 20 descriptors més rellevants entre el sistema DM96 i el BX43. Es mostra el coeficient R de Pearson, el coeficient de determinació (R^2), l'interval de confiança (IC) del 95 %, el p valor així com la recta d'ajust.....	263
Taula 9.2: Coeficients de determinació (R^2) dels 20 descriptors per a cada tipus cel·lular limfoide obtinguts entre el sistema DM96 i el BX43.	265
Taula 9.3: Valors quantitius dels descriptors geomètrics més rellevants obtinguts amb el sistema BX43. Els valors que es mostren són les medianes (rang interquartílic).....	267
Taula 9.4: Coeficient R de Pearson, recta d'ajust, interval de confiança del 95 % i coeficients de determinació (R^2) dels tres descriptors per a cada tipus cel·lular limfoide obtinguts entre el sistema DM96 i el BX43.	288

Llistat d'il·lustracions

Figura 1.1: Representació esquemàtica del flux de treball al laboratori d'Hematologia.....	30
Figura 1.2: Esquema de la Tesi. DM96 i BX43 fan referència als dos sistemes d'adquisició d'imatges de cèl·lules limfoides de sang perifèrica utilitzats.	40
Figura 2.1: Representació esquemàtica de les diferents línies cel·lulars maduratives a partir d'una cèl·lula mare pluripotencial. Font: Erasmus MC Rotterdam.....	45
Figura 2.2: Esquema del flux de treball al laboratori d'Hematologia.	46
Figura 2.3: Frotis de sang perifèrica.....	47
Figura 2.4: Frotis de sang perifèrica tenyida amb May Grünwald-Giemsa.	47
Figura 2.5: Hematies, limfòcit i plaqueta observats a sang perifèrica. Imatge adquirida a x1000.....	48
Figura 2.6: Leucòcits (glòbuls blancs) que trobem a la sang perifèrica en condicions normals.	49
Figura 2.7: Cèl·lula limfoide reactiva.	51
Figura 2.8: Mètodes d'anàlisi cel·lular i objectius de funcionament.	53
Figura 2.9: Autoanalitzador hematològic Advia 2120i®.	54
Figura 2.10: Representació gràfica de les diferents poblacions leucocitàries de la sang perifèrica classificades en funció de la seva mida i activitat peroxidasa obtingut del canal mieloperoxidasa de l'Advia 2120i®.....	54
Figura 2.11: Resultats que s'obtenen de l'autoanalitzador Advia 2120i® d'un pacient amb una neoplàsia limfoide B. WBC: recompte glòbuls blancs, Lymph: recompte limfòcits absoluts, ATYPS: alarma de cèl·lules atípiques.	55
Figura 2.12: Resultats que s'obtenen de l'autoanalitzador Advia 2120i® d'un pacient amb una leucèmia limfoide aguda.....	56
Figura 2.13: Visió externa del CellaVision® DM96 amb el monitor i l'aparell que conté a l'interior el microscopi robotitzat.....	57
Figura 2.14: Interior del CellaVision®DM96.....	58
Figura 2.15: Software d'adquisició i preclassificació.	58
Figura 2.16: Supervisió facultativa de la preclassificació.	58
Figura 2.17: Cobas ® m511 amb tecnologia Bloodhound.....	60
Figura 2.18: Estació de visualització del Cobas ® m511.....	61
Figura 2.19: Cèl·lula blàstica limfoide.....	64
Figura 2.20: Cèl·lules limfoides anormals característiques de les Neoplàsies Limfoides B i T.	65
Figura 3.1: Esquema del sistema per a l'anàlisi morfològica dels frotis de sang perifèrica (FSP) a les dues plataformes.	70
Figura 3.2: Conjunt de cèl·lules limfoides estudiades a la present <i>Tesi Doctoral</i> . Imatges obtingudes amb el CellaVision DM96 a x1.000 augments.....	73
Figura 3.3: Microscopi òptic Olympus BX43 amb càmera DP73 utilitzat en aquesta Tesi.	74
Figura 3.4: La imatge digital estructurada mitjançant una matriu formada per un nombre M x N de píxels quadrats de la mateixa mida.	75
Figura 3.5: Part a: Imatge en escala de grisos d'una cèl·lula plasmàtica. Part b: valors de la intensitat en una escala entre 0 i 255 per a una zona petita del nucli. Part c: Visualització de la intensitat dels píxels corresponents d'aquesta zona.....	75
Figura 3.6: Protocol estandarditzat d'adquisició d'imatges utilitzat durant la recerca.....	77
Figura 3.7: Diagrama de les diferents taules que conformen la base de dades.	78
Figura 3.8: Diagrama que resumeix les etapes de l'algoritme de segmentació (Font: S.Alfárez <i>et al.</i> (59))	81
Figura 3.9: Segmentació d'una cèl·lula limfoide anormal (tricleucòcit).....	84
Figura 3.10: Etapes per calcular el descriptor citoplasmàtic <i>hairiness</i> a la imatge d'un tricleucòcit. Després de la segmentació de la cèl·lula (I), es selecciona la regió externa de la cèl·lula (II). L'histograma del component verd d'aquesta regió mostra una distribució bimodal (III), sent la primera (70-215 d'intensitat)	

qui conté la major part de les projeccions citoplasmàtiques. Per últim, s'estableix la presència d'aquestes projeccions (IV) i es quantifica l'àrea. Font: Tesi Doctoral S. Alférez (3).	88
Figura 3.11: Part visible de l'espectre electromagnètic.	91
Figura 3.12: Representació dels diferents espais de colors utilitzats a la <i>Tesi Doctoral</i> .	94
Figura 3.13: Cercle representatiu del model de color HSV, on es representa la tonalitat o hue (H) com l'angle mesurat (en una escala entre 0 i 1) al voltant del cercle amb referència al color vermell (hue = 0). La saturació (S) es representa com la distància radial al centre del cercle que té un valor de hue = 0.	96
Figura 3.14: Imatge en escala de grisos d'un limfòcit anormal de sang perifèrica (esquerra) i el seu corresponent histograma (dreta).	98
Figura 3.15: Histogrames de cada canal de color de la imatge del tricoleucòcit.	99
Figura 3.16: Cèl·lules limfoides de sang perifèrica amb nuclis de diferent textura.	101
Figura 3.17: Càlcul la matriu de co-ocurrència del nivell de gris (<i>Gray Level Co-occurrence Matrix</i> o GLCM).	103
Figura 3.18: Càlcul de l'energia a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de la saturació del model de color HSV per la regió del nucli.	104
Figura 3.19: Càlcul de la correlació a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de blau del model de color RGB per la regió de la cèl·lula.	106
Figura 3.20: Càlcul de la probabilitat màxima a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de magenta del model de color CMGN per la regió de la cèl·lula.	107
Figura 3.21: Procés de dilatació (part esquerra de la corba pseudo-granulomètrica).	109
Figura 3.22: Corba pseudo-granulomètrica per a la regió del citoplasma d'una cèl·lula blàstica limfoide considerant el negre com a component de color.	110
Figura 3.23: Corbes granulomètriques i pseudo-granulomètriques per a la regió del citoplasma d'un limfòcit gran granular (imatge 1) i un limfòcit anormal vellós d'un limfoma de la zona marginal esplènica considerant el "u" com a component de color.	111
Figura 3.24: Diagrama de caixes del perímetre nuclear de 12 tipus diferents de cèl·lules limfoides.	118
Figura 3.25: Corba ROC.	127
Figura 4.1: Esquema del Capítol 4.	133
Figura 4.2: Exemples de cada un dels tipus de cèl·lules limfoides incloses a l'estudi. Imatges obtingudes amb el CellaVision DM96 a x1.000 augmentos.	134
Figura 4.3 : Anàlisi d'imatges de <i>l'entropia de l'histograma del component magenta de la cèl·lula</i> . (I) Imatges originals de dues cèl·lules de la síndrome de Sézary (SS). (II) Component magenta en escala de grisos. (III) Histogrames de les intensitats de magenta de la cèl·lula.	138
Figura 4.4: Exemples d'imatges de cèl·lules individuals que corresponen al valor més alt (A) i més baix (B) de la <i>relació nucli/citoplasma</i> i els corresponents diagrames de caixes.	144
Figura 4.5: Diagrames de caixes obtinguts per al descriptor del <i>perímetre nuclear</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.	146
Figura 4.6: Diagrames de caixes obtinguts per al descriptor del <i>diàmetre equivalent cel·lular</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.	146
Figura 4.7: Diagrames de caixes obtinguts per al descriptor del <i>hairiness</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.	147
Figura 4.8: Diagrames de caixes obtinguts per al descriptor de la <i>circularitat nuclear</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.	147
Figura 4.9: Anàlisi d'imatges de <i>la curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula</i> (4a posició). (I) Imatges originals de limfòcits normals. (II) Components de color en escala de grisos. (III) Corbes pseudo-granulomètriques. B, baix; A, alt.	148
Figura 4.10: Diagrames de caixes obtinguts per al descriptor de la <i>curtosi de la corba pseudo-granulomètrica del component cian de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.	149
Figura 4.11: Anàlisi d'imatges de <i>l'asimetria de l'histograma del component blau de la cèl·lula</i> (5a posició). (I) Imatges originals. (II) Components de color blau en escala de grisos. (III) Histogrames. B, baix; A, alt.	150

Figura 4.12: Diagrames de caixes obtinguts per al descriptor de <i>l'asimetria de l'histograma del component blau de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	151
Figura 4.13: Diagrames de caixes obtinguts per al descriptor de <i>la information measure of correlation I del component magenta de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	152
Figura 4.14: Anàlisi d'imatges de la desviació estàndard de l'histograma del component blau de la cèl·lula (7a posició). (I) Imatges originals de dos prolimfòcits B. (II) Components de color blau en escala de grisos. (III) Histogrames de les intensitats de blau. B, baix; A, alt.....	153
Figura 4.15: Diagrames de caixes obtinguts per al descriptor de <i>la desviació estàndard del component blau de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	153
Figura 4.16: Diagrames de caixes obtinguts per al descriptor de <i>l'entropia de l'histograma de component magenta de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	155
Figura 4.17: Anàlisi d'imatges de <i>l'asimetria de l'histograma del component verd de la cèl·lula</i> (9a posició). (I) Imatges originals de dues cèl·lules limfoides anormals d'un limfoma de cèl·lules del mantell. (II) Components de color verd en escala de grisos. (III) Histogrames de les intensitats de verd. B, baix; A, alt.....	156
Figura 4.18: Diagrames de caixes obtinguts per al descriptor de <i>l'asimetria de l'histograma del component verd de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	157
Figura 4.19: Diagrames de caixes obtinguts per al descriptor de <i>la mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	158
Figura 4.20: Diagrames de caixes obtinguts per al descriptor de <i>la mitjana de l'histograma del component blau de la regió del nucli</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	158
Figura 4.21: Diagrames de caixes obtinguts per al descriptor del <i>cluster shade del component de la saturació de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	159
Figura 4.22: Diagrames de caixes obtinguts per al descriptor de <i>la mitjana de l'histograma del component magenta de la regió del citoplasma</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi. ...	160
Figura 4.23: Diagrames de caixes obtinguts per al descriptor de <i>la homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	160
Figura 4.24: Diagrames de caixes obtinguts per al descriptor de <i>la curtosi de l'histograma del component de la saturació de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi.....	161
Figura 4.25: Diagrames de caixes obtinguts per al descriptor de <i>la Information measure of correlation I del component cian de la regió del nucli</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi. .	162
Figura 4.26: Diagrames de caixes obtinguts per al descriptor de <i>la desviació estàndard del component de la saturació de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi..	162
Figura 4.27: Diagrames de caixes obtinguts per al descriptor de <i>la homogeneïtat del component negre de la regió del citoplasma</i> als diferents grups de cèl·lules limfoides incloses a l'estudi.....	163
Figura 5.1: Esquema de la primera etapa de l'estudi.....	173
Figura 5.2: Exemples d'imatges de les diferents categories qualitatives considerades per a les quatre característiques citològiques. Per a la cromatina cerebriforme s'han inclòs dues imatges, una per a il·lustrar la cèl·lula de Sézary clàssica (SS) i l'altra de la variant petita (SS-L). Les fletxes indiquen l'absència i presència de grànuls al citoplasma. Les imatges han estat adquirides pel DM96. n, nombre d'imatges..	176
Figura 5.3: Esquema de la segona etapa de l'estudi.....	178
Figura 5.4: Anàlisi d'imatges de <i>la correlació del component cian de la regió del nucli</i>	184
Figura 5.5: Valors de <i>la correlació del component cian del nucli (CCN)</i> calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.....	185
Figura 5.6: Anàlisi d'imatges de <i>la desviació estàndard de la corba granulomètrica del component cian de la regió del nucli</i>	186
Figura 5.7: Valors de <i>la desviació estàndard de la corba granulomètrica del component cian del nucli (DGCN)</i> calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.....	187
Figura 5.8: Valors de <i>l'asimetria de l'histograma del component u del citoplasma (ASUC)</i> calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.....	188

Figura 5.9: Anàlisi d'imatges de <i>l'asimetria del component de color u de la regió del citoplasma</i>	189
Figura 5.10: Càlcul del descriptor <i>hairiness</i> per a un tricoleucòcit.....	191
Figura 5.11: Valors del <i>hairiness</i> calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.....	191
Figura 5.12: Corba ROC de la <i>correlació del component cian del nucli</i>	193
Figura 5.13: Corba ROC de la <i>DE de la corba granulomètrica del component cian del nucli</i>	194
Figura 5.14: Corba ROC de <i>l'asimetria de l'histograma del component u del citoplasma</i>	195
Figura 5.15: Corba ROC del <i>hairiness</i>	196
Figura 6.1: Esquema del Capítol 6.	206
Figura 6.2: Exemples de cada un dels tipus de cèl·lules limfoides incloses a l'estudi. Imatges obtingudes amb el microscopi Olympus BX43 a x1.000 augments.	207
Figura 6.3: Comparació del descriptor <i>hairiness</i> per la mateixa imatge (tricoleucòcit) adquirida amb els dos sistemes d'adquisició d'imatges (BX43 i DM96).	213
Figura 6.4: Càlcul del descriptor <i>hairiness</i> per a una imatge d'un limfòcit anormal de la leucèmia limfàtica crònica (LLC) adquirida amb el sistema BX43 i el corresponent histograma del component verd de la regió externa de la cèl·lula.....	214
Figura 7.1 Esquema de la primera etapa de l'estudi.....	222
Figura 7.2: Esquema de la segona etapa de l'estudi.....	225
Figura 7.3: Diagrames de caixes obtinguts per a la <i>correlació del component cian del nucli (CCN)</i> per als 15 grups de cèl·lules limfoides.....	230
Figura 7.4: Diagrames de caixes obtinguts per a la <i>desviació estàndard de la corba granulomètrica del component cian del nucli (DGCN)</i> per als 15 grups de cèl·lules limfoides.	230
Figura 7.5: Diagrames de caixes obtinguts per a <i>l'asimetria de l'histograma del component u del citoplasma (ASUC)</i> per als 15 grups de cèl·lules limfoides.....	231
Figura 7.6: Diagrames de caixes obtinguts per al <i>hairiness</i> per als 15 grups de cèl·lules limfoides.	232
Figura 7.7: Corba ROC de la <i>correlació del cian del nucli</i>	234
Figura 7.8: Corba ROC de la <i>DE de la corba granulomètrica del cian del nucli</i>	235
Figura 7.9: Corba ROC de <i>l'asimetria de l'histograma del component u del citoplasma</i>	236
Figura 7.10: Corba ROC del descriptor <i>hairiness</i>	237
Figura 9.1: Correlacions de Pearson dels 20 descriptors entre el sistema DM96 i el BX43 per a tots els grups cel·lulars.	264
Figura 9.2: Diagrama de caixes corresponents a la <i>relació nucli/citoplasma</i> per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.	266
Figura 9.3: Diagrama de caixes corresponents al <i>perímetre nuclear</i> per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.	268
Figura 9.4: Diagrama de caixes corresponents al <i>diàmetre equivalent</i> per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.	269
Figura 9.5: Diagrama de caixes corresponents a la <i>curtosi de la corba pseudogranulomètrica del component cian de la cèl·lula</i> per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.....	270
Figura 9.6: Diagrames de caixes corresponents a <i>l'asimetria de l'histograma del component blau de la cèl·lula</i> per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.....	271
Figura 9.7: Diagrames de caixes corresponents a la <i>Information measure of correlation 1 del component magenta de la cèl·lula</i> per als 12 grups de cèl·lules limfoides amb el sistema BX43.....	272
Figura 9.8: Diagrama de caixes corresponents a la <i>desviació estàndard de l'histograma del component blau de la cèl·lula</i> per als 12 grups de cèl·lules limfoides amb el sistema BX43.	273
Figura 9.9: Diagrames de caixes obtinguts per al descriptor de <i>l'entropia de l'histograma del magenta de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.....	274
Figura 9.10: Diagrames de caixes obtinguts per al descriptor de <i>l'asimetria de l'histograma del component verd de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	275

Figura 9.11: Diagrames de caixes obtinguts per al descriptor de la <i>mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43	276
Figura 9.12: Diagrames de caixes obtinguts per al descriptor de la <i>mitjana de l'histograma del component blau del nucli</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43....	277
Figura 9.13: Diagrames de caixes obtinguts per al descriptor del <i>cluster shade de la saturació de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	278
Figura 9.14: Diagrames de caixes obtinguts per al descriptor de la <i>mitjana de l'histograma del component magenta de la regió del citoplasma</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	279
Figura 9.15: Diagrames de caixes obtinguts per al descriptor de la <i>homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.....	280
Figura 9.16: Diagrames de caixes obtinguts per al descriptor de la <i>curtosi de l'histograma de la saturació de la regió de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	281
Figura 9.17: Diagrames de caixes obtinguts per al descriptor de la <i>Information measure of correlation 1 del component cian de la regió del nucli</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	282
Figura 9.18: Diagrames de caixes obtinguts per al descriptor <i>hairiness</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	283
Figura 9.19: Diagrames de caixes obtinguts per al descriptor de la <i>desviació estàndard de l'histograma de la saturació de la cèl·lula</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	284
Figura 9.20: Diagrames de caixes obtinguts per al descriptor de la <i>circularitat nuclear</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.....	285
Figura 9.21: Diagrames de caixes obtinguts per al descriptor de la <i>homogeneïtat del component negre del citoplasma</i> per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.	286
Figura 9.22: Correlacions de Pearson dels tres descriptors entre el sistema DM96 i el BX43 per a tots els grups cel·lulars	287

Llistat d'abreviatures

ASUC	Asimetria de l'histograma del component u de la regió del citoplasma
AUC	Àrea sota la corba
BD	Base de dades
BL	Blasts Limfoide
CCN	Correlació del component cian de la regió del nucli
CIE	Comissió Internacional d'Il·luminació
CLR	Cèl·lula Limfoide Reactiva
CMGN	Cian, Magenta, Groc, Negre
CP	Cèl·lula Plasmàtica
DE	Desviació estàndard
DGCN	Desviació estàndard de la corba granulomètrica del component cian de la regió del nucli
FSP	Frotis de Sang Perifèrica
GLCM	<i>Gray Level Co-occurrence Matrix</i> o Matriu de co-ocurrència de nivell de grisos
HCL	<i>Hairy Cell Leukemia</i> o Tricoleucèmia
LA	Limfòcit anormal
LCM	Limfoma de Cèl·lules del Mantell
LF	Limfoma Fol·licular
LI	Limfòcit
LLA	Leucèmia Limfoblàstica Aguda
LLC	Leucèmia Limfàtica Crònica
LLGG	Leucèmia de Limfòcits Grans Granulars
LM-B	Limfoma de Cèl·lules del Mantell amb morfologia blàstica
LM-C	Limfoma de Cèl·lules del Mantell amb morfologia clàssica
LPB	Leucèmia Prolimfocítica B
LPT	Leucèmia Prolimfocítica T
LZME	Limfoma de la Zona Marginal Esplènic
MGG	May Grünwald-Giemsa
MM	Mieloma Múltiple
MO	Monòcit
N/C	Nucli/Citoplasma
NL	Neoplàsia Limfoide
OMS	Organització Mundial de la Salut
PDI	Processament Digital d'Imatges
RI	Regió d'interès

ROC	<i>Receiver operating characteristics</i>
SP	Sang Perifèrica
SS	Síndrome de Sézary
SS-L	Síndrome de Sézary variant Lutzner
VPN	Valor predictiu negatiu
VPP	Valor predictiu positiu

Capítol 1

1 Motivació i objectius

1.1 Introducció

Les malalties hematològiques greus, especialment leucèmies i limfomes, són comunes a tots els períodes de la vida. La detecció precoç de la presència de cèl·lules leucèmiques a la sang perifèrica (SP) i la possibilitat posterior d'un tractament ràpid són essencials per a la supervivència del pacient. L'anàlisi morfològica cel·lular és el punt de partida per a l'orientació diagnòstica de més del 80 % de les malalties hematològiques.

Els diferents tipus de cèl·lules sanguínies normals són relativament fàcils de reconèixer per la seva morfologia, però les malignes, especialment les limfoides, són més difícils, ja que les diferències morfològiques entre els diferents tipus són subtils i el seu reconeixement exigeix gran experiència i habilitat. Tot i que la morfologia s'utilitza rutinàriament als laboratoris clínics amb finalitats diagnòstiques, les anormalitats observades estan subjectes a la interpretació amb certa subjectivitat (1). D'altra banda, no existeixen valors objectius per a definir variables citològiques (2).

El diagnòstic hematològic comença amb l'anàlisi morfològica i continua amb altres procediments més complexos com són la citometria de flux mitjançant l'ús d'anticossos monoclonals i els estudis genètics i moleculars, que estan només disponibles en laboratoris clínics altament especialitzats a causa dels costosos equips i els coneixements científics requerits. La Figura 1.1 mostra el flux de treball al laboratori d'Hematologia.

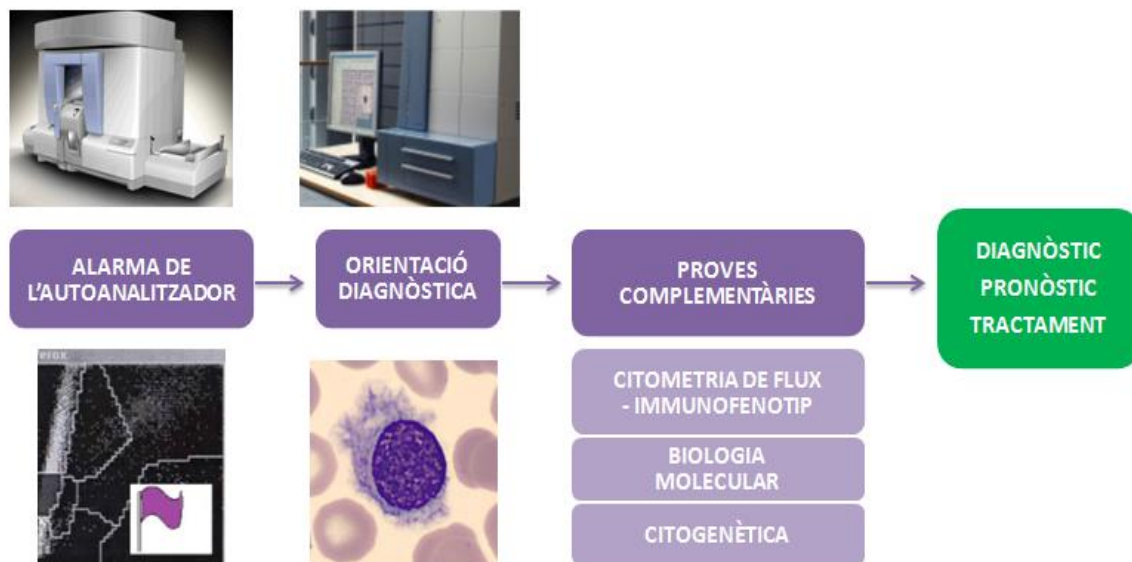


Figura 1.1: Representació esquemàtica del flux de treball al laboratori d'Hematologia.

La detecció i correcta classificació de les cèl·lules limfoides anormals és d'elevada importància per a poder donar una correcta orientació diagnòstica inicial. Els autoanalitzadors únicament mostren alarmes que poden indicar la possible presència de cèl·lules anormals. Recentment, han aparegut equips que realitzen preclassificacions automàtiques de cèl·lules de SP basats en el processament digital d'imatges (PDI), els quals mostren una alta eficàcia en el reconeixement de les cèl·lules normals de la SP com ara neutròfils, limfòcits, monòcits, eosinòfils i basòfils, però no identifiquen les cèl·lules limfoides anormals.

En els últims anys nous treballs han estat reportats a la literatura tractant d'omplir el buit en el reconeixement automàtic de les cèl·lules normals i anormals.

L'any 2011 es va iniciar una col·laboració entre membres del grup de recerca CoDALab (Control, Modelització, Identificació i Aplicacions) de la Universitat Politècnica de Catalunya (UPC), dirigit pel Prof. José Rodellar i la Dra. Anna Merino, la qual desenvolupa la seva tasca assistencial, docent i investigadora a la Unitat d'Hematologia i Citologia del Laboratori Core de l'Hospital Clínic de Barcelona. Va ser fruit de les serioses limitacions observades en el reconeixement de les cèl·lules malignes que presentava el software disponible al laboratori clínic que es va iniciar la primera línia de recerca, la qual va explorar l'aplicació de tècniques de PDI i reconeixement de patrons per a la classificació automàtica de cèl·lules limfoides de sang perifèrica i va donar lloc a la publicació de la primera tesi doctoral per *Alferez* (3).

La present *Tesi Doctoral* s'engloba dins d'un projecte concedit pel Ministeri d'Economia i Competitivitat d'Espanya al grup de recerca CoDALab seguint la línia de recerca iniciada. Aquest treball ha estat possible gràcies a una beca pre-doctoral (FPI-UPC) concedida per la UPC a la doctoranda. El treball s'ha realitzat en part a CoDALab, al Departament de Matemàtiques (entre l'Escola Industrial i el Campus Besós) així com a la Unitat d'Hematologia i Citologia del Laboratori Core de l'Hospital Clínic de Barcelona.

1.2 Motivació

La present *Tesi Doctoral* s'emmarca en els objectius globals del grup de recerca d'avançar en el desenvolupament de sistemes de reconeixement de les diferents classes de cèl·lules circulants a la sang. Els objectius que es plantegen en aquesta Tesi sorgeixen d'haver observat la gran varietat de descriptors quantitius que es troben implicats en el processament digital d'imatges cel·lulars i que poden tenir una significació prou rellevant per a la classificació automàtica dels diferents grups.

A mesura que s'han anat incrementant el nombre de grups cel·lulars objecte de classificació, el ventall de descriptors explorats ha anat en augment. A més a més, la diferenciació morfològica entre les cèl·lules limfoides anormals continua sent un problema sense resoldre que requereix experiència i habilitat, ja que són les cèl·lules més difícils de reconèixer si s'utilitzen només característiques morfològiques (4).

En el processament digital d'imatges, els descriptors assignen valors numèrics a característiques morfològiques cel·lulars que tenen una base visual i que generalment es relacionen amb la geometria, el color i la textura, com es veurà en detall durant la presentació de tota la recerca. Mentre que els descriptors geomètrics tenen generalment una interpretació fàcilment intuïtiva, les eines i els algorismes de l'anàlisi d'imatges i l'aprenentatge automàtic (*machine learning*) generen uns descriptors quantitius que tenen una base matemàtica ben establerta i poden donar classificadors eficients, però són difícils d'interpretar en termes morfològics i de traduir en els termes utilitzats pels citòlegs dels laboratoris clínics.

D'altra banda, la revisió de la literatura ha posat de manifest que en la majoria de treballs sobre el reconeixement automàtic de cèl·lules sanguínies anormals, no s'analitzen ni s'interpreten els descriptors que utilitza el classificador, és a dir, no intenten esbrinar la seva significança ni contribució cap a una major objectivitat en la morfologia cel·lular.

Pel que fa als tipus de cèl·lules limfoides estudiades, determinades cèl·lules limfoides anormals tals com les velloses (tricoleucèmia i limfoma de la zona marginal esplènica leucemitzat) han estat escassament analitzades. D'altra banda, els limfòcits reactius no s'han estudiat prèviament en treballs de reconeixement automàtic, malgrat que alguns estudis han posat de manifest la variabilitat en la classificació per part del citòleg expert d'una mateixa sèrie d'imatges de limfòcits normals, reactius i anormals (5). Això demostra la imperiosa necessitat d'una estandardització basada en paràmetres objectius i quantitius de tots ells.

Fins ara el nombre màxim de grups de cèl·lules limfoides (incloent anormals) reconeguts automàticament als treballs publicats ha estat un màxim de cinc: limfòcits normals, blasts i limfòcits anormals de les següents neoplàsies limfoides: leucèmia limfàtica crònica, leucèmia de cèl·lules del mantell i limfoma fol·licular. La present Tesi, dins de l'ampli ventall de cèl·lules sanguínies, se centra en les cèl·lules limfoides anormals associades a les neoplàsies limfoides.

Una de les principals limitacions dels estudis previs és que utilitzen conjunts d'imatges petits en comparació amb el número de tipus de cèl·lules a classificar. A més a més, exceptuant els treballs del grup de Tuzel, Yang i Foran (6–8), no es descriu una etapa de validació dels mètodes pacient a pacient, indispensable per a avaluar les metodologies de forma més real en un escenari més proper al diagnòstic mitjançant la identificació de les cèl·lules anormals.

Així doncs, la motivació de la Tesi és contribuir a la caracterització morfològica de les cèl·lules limfoides anormals de sang perifèrica mitjançant la quantificació de variables citològiques eficients en la classificació dels diferents tipus cel·lulars. Com es precisarà més endavant, es pretén enfocar la recerca des de dues perspectives complementàries:

- (1) partint de les eines de l'aprenentatge automàtic, es buscaran descriptors geomètrics, de color i de textura que tinguin una formulació quantitativa explícita i una interpretació qualitativa raonable en termes morfològics visuals;
- (2) partint d'algunes anormalitats cel·lulars establertes *a priori* pels citòlegs, les quals estan associades a cèl·lules i malalties específiques i de particular interès clínic, l'objectiu serà identificar descriptors quantitius específics. Aquests descriptors haurien de descriure objectivament característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius.

L'aplicabilitat dels resultats de la present *Tesi Doctoral* rau en el potencial de reduir la subjectivitat associada a la morfologia de sang perifèrica. L'estandardització de l'ús de descriptors quantitius podria ser interessant per donar suport al diagnòstic morfològic diferencial i, molt especialment, quan l'immunofenotip de les cèl·lules limfoides per citometria de flux no és conclouent.

En la present recerca s'han inclòs més de 200 pacients diferents i fins a 16 grups cel·lulars limfoides diferents. Amb el sistema CellaVision DM96 s'han adquirit automàticament unes 12.000 imatges de cèl·lules limfoides normals, reactives, anormals i blàstiques mentre que amb el microscopi Olympus BX43, s'han adquirit manualment unes 9.000 imatges. Els estudis basats en l'anàlisi d'imatges publicats a la literatura no inclouen un nombre tant elevat de diferents tipus de neoplàsies limfoides, d'imatges ni de pacients.

Per a la descripció i diferenciació cel·lular s'han analitzat gairebé 2.700 descriptors entre geomètrics, de color i de textura (estadístics de primer i segon ordre, granulomètrics, *Wavelet* i Gabor). S'han considerat sis espais de color: RGB, CMGN, HSV, XYZ, Lab i Luv. El conjunt de descriptors extrets ha estat capaç de descriure la mida, la forma, la textura i el color dels diferents tipus de cèl·lules limfoides incloses en la recerca. Alguns dels estudis previs han considerat alguns dels espais de color anteriorment mencionats però cap ha inclòs els sis. Pel que fa a les característiques citològiques analitzades, la majoria dels estudis

anteriors s'han centrat en característiques geomètriques i són pocs els autors que han desenvolupat una metodologia per a característiques de textura del nucli i del citoplasma (9,10).

L'apartat següent presenta una anàlisi de l'estat de l'art dels treballs més significatius amb aportacions relacionades amb descriptors per a la diferenciació de les diferents neoplàsies limfoides. El contingut d'aquest apartat ens ajudarà a concretar més la motivació de la nostra recerca i a plantejar els objectius.

1.3 Estat de l'art

Aquest apartat se centra a presentar els diferents treballs realitzats fins ara que han utilitzat descriptors geomètrics, de color i de textura per tal de diferenciar entre les diferents neoplàsies limfoides. Com s'ha comentat anteriorment, la microscòpia digital i l'anàlisi computacional d'imatges poden proporcionar una avaluació morfològica més precisa i objectiva, transformant els paràmetres qualitius citològics en valors quantitius. La secció 1.3.1 presenta les publicacions sobre descriptors geomètrics mentre que la 1.3.2 sobre els descriptors de color i de textura. Per últim, la secció 1.3.3 conté una revisió dels treballs que s'han centrat en el reconeixement automàtic de cèl·lules limfoides anormals.

1.3.1 Descriptors geomètrics

Els descriptors geomètrics s'han utilitzat anteriorment per a la caracterització quantitativa de limfòcits en els teixits tiroïdals, en concret, en el limfoma de cèl·lules petites de la glàndula tiroide. L'estudi va consistir a comparar els paràmetres nuclears (diàmetre, irregularitat i àrea) dels limfòcits i va demostrar que l'àrea nuclear era el descriptor òptim per tal de diferenciar entre els limfòcits petits neoplàsics en els limfomes tiroïdals i els limfòcits reactius presents a la tiroïditis de Hashimoto (11).

D'altra banda, l'anàlisi d'imatges també s'ha utilitzat per avaluar l'efecte de l'albumina sobre la morfologia cel·lular dels limfòcits de sang perifèrica en el cas de la leucèmia limfàtica crònica (LLC) (12). Lunning *et al.* demostra que l'addició d'albumina a la sang provoca canvis en les característiques citològiques, disminuint no només les àrees cel·lulars, sinó també les nuclears, al mateix temps que augmenta la relació nucli/citoplasma (N/C).

Benattar *et al.* (13) proposa un sistema de puntuació per als limfòcits de les neoplàsies limfoides tipus B construït a partir de diversos descriptors morfomètrics: forma cel·lular, forma cel·lular, àrea cel·lular, relació nucli/citoplasma, relació nuclear del vermell/blau, relació citoplasmàtica del verd/blau i la proporció de cèl·lules amb nuclèol.

Sabino *et al.* (14) estudia el problema de reconeixement de leucòcits mitjançant el càlcul de descriptors de forma i mida com el perímetre, l'àrea, la circularitat i la relació N/C. Classifica cinc tipus de leucòcits normals i només un tipus de cèl·lula limfoide neoplàsica, la leucèmia limfàtica crònica (LLC), utilitzant màquines de suport vectorial.

Angulo *et al.* (9,10) extreu diversos paràmetres quantitius dels limfòcits per a definir algunes característiques morfològiques qualitatives: mides nuclears i cel·lulars, relació N/C, excentricitat nuclear, formes nuclears normal i irregular mitjançant paràmetres senzills (circularitat, excentricitat) i l'anàlisi específica dels lòbuls nuclears i d'altres irregularitats així com el nombre de nuclèols presents. El treball compara diferents tipus de cèl·lules limfoides anormals, classificant els seus descriptors en categories.

En una publicació posterior, Jahanmehr *et al.* (2) utilitza l'anàlisi d'imatges per extreure descriptors geomètrics per quantificar alguns paràmetres citològics dels limfòcits malignes de les següents tres neoplàsies limfoides tipus B: LLC, leucèmia de cèl·lules del mantell (LCM) i leucèmia prolimfocítica B. Concretament, l'àrea cel·lular, el diàmetre cel·lular, l'àrea del citoplasma, l'àrea nuclear, la relació N/C i la densitat nuclear són avaluades. L'estudi demostra que aquests descriptors poden ser útils per diferenciar les neoplàsies limfoides.

1.3.2 Descriptors de color i de textura

Des de la introducció per Haralick, la matriu de co-ocurrència del nivell de gris (en anglès *GLCM*) ha estat àmpliament utilitzada com a mesura de textura en imatges mèdiques, com ara imatges d'ultrasons per a neoplàsies sòlides (15) i imatges de medul·la òssia per distingir quatre tipus de cèl·lules precursors d'eritròcits (16).

No obstant això, pocs estudis amb *GLCM* s'han realitzat amb imatges de sang perifèrica i principalment han diferenciat entre els leucòcits normals i les cèl·lules blàstiques limfoides (17). Yamamoto *et al.* van calcular cinc atributs texturals basats en la *GLCM* (entropia energètica, entre d'altres) per diferenciar entre leucòcits normals (cinc subtipus) i cèl·lules de la LLC (18).

Sabino *et al.* (14) estudia el problema de reconeixement de leucòcits mitjançant descriptors de textura sobre l'espai CIELAB basats en la *GLCM* aplicada sobre blocs de diferents mides per a una versió en nivells de gris de la imatge cel·lular. En treballs posteriors (19,20), els autors estenen el mètode amb l'aplicació dels descriptors de textura als components de color RGB, i es selecciona de nou els descriptors més importants per a una classificació dels cinc tipus de leucòcits normals i només un tipus de cèl·lula limfoide neoplàsica (LLC) utilitzant màquines de suport vectorial. Els resultats són excel·lents per al leucòcits normals, però l'efectivitat disminueix per a les cèl·lules de la LLC.

Per tal de descriure les cèl·lules limfoides neoplàsiques i les cèl·lules blàstiques, Tuzel *et al.* (6,8) realitza una representació cel·lular que caracteritza la seva estructura de textura utilitzant textons dins del nucli i el citoplasma mitjançant la construcció de dos histogrames de textons. Aquest treball classifica tres tipus de cèl·lules limfoides anormals tipus B (LLC, LCM, limfoma fol·licular (LF)), blasts i limfòcits normals utilitzant només descriptors de textura i una combinació entre descriptors de forma, mida i textura.

Angulo *et al.* (9,10) extreu diversos paràmetres quantitius dels limfòcits per a definir algunes característiques morfològiques qualitatives com la densitat de la cromatina mitjançant corbes granulomètriques, la basofília citoplasmàtica (mitjançant la mitjana de cada un dels components de color de l'espai de color *Lab*), la granulació citoplasmàtica i la forma del citoplasma a partir de granulometria. El treball compara diferents tipus de cèl·lules limfoides anormals, classificant els seus descriptors en categories. Les característiques nuclears són un problema distintiu que suporta el diagnòstic morfològic. En el seu mètode, Angulo *et al.* analitzen la textura nuclear basant-se en la noció de granulometria sobre la

luminància (component L de l'espai de color del *Lab*) i estableix quatre grups: cromatina molt laxa, laxa, densa i molt densa.

1.3.3 Reconeixement de cèl·lules limfoides anormals

Els treballs de Comaniciu *et al.* i Foran *et al.* (21,22) presenten un sistema d'ajuda a la presa de decisions a la patologia clínica basat en la consulta d'imatges mitjançant exemple. S'utilitza un mètode d'agrupament no paramètric sobre els components de l'espai *Luv* per a segmentar únicament la regió nuclear. El sistema inclou paràmetres tals com: àrea, descriptors el·líptics de Fourier i descriptors de textura del nucli per a la caracterització de cèl·lules limfoides neoplàsiques o anormals (LLC, LF i LCM), així com de limfòcits normals, a través de mesures de similitud i l'optimització de la probabilitat d'una correcta classificació. El sistema s'avalua utilitzant validació creuada (*10-fold cross validation*) mostrant un bon rendiment en comparació amb el citòleg expert.

Saeedizadeh *et al.* (23) presenta un mètode per al reconeixement automàtic de cèl·lules de mieloma en imatges obtingudes per microscòpia utilitzant l'algoritme *bottleneck*, la transformada de Watershed modificada i el classificador de màquines de suport vectorial. Ho apliquen a 50 imatges digitals obtingudes a partir d'extensions d'aspirats de moll d'os que contenen 678 cèl·lules: 132 cèl·lules plasmàtiques normals, 256 cèl·lules de mieloma i 29 d'altres tipus de cèl·lules. Obtenen una sensibilitat del 96,5 %, una especificitat del 93,0 % i una exactitud del 95,3 % en la identificació de les cèl·lules de mieloma.

Treballs previs en la detecció automàtica de cèl·lules blàstiques limfoides utilitzen una metodologia a partir d'imatges de cèl·lules de la sang, obtenint la regió d'interès del nucli mitjançant agrupament del color i la classificació del veí més pròxim. Calculen els següents descriptors de la regió nuclear: rugositat del perímetre, forma, color (mitjanes dels components dels espais RGB i HSV), i quatre descriptors de segon ordre dels nivells de grisos de la imatge. Els descriptors s'utilitzen per a classificar imatges de blasts limfoides, mieloides i limfòcits normals mitjançant màquines de suport vectorial. El mètode aconsegueix una notable exactitud (24,25), però l'estudi es realitza segmentant només el nucli, amb una classificació de només dos tipus, i no especifica el número d'imatges per cèl·lula.

González *et al.* (26) presenta una estratègia per identificar diferents tipus de blasts a partir d'imatges de cèl·lules del moll de l'os, extraient descriptors geomètrics, de textura i valors propis obtinguts mitjançant Anàlisi de Components Principals (PCA), sobre cada component de color de la imatge. Els experiments de classificació utilitzen diferents algoritmes (AdaBoost, RandomForest i altres d'*ensemble learning*), obtenint una elevada exactitud per a diferenciar entre vàries leucèmies agudes mieloides i limfoides. Aquest treball utilitza la tinció Wright sobre imatges del moll de l'os i, tot i que hi ha un número acceptable de mostres d'imatges de cèl·lules totals, la distribució de les mateixes per a cada subtipus no és homogènia.

Madhloom *et al.* (27) segmenta els limfòcits i blasts limfoides mitjançant la umbralització de l'espai de color HSV, i l'aplicació de la reconstrucció morfològica. Els descriptors més rellevants i menys

correlacionats (entre descriptors de forma i estadístics de textura del nucli i la cèl·lula) se seleccionen mitjançant discriminació de Fisher. Així, aquests descriptors s'utilitzen en un classificador kNN per a reconèixer automàticament entre imatges de limfòcits normals i blasts limfoides, obtenint una exactitud excel·lent. No obstant, el número d'imatges de cèl·lules estudiades és molt baix (n=260), no s'assenyala quants pacients s'han inclòs, la segmentació es realitza només per al nucli i es realitza una classificació que inclou només dos tipus de cèl·lules.

Una metodologia per al reconeixement automàtic d'imatges de blasts limfoides i mieloides es presenta a Aimi *et al.* (28), on extreuen descriptors de mida, forma i color del nucli i del citoplasma. Posteriorment, aquests descriptors es classifiquen utilitzant dues tècniques: ANN (*Artificial Neural Network*) multicapa i *simplified Fuzzy ARTMAP* ANN, donant una exactitud molt bona per al ANN multicapa. Tot i així, el treball utilitza un número relativament baix de cèl·lules als experiments (n=200) i no s'especifica el número de pacients inclosos a l'estudi, ni la tinció utilitzada.

Pel que fa al nostre grup de recerca, en un primer estudi realitzat a l'any 2013, es van extreure 44 descriptors (geomètrics, de color i textura) de 340 imatges digitals i mitjançant *clustering Fuzzy C-means*, es va assolir una classificació precisa dels següents grups: limfòcits normals (83,3 %) i dos tipus de cèl·lules limfoides anormals: tricoleucèmia (98 %) i leucèmia limfàtica crònica (71,3 %) (29). En un treball posterior (30), es van incloure 1.650 imatges digitals de les següents cèl·lules limfoides: limfòcits normals i quatre tipus de limfòcits anormals: tricoleucèmia, leucèmia limfàtica crònica, leucèmia de cèl·lules del mantell i leucèmia prolimfocítica B. Un total de 113 descriptors (geomètrics, estadístics de primer i segon ordre i granulomètrics) van ser extrets per al reconeixement de les cèl·lules limfoides mitjançant una anàlisi discriminant linear amb una validació creuada posterior. L'exactitud de la classificació del conjunt de validació va ser del 85,3 %.

A l'any 2016 es va realitzar un altre estudi (31), el qual incloïa per primera vegada els limfòcits reactius i els limfòcits anormals del limfoma fol·licular, juntament amb els cinc grups cel·lulars limfoides prèviament inclosos als estudis previs. En aquest estudi, es va incloure un major nombre d'imatges cel·lulars (4.389) obtingudes de 105 pacients i es van extreure 13 descriptors geomètrics i 2.451 descriptors de color i textura (utilitzant sis espais de color diferents sobre tres regions d'interès: nucli, citoplasma i cèl·lula sencera). L'exactitud de la classificació va ser del 91,2 %.

1.4 Objectius

L'**objectiu general** de la present *Tesi Doctoral* és obtenir nous descriptors quantitativs mitjançant processament digital d'imatges i aprenentatge automàtic per a la diferenciació de cèl·lules limfoïdes normals, reactives i anormals de sang perifèrica, contribuint a una anàlisi objectiva de la citologia sanguínia.

Aquest objectiu es desdobra en els següents **objectius específics** complementaris:

- (1) Partint de les eines d'aprenentatge automàtic, es pretén obtenir descriptors geomètrics, de color i de textura que tinguin una formulació quantitativa explícita i una interpretació qualitativa en termes morfològics visuals.

(1A) Obtenir descriptors quantitativs de les diferents regions d'interès (nucli, citoplasma, cèl·lula i regió externa de la cèl·lula) d'imatges de cèl·lules limfoïdes circulants a sang perifèrica.

(1B) Avaluar la capacitat discriminant dels descriptors quantitativs mitjançant tècniques estadístiques per a la diferenciació entre cèl·lules limfoïdes reactives, anormals i blàstiques de sang perifèrica.

- (2) Cercar descriptors quantitativs citològics per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, la tricoleucèmia i el limfoma de la zona marginal esplènica, la síndrome de Sézary i la leucèmia de limfòcits grans granulars.

(2A) Identificar descriptors quantitativs específics, els quals haurien de descriure objectivament característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius, com la cromatina i condensada a la leucèmia limfàtica crònica, cromatina cerebriforme a la síndrome de Sézary, vellositats citoplasmàtiques a la tricoleucèmia i al limfoma de la zona marginal esplènica i la granulació citoplasmàtica a la leucèmia de limfòcits grans granulars.

(2B) Establir escales numèriques objectives per cada descriptor per tal de relacionar-lo amb la corresponent característica qualitativa i cèl·lula limfoïde específica mitjançant una anàlisi estadística.

(2C) Avaluar l'eficiència diagnòstica dels quatre descriptors per a la classificació de frotis de sang perifèrica en una determinada característica citològica: cromatina madura i condensada, cromatina cerebriforme, granulació citoplasmàtica i vellositats citoplasmàtiques.

Per assolir aquests objectius s'ha utilitzat el CellaVision DM96 com a sistema principal per adquirir les imatges digitals dels frotis de sang perifèrica. Es tracta d'un analitzador basat en tècniques d'imatge i d'aprenentatge automàtic que actualment s'utilitza a molts laboratoris clínics. Així, els resultats de la Tesi s'emmarcaran dins l'estat de l'art de les aplicacions d'aquestes tècniques per a la valoració morfològica de les cèl·lules sanguínies. En un segon nivell, s'explorará la utilització d'un microscopi convencional per tal de valorar si els resultats obtinguts són reproduïbles amb imatges adquirides en un sistema més convencional i a l'abast de laboratoris més petits.

1.5 Organització de la Tesi i contribucions

La Figura 1.2 il·lustra l'organització de la present Tesi, que permet seguir el fil conductor de la mateixa. En concret, aquesta *Tesi Doctoral* s'ha centrat en la caracterització morfològica de cèl·lules limfoides normals, reactives, anormals i blàstiques de sang perifèrica.

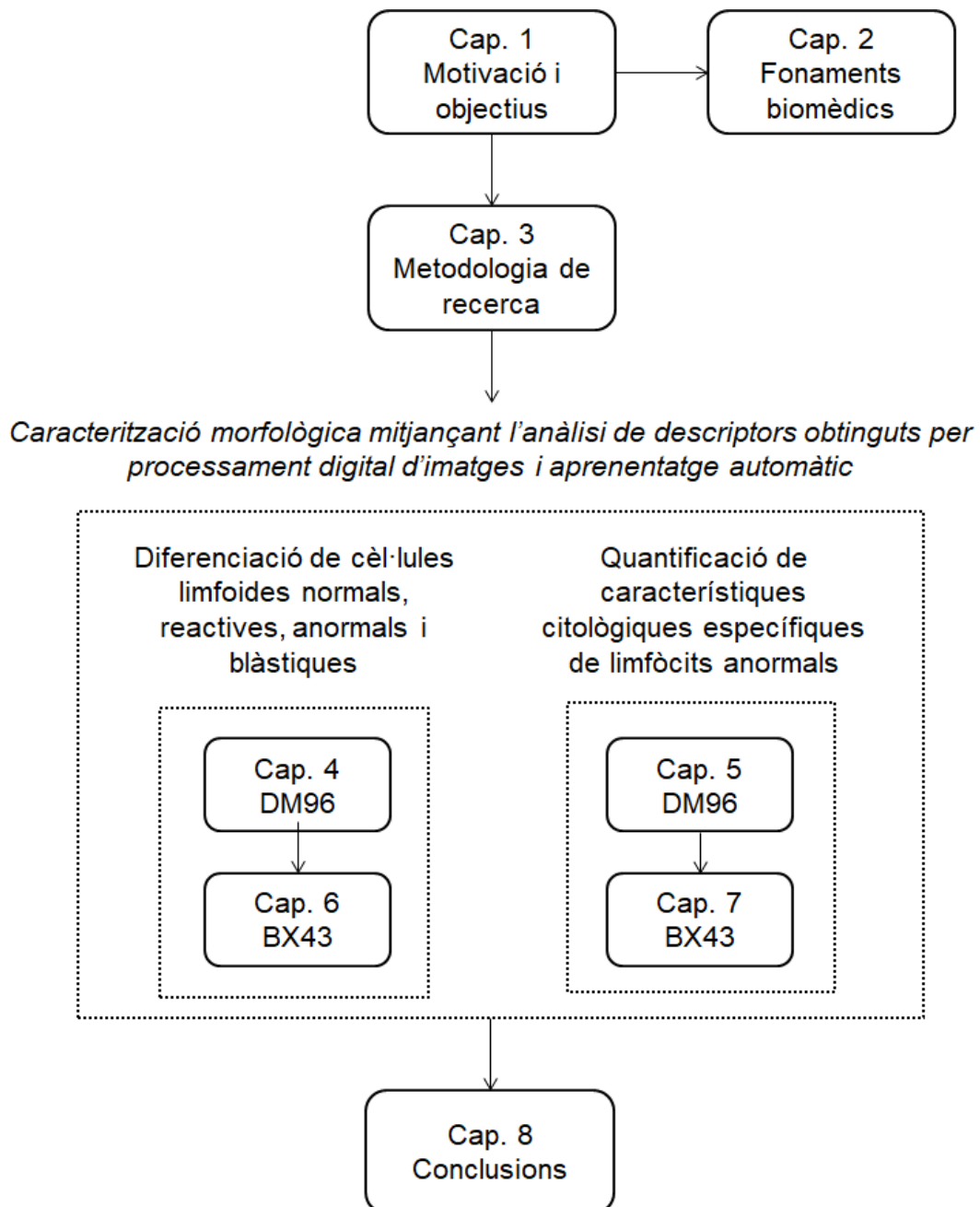


Figura 1.2: Esquema de la Tesi. DM96 i BX43 fan referència als dos sistemes d'adquisició d'imatges de cèl·lules limfoides de sang perifèrica utilitzats.

La present Tesi consta de nou capítols (l'últim està format per dos annexos) que comencen amb el present Capítol 1 on es descriu la motivació, l'estat de l'art sobre els descriptors utilitzats en la identificació de les diferents cèl·lules sanguínies a partir del processament digital d'imatges, així com els objectius de la Tesi. El Capítol 2 inclou una revisió d'alguns conceptes bàsics sobre l'anàlisi morfològica de la sang perifèrica, centrant-se en les cèl·lules limfoides anormals mentre que el Capítol 3 detalla la metodologia de la recerca.

Del Capítol 4 al Capítol 7 es realitza una caracterització morfològica dels diferents tipus de cèl·lules limfoides de sang perifèrica analitzant descriptors geomètrics, de color i de textura obtinguts mitjançant processament digital d'imatges. Aquests capítols són el nucli principal de la Tesi.

Cal tenir en compte que hi ha una distribució especial de lectura segons dues qüestions paral·leles:

1. Els Capítols 4 i 6 tenen com a objectiu analitzar i interpretar els 20 descriptors que s'han seleccionat en el processament digital d'imatges per a la diferenciació de 12 tipus diferents de cèl·lules limfoides, incloent cèl·lules limfoides normals, reactives, anormals i blàstiques.
2. En els Capítols 5 i 7 el punt de partida no és aquest, sinó que es parteix d'algunes anormalitats cel·lulars establertes *a priori* pels citòlegs, les quals estan associades a cèl·lules i malalties hematològiques específiques i de particular interès clínic. L'objectiu serà identificar descriptors quantitius específics, els quals haurien de descriure objectivament característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius.

En concret, el Capítol 4 presenta el primer treball realitzat, el qual analitza i interpreta els 20 descriptors més rellevants per a la diferenciació entre els següents 12 grups de cèl·lules limfoides incloent limfòcits normals, cèl·lules limfoides reactives, anormals de neoplàsies limfoides B i T i blasts limfoides. El treball utilitza el sistema d'adquisició d'imatges automàtic CellaVision® DM96. Els resultats són la base del primer article derivat de la present Tesi i publicat al *Journal of Clinical Pathology* (2017).

Al Capítol 5 la metodologia d'adquisició de les imatges és la mateixa que al Capítol 4, però l'objectiu del treball és obtenir descriptors quantitius citològics per a la identificació de característiques citològiques específiques dels limfòcits anormals de la leucèmia limfàtica crònica, la tricoleucèmia i el limfoma de la zona marginal esplènica, la síndrome de Sézary i la leucèmia de limfòcits grans granulars així com l'establiment d'escala numèriques objectives.

Per tant, el Capítol 5 implica l'extracció d'un nombre considerable de descriptors geomètrics, de color i de textura per tal de trobar un descriptor que permeti la quantificació de cada una de les característiques citològiques considerades. En el treball s'han inclòs un total de 15 tipus de cèl·lules limfoides, incloent cèl·lules limfoides normals, reactives i neoplàsiques. Es diferencien els dos subtipus de cèl·lules de Sézary (clàssiques i variant Lutzner) i els dos subtipus de cèl·lules del mantell (clàssica i blàstica). Mitjançant una anàlisi estadística, s'han identificat i seleccionat els millors descriptors per a la identificació de: (1)

anomalies limfoides nuclears, com la cromatina madura i condensada o la cerebriforme i (2) característiques del citoplasma, com la granulació o les vellositats.

Els descriptors identificats en aquest Capítol 5 són nous pel reconeixement de les cèl·lules limfoides anormals de determinades malalties hematològiques. A més a més, no hi ha publicacions anteriors que hagin quantificat les anomalies limfoides nuclears presentades en aquesta recerca. Per últim, s'ha realitzat una avaluació de l'eficiència diagnòstica dels quatre descriptors per a la classificació del frotis de sang perifèrica en una determinada categoria diagnòstica en funció de les característiques citològiques que predominen (cromatina madura i condensada, cromatina cerebriforme, granulació citoplasmàtica o vellositats citoplasmàtiques). Els resultats obtinguts al Capítol 5 han estat publicats recentment a l'*American Journal of Clinical Pathology* (2019).

En relació al Capítol 6, aquest demostra que els resultats presentats al Capítol 4 amb el sistema DM96 són reproduïbles en imatges digitals obtingudes mitjançant un microscopi convencional, com és el microscopi Olympus BX43 amb la càmera digital DP73. S'ha utilitzat un significatiu nombre d'imatges digitals adquirides de forma manual (9.000) i comparat el poder discriminant dels 20 descriptors presentats al Capítol 4 en un total de 12 tipus diferents de cèl·lules limfoides anormals.

El Capítol 7 reproduïx els resultats obtinguts amb els descriptors obtinguts al Capítol 5 però utilitzant les imatges digitals obtingudes mitjançant el sistema BX43.

El Capítol 8 resumeix les conclusions finals de la Tesi, les seves contribucions i algunes perspectives de futur.

Per últim, el Capítol 9 conté dos annexos. A l'annex A es mostren els resultats de la correlació així com l'anàlisi estadística descriptiva dels 20 descriptors presentats als Capítols 4 i 6 entre els dos sistemes d'adquisició d'imatges: DM96 i BX43. A l'annex B es presenten els resultats de la correlació de tres dels quatre descriptors analitzats al Capítol 5 amb utilitat en la quantificació de característiques citològiques específiques obtingudes mitjançant els dos sistemes d'adquisició d'imatges (DM96 i BX43).

Capítol 2

2 Fonaments biomèdics del reconeixement de cèl·lules sanguínies

2.1 Introducció

En la present *Tesi Doctoral* es realitza una caracterització morfològica de diferents tipus de cèl·lules limfoides de sang perifèrica mitjançant l'anàlisi de descriptors geomètrics, de color i de textura obtinguts mitjançant processament digital d'imatges (PDI). Aquest capítol presenta una revisió d'alguns conceptes bàsics sobre l'anàlisi morfològica de la sang perifèrica, centrant-se en les cèl·lules limfoides anormals.

El capítol s'organitza de la manera següent. La secció 2.2 presenta el concepte d'hematopoesi i la secció 2.3 detalla el procés de l'anàlisi morfològica de les cèl·lules sanguínies de sang perifèrica. La secció 2.4 conté els mètodes d'anàlisi cel·lular automatitzada, la secció 2.5 detalla els sistemes automatitzats de morfologia digital i la secció 2.6 descriu les diferents neoplàsies limfoides i la seva classificació.

2.2 Hematopoesi

La hematopoesi es defineix com el procés de la producció, desenvolupament, diferenciació i maduració de totes les cèl·lules sanguínies (32) (Fig. 2.1). Aquestes s'originen al moll de l'os i deriven d'una cèl·lula mare hematopoètica (*stem cell*), la qual té dues funcions fonamentals: l'auto renovació i la diferenciació.

La proliferació cel·lular suposa l'amplificació del nombre de cèl·lules madures produïdes a partir d'una cèl·lula que s'ha compromès amb una estirp cel·lular en concret. En condicions normals, existeix un equilibri entre la quiescència i la capacitat d'auto renovació de les cèl·lules mare hematopoètiques, existint control de la proliferació, l'apoptosi i la diferenciació dels progenitors a cèl·lules més madures. La diferenciació cel·lular per la seva banda implica el desenvolupament progressiu de característiques bioquímiques, funcionals i estructurals específiques per a un tipus de cèl·lula (33–35).

Els leucòcits que circulen a la sang perifèrica (SP) es divideixen en diferents estirps cel·lulars: mieloide, monocítica i limfoide. Aquestes cèl·lules expressen una sèrie d'antígens a la seva superfície, al citoplasma o al nucli, que van variant durant el transcurs del seu procés maduratiu. La detecció d'aquests antígens mitjançant anticossos monoclonals, ha permès molts avenços en el coneixement dels leucòcits i el seu procés maduratiu. Aquests anticossos monoclonals han estat caracteritzats i descrits mitjançant una numeració denominada llista de molècules CD (*cluster of differentiation*) humanes. El principi de la citometria de flux es basa en que cèl·lules amb antígens específics, poden ser identificades mitjançant anticossos monoclonals marcats amb fluorocroms. Aquesta tècnica s'aplica al diagnòstic i a la classificació de les neoplàsies hematològiques, així com per al seguiment de la malaltia residual (33).

2.3 Anàlisi morfològica de la sang perifèrica

Donat que la SP és un fluid fàcilment accessible, l'observació visual al microscopi és clau per a una primera detecció d'anomalies a les cèl·lules circulants (36). L'anàlisi morfològica de l'extensió o frotis de sang perifèrica conté el percentatge de cada una de les diferents sub poblacions leucocitàries.

Aquesta prova pot ésser sol·licitada directament pel clínic davant d'una sospita diagnòstica concreta i sempre que existeixin limfadenopaties i/o esplenomegàlia. També és tasca del facultatiu del laboratori indicar la realització de l'anàlisi morfològica quan els resultats de l'hemograma presenten alguna anomalia quantitativa, o bé apareix alguna alarma a l'autoanalitzador (Fig. 2.2).

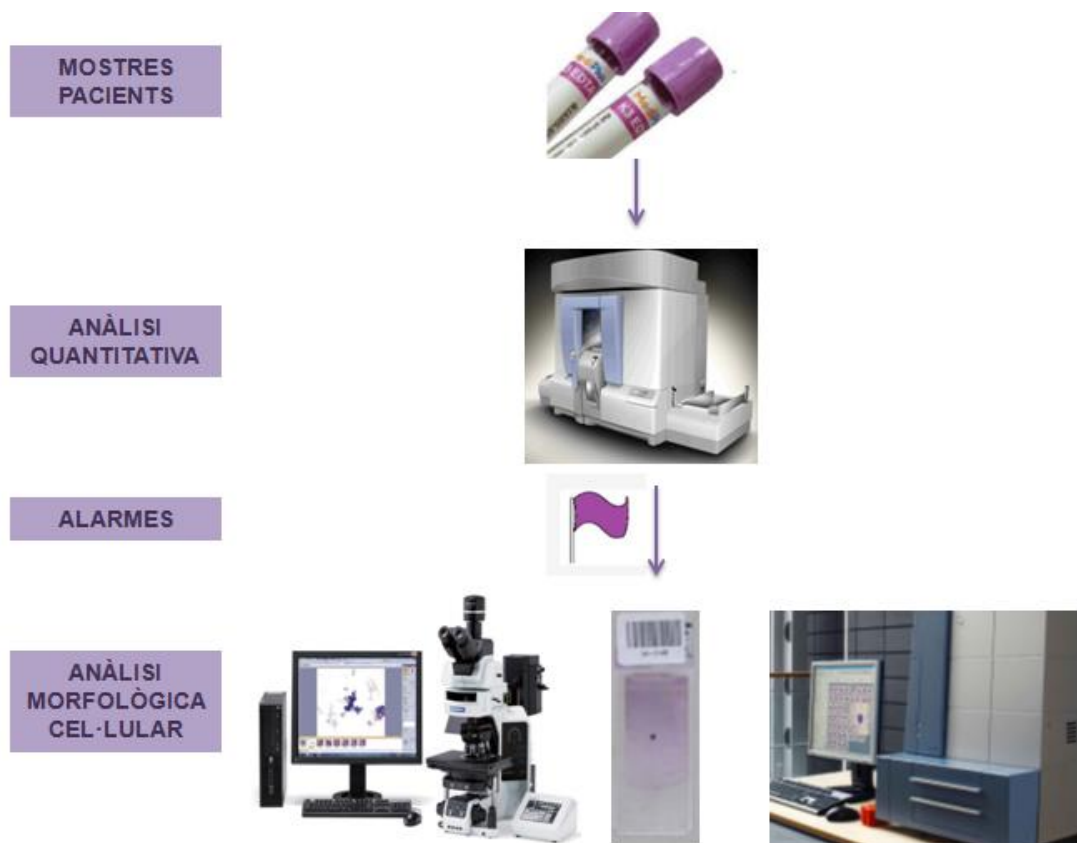


Figura 2.2: Esquema del flux de treball al laboratori d'Hematologia.

2.3.1 Frotis de sang perifèrica

Per a observar al microscopi els diferents components cel·lulars de la SP es realitza un frotis de sang perifèrica. Es diposita una gota de sang sobre un portaobjectes i es col·loca l'extrem d'un segon portaobjectes sobre la gota formant un angle de 45° respecte al primer. L'angle d'inclinació del portaobjectes que llisca sobre la gota de sang condiona el gruix de l'extensió. Quant major sigui l'angle, més gruixuda i curta serà l'extensió (Fig. 2.3).



Figura 2.3: Frotis de sang perifèrica.

2.3.2 Fixació

Un cop realitzat el frotis de SP, i quan no es procedeix a la immediata tinció, les cèl·lules s'han de fixar. Per a la fixació de les extensions s'utilitza metanol absolut durant un període de temps entre 10 i 20 minuts (34). El mètode de tinció més comunament utilitzat és el denominat May Grünwald-Giemsa.

2.3.3 Tinció amb May Grünwald-Giemsa

La tinció de May Grünwald-Giemsa (MGG) consisteix en : a) submergir les extensions en May Grünwald pur durant un minut i mig, b) rentar les extensions (amb aigua de l'aixeta), c) submergir els frotis ja rentats en una segona cubeta que conté Giemsa al 10 % (dilució en aigua destil·lada i filtració prèvia a l'ús) durant 10 minuts, d) submergir les extensions durant uns segons en aigua de l'aixeta (segon rentat per retirar el Giemsa) i, finalment, deixar assecar.

Petites variacions del pH poden modificar els resultats del procediment de tinció. Així, quan el pH és massa baix els components basòfils de les cèl·lules no es tenyeixen adequadament, fet que provoca que el citoplasma dels limfòcits sigui molt pàl·lid i els eosinòfils presentin una granulació brillant. Per contra, si el pH és massa alt els hematies policromàtics s'assemblen als normals i els polimorfonuclears presenten una granulació intensament tenyida, que pot confondre's amb la granulació tòxica (34).

A la Figura 2.4 es mostra un frotis de sang perifèrica tenyit amb MGG, el qual permet una bona observació de les característiques morfològiques de les cèl·lules sanguínies.



Figura 2.4: Frotis de sang perifèrica tenyida amb May Grünwald-Giemsa.

2.3.4 Observació al microscopi

L'observació al microscopi de les tres sèries hematopoètiques (eritroblàstica, granulopoètica i megacariocítica) de SP ens permet la detecció de diferents patologies hematològiques, que es posen de manifest mitjançant alteracions citològiques característiques. Per a l'observació dels components cel·lulars de l'extensió correctament tenyida i valorar el detall de l'observació utilitzarem l'objectiu d'immersió de 100 augments (total x 1000).

Quan s'observa al microscopi una extensió de SP, per a valorar la morfologia eritrocitària ens hem de fixar en a) el número, b) la mida, c) la forma i d) el grau d'hemoglobinització dels hematies. El gruix de l'extensió està influenciat per la viscositat de la mostra de sang, que depèn en part del contingut eritrocitari en hemoglobina (34). La zona de l'extensió més favorable per a l'observació morfològica dels elements sanguinis és la que s'anomena "barbes", on els hematies es troben uns a prop dels altres, però sense solapar-se (34).

Si l'hemoglobina està anormalment elevada, el que succeeix per exemple a la *Policitemia Vera*, l'extensió acostuma a quedar uniformement gruixuda degut a la major viscositat sanguínia, i els hematies s'observen uns sobre uns altres. D'altra banda, en situacions d'anèmia, l'extensió acostuma a ser prima i amb una separació exagerada entre els hematies.

Un exemple del que podem veure a la SP amb un microscopi òptic es mostra a la Figura 2.5, on s'observa un limfòcit (glòbul blanc) al centre, envoltat majoritàriament per hematies (glòbuls vermells) i algunes plaquetes. La mida del limfòcit és similar a la de l'hematia (7 μm), mostra una relació nucli/citoplasma elevada, un nucli de perfil arrodonit de cromatina madura, i un citoplasma escàs i basòfil.

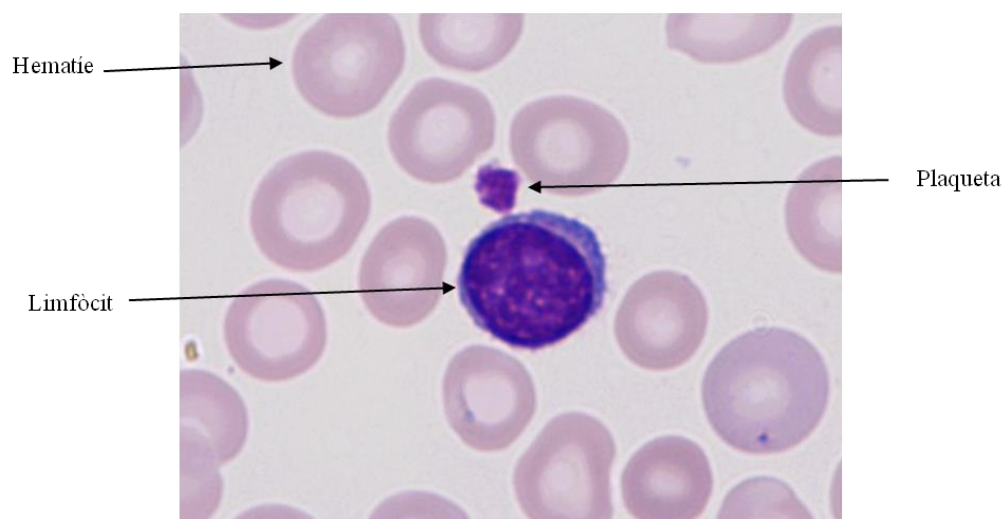


Figura 2.5: Hematies, limfòcit i plaqueta observats a sang perifèrica. Imatge adquirida a x1000.

2.3.5 Cèl·lules benignes de la sang perifèrica

A continuació descriurem l'aspecte morfològic i les funcions dels diferents tipus cel·lulars benignes presents a la SP: els hematies, les plaquetes i els glòbuls blancs. Els glòbuls blancs o leucòcits es divideixen en granulòcits o polimorfonuclears (neutròfils segmentats, eosinòfils i basòfils) i cèl·lules mononucleades (limfòcits i monòcits) (Fig. 2.6).



Figura 2.6: Leucòcits (glòbuls blancs) que trobem a la sang perifèrica en condicions normals.

Els hematies o glòbuls vermells

Són cèl·lules sense nucli i contenen un component essencial, que és una hemoproteïna d'unió a l'oxigen: l'hemoglobina. La funció principal d'aquestes cèl·lules és assegurar el transport d'oxigen i CO₂ entre els alvèols pulmonars i els teixits. Són cèl·lules amb un diàmetre aproximat de 7,5 µm, d'aspecte homogeni, amb una zona pàl·lida central que els hi confereix una forma de disc bicòncav, propietats que permeten que es deformin i puguin travessar els vasos sanguinis més prims, els capil·lars. En condicions normals, el nombre d'hematies a SP és de $5 \times 10^6/\text{mm}^3$, trobant-se lleugerament més elevat en els homes. La seva coloració a l'extensió de SP és deguda a que l'hemoglobina es tenyeix amb l'eosina, un dels colorants del May Grünwald-Giemsa.

Les plaquetes

Les cèl·lules precursors de les plaquetes són els megacariòcits, de mida gran i nucli polilobulat. Les plaquetes normals no tenen nucli, tenen un diàmetre d'1-3 µm i contenen una fina granulació azuròfila, que també és característica del megacariòcit. La seva funció principal es troba en els fenòmens d'iniciació de la coagulació, tant a l'adhesió com a l'agregació.

Els neutròfils

Són les cèl·lules polinuclears més nombroses, del 40 al 75 % del total de glòbuls blancs. Tenen un diàmetre d'uns 12 µm, el nucli acostuma a estar trilobulat (el nombre de lòbuls varia de 2 a 5 i és un índex de maduració de la cèl·lula). El citoplasma conté granulació i és incolor, ja que no es tenyeix amb la tinció MGG. S'originen al moll de l'os, circulen a SP durant 6 – 10 hores i després passen dels capil·lars als teixits, on la seva funció principal és la fagocitosi.

Els eosinòfils

Representen de l'1 al 3 % dels glòbuls blancs i la seva granulació específica és acidòfila. Tenen un diàmetre de 10 a 14 µm, el nucli és generalment bilobulat i el citoplasma ataronjat amb MGG d'aspecte granulós degut a que capten l'eosina de la tinció. La funció dels eosinòfils és la participació, en sinèrgia amb altres cèl·lules, en reaccions d'hipersensibilitat immediates i retardades.

Els basòfils

Són les cèl·lules menys nombroses dels polinuclears (0 – 1 % del total dels glòbuls blancs). Tenen un diàmetre de 10 a 14 µm. El seu nucli és irregular, pot prendre una forma de trèvol (d'1 a 4 lòbuls), però habitualment es troba emmascarat per les nombroses granulacions de color púrpura intens (capten els colorants bàsics del MGG). Les funcions d'aquestes cèl·lules són les manifestacions al·lèrgiques de tipus immediat.

Els monòcits

Aquestes cèl·lules constitueixen entre el 2 i el 10 % dels leucòcits. Són rodones, amb un diàmetre de 15 a 20 µm. El citoplasma és d'un color blau grisaci amb MGG i té un aspecte poc granulós (de sorra fina). El nucli és central, de forma arronyonada o de C. Funcionalment, pertanyen al sistema mononuclear fagocític.

Els limfòcits

Són els leucòcits més nombrosos a SP després dels neutròfils (25 – 33 %). Són mononucleats amb una relació nucli/citoplasma elevada. Són de mida petita, al voltant d'uns 7 µm de diàmetre amb un nucli esfèric i dens, que ocupa gairebé la totalitat de la cèl·lula, tot i que un reduït percentatge d'aquests poden tenir una mida superior (12 – 16 µm), amb un nucli rodó de cromatina condensada o madura i un citoplasma basòfil molt escàs. En els limfòcits de mida superior, la cromatina és menys condensada i el citoplasma és més abundant.

Tots els limfòcits són semblants pel que fa a la morfologia però existeixen diferents grups de limfòcits que es poden diferenciar gràcies als marcadors antigènics de membrana: els limfòcits B i els limfòcits T. Es descriu un tercer grup que s'assembla als limfòcits T: les cèl·lules NK o "*natural killer*". La població limfocitària sanguínia compren: 8 – 12 % de limfòcits B, 70 – 80 % de limfòcits T i 5 – 15 % de limfòcits NK. Aquestes cèl·lules són responsables de les respostes immunes específiques.

En situació normal, com hem comentat, la forma d'un limfòcit és regular i rodona i existeix una petita franja citoplasmàtica perifèrica però la morfologia dels limfòcits està subjecte a una gran variabilitat degut a: (1) estímuls immunològics de malalties inflamatòries o infeccioses, especialment víriques i (2) malalties neoplàsiques (leucèmies i limfomes). Així, la circulació en sang de limfòcits amb diferents anomalies morfològiques pot ser deguda a malalties benignes o, pel contrari, malignes.

Segons les recomanacions recentment publicades pel Comitè Internacional d'Estandardització en Hematologia (ICSH) (37), es descriuen com a limfòcits reactius als que deriven d'una causa benigna, i com a limfòcits anormals als derivats de malalties hematològiques malignes, que s'originen als ganglis limfàtics, tals com els diferents tipus de limfomes, que poden ser de tipus B o T. Els limfòcits B i T provenen del moll de l'os, els primers maduren als ganglis limfàtics i els segons al timus (38).

A continuació veurem les cèl·lules limfoides reactives.

Els limfòcits reactius són de mida gran, tenen un nucli de cromatina laxa i un citoplasma ampli, intensament basòfil, especialment als punts de contacte amb els hematies veïns. És freqüent observar-les en frotis de pacients que pateixen infeccions víriques o bacterianes. A la Figura 2.7 podem veure una cèl·lula limfoide reactiva (CLR).

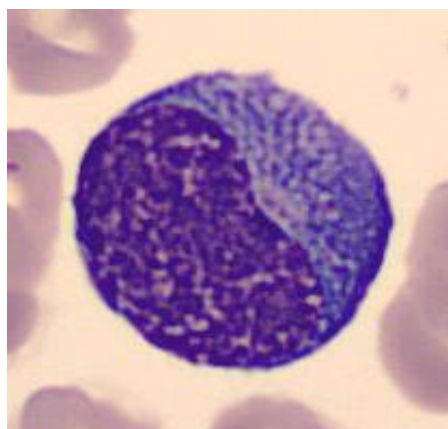


Figura 2.7: Cèl·lula limfoide reactiva.

2.3.6 Rellevància del frotis de sang perifèrica en el diagnòstic de les neoplàsies limfoides

Existeixen una sèrie de protocols específics de realització de l'anàlisi morfològica cel·lular manual, ja que es tracta d'una prova laboriosa (35). En el cas dels limfomes o neoplàsies limfoides de cèl·lules madures, la xifra de leucòcits a sang perifèrica acostuma a estar alterada, observant-se leucocitosi (augment de la xifra de leucòcits) amb limfocitosi. Per exemple, sovint el diagnòstic de la leucèmia limfàtica crònica (LLC) es realitza a partir d'una troballa accidental d'una leucocitosi amb limfocitosi en una analítica de control en un pacient d'edat avançada amb la detecció de limfòcits anormals al frotis de sang perifèrica.

L'estudi citològic ens permetrà realitzar una orientació diagnòstica morfològica. La informació de la morfologia de les cèl·lules limfoides anormals és de gran interès en el diagnòstic inicial i indica quines proves complementàries s'han de realitzar, ja que el diagnòstic definitiu es realitzarà integrant tots els resultats: aspirat del moll de l'os, estudis immunofenotípics, citogenètics i de biologia molecular (35,38–40).

Tot i així, com l'estudi morfològic és el primer pas en el diagnòstic, el reconeixement de limfòcits anormals al frotis de sang perifèrica pot contribuir a un diagnòstic ràpid de les neoplàsies limfoides, permetent una ràpida intervenció terapèutica, la qual influirà en el pronòstic (41). Es requereix experiència per a diferenciar morfològicament entre diferents tipus de cèl·lules limfoides anormals i, a més a més, no existeixen valors objectius per a definir variables citològiques (2).

2.4 Mètodes d'anàlisi cel·lular automatitzada

L'anàlisi cel·lular ens permet detectar la presència de cèl·lules malignes a la sang, el que és especialment rellevant quan el pacient visita el servei d'Urgències d'un hospital. Per a poder realitzar tant el diagnòstic com el seguiment d'un pacient, d'una banda, necessitem conèixer les dades de l'hemograma o el recompte sanguini complet, i d'altra banda, es requereix realitzar el recompte diferencial de les diferents subpoblacions leucocitàries (42).

Els mètodes fonamentals del recompte cel·lular són els següents: el Principi de Coulter o Impedància, la Citometria de flux i la Fluorescència (35).

Els objectius dels autoanalitzadors hematològics són, d'una banda, l'exactitud en la mesura del nombre total cel·lular, i per l'altra banda, evitar els falsos negatius en la identificació de cèl·lules anòmales, que són marcades per una alarma que indica la necessitat de revisió del frotis de sang perifèrica per part del professional del laboratori (Fig. 2.8). Aquestes alarmes apareixen quan existeixen cèl·lules que no identifiquen els autoanalitzadors (Fig. 2.9), com són les cèl·lules limfoides reactives, blàstiques o cèl·lules limfoides anormals (43).

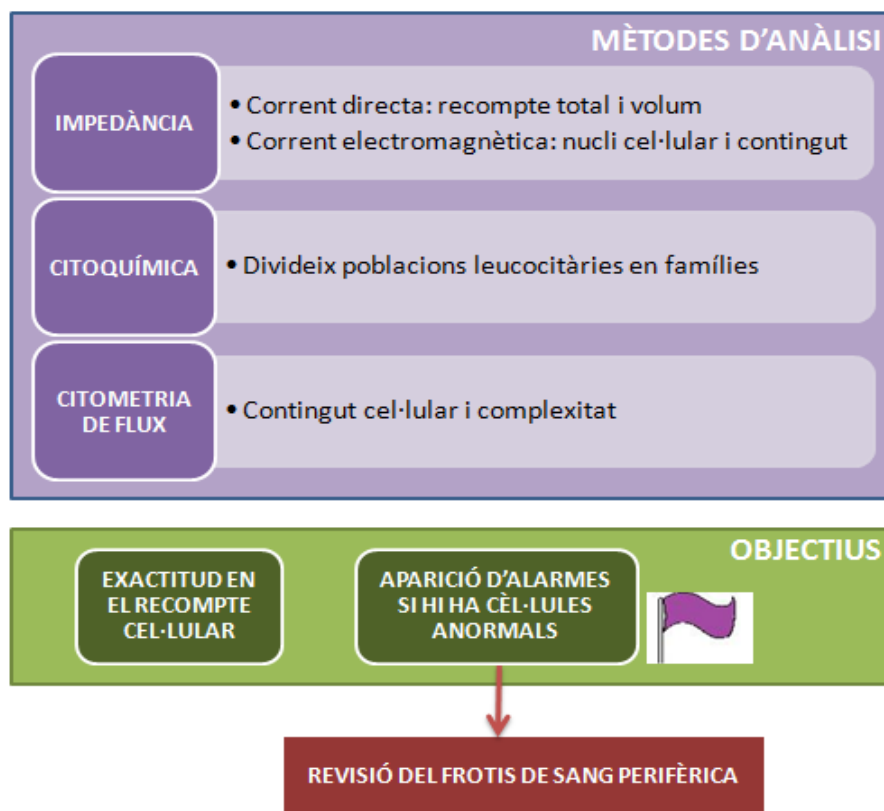


Figura 2.8: Mètodes d'anàlisi cel·lular i objectius de funcionament.



Figura 2.9: Autoanalitzador hematològic Advia 2120i®.

La Figura 2.10 mostra la representació gràfica de les diferents poblacions leucocitàries de sang perifèrica classificades en funció de la seva mida i activitat peroxidasa obtingut del canal mieloperoxidasa de l'Advia 2120i®. Les cèl·lules blàstiques, les cèl·lules limfoides reactives i algunes cèl·lules limfoides anormals se situen a la regió 5, fet que es reflexa per una alarma a l'autoanalitzador, el qual no és capaç d'identificar-les.

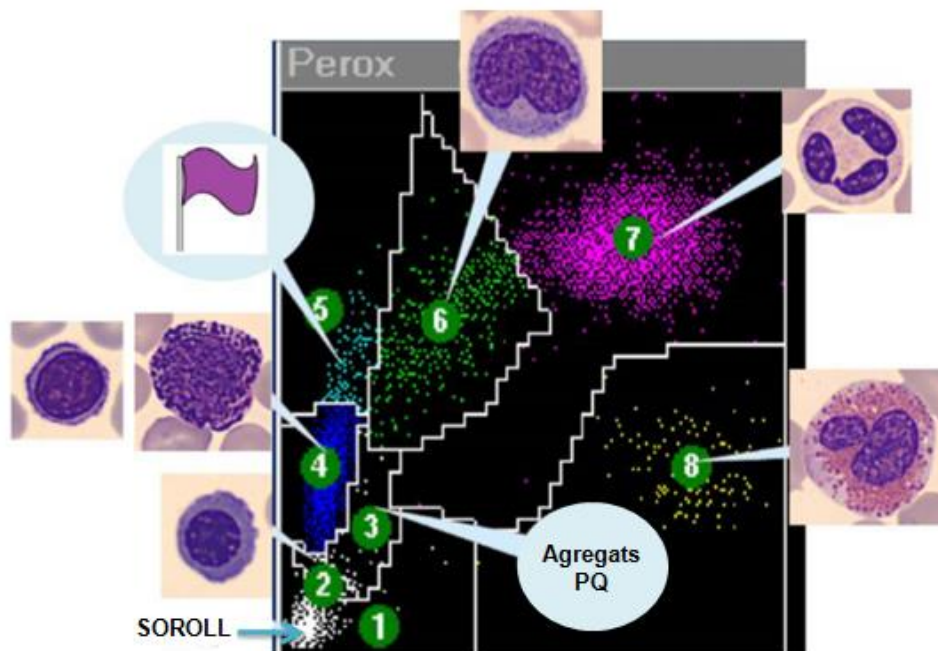


Figura 2.10: Representació gràfica de les diferents poblacions leucocitàries de la sang perifèrica classificades en funció de la seva mida i activitat peroxidasa obtingut del canal mieloperoxidasa de l'Advia 2120i®.

El recompte diferencial dels diferents leucòcits (*WBC Differential*) mostrant un 90% de limfòcits. A la columna següent hi ha les alarmes de morfologia (*Morphology Flags*) i destaca l'alarma de cèl·lules anormals (ATYPS: ++) que correspondria a cèl·lules limfoides anormals. Pel que fa a les gràfiques, cal observar que s'ha encerclat la població que correspondria a les cèl·lules limfoides anormals.

La Figura 2.12 conté els resultats d'un pacient amb una leucèmia aguda limfoide. Destaca la presència d'un 66,2 % de limfòcits i un 32 % de LUC (*Large Unstained Cells*: podrien correspondre a cèl·lules limfoides reactives, anormals, blàstiques o plasmàtiques). A la columna següent hi ha les alarmes de morfologia (*Morphology Flags*) i destaca l'alarma de cèl·lules anormals (ATYPS: ++). L'observació al microscopi ens demostraria que en aquest cas correspondrien a limfoblasts.

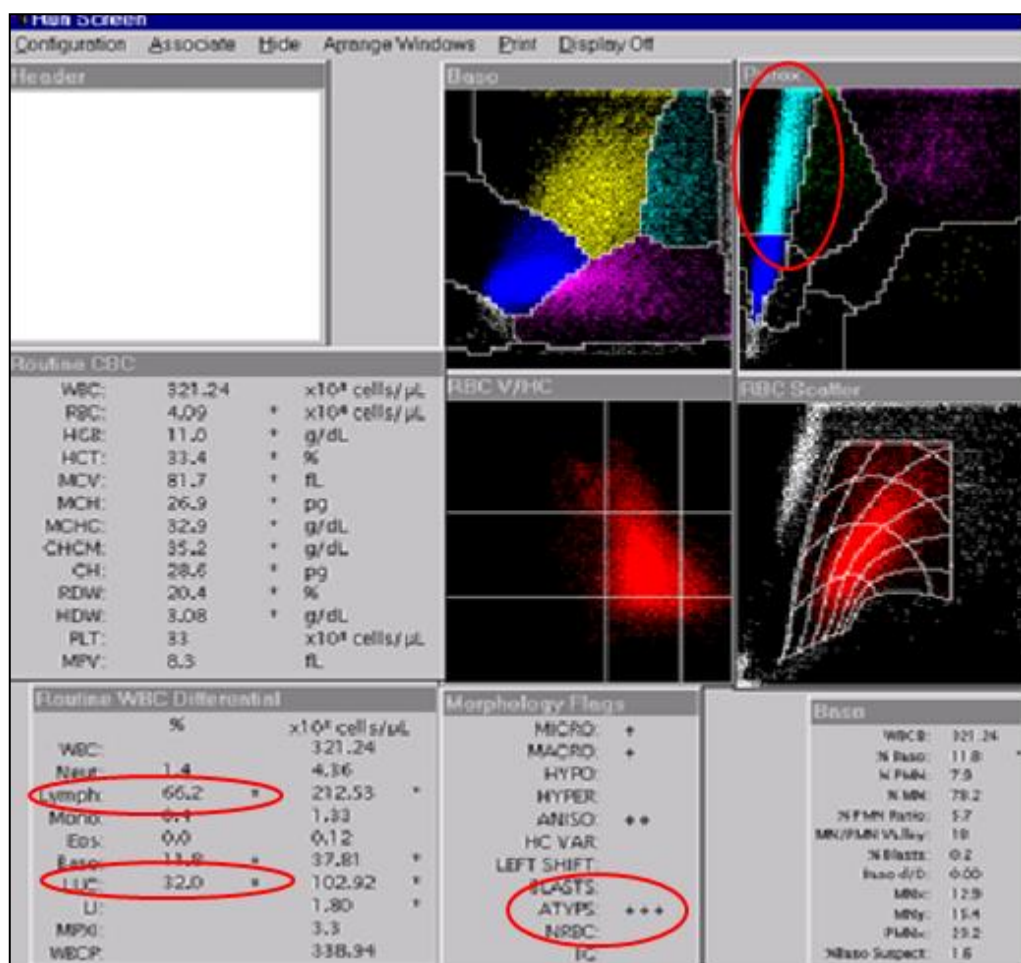


Figura 2.12: Resultats que s'obtenen de l'autoanaltzador Advia 2120i® d'un pacient amb una leucèmia limfoide aguda.

2.5 Sistemes automatitzats de morfologia digital

Des de principis del segle XXI s'inicia una nova tendència cap a l'estudi dels mètodes de processament d'imatges digitals, sortint al mercat instruments capaços de realitzar la preclassificació de les diferents subpoblacions leucocitàries amb temps d'anàlisi curts i de manera eficient. Aquests sistemes utilitzen microscòpia motoritzada, processament digital d'imatges i tècniques de reconeixement de patrons per a la identificació automàtica de cèl·lules nucleades i a l'hora realitzen una valoració morfològica de la sèrie vermella. Posteriorment, mostren les imatges a una pantalla per a la confirmació o reclassificació per part del facultatiu responsable (46).

MedicaEasyCell® *assistant* és un sistema d'anàlisi d'imatges cel·lulars (Sysmex America, Inc, Mundelein, IL) (47) que detecta automàticament els leucòcits del frotis de sang perifèrica i classifica els normals, les cèl·lules trencades i els eritroblasts. Aquest sistema utilitza processament digital d'imatges i intel·ligència artificial. Les imatges són classificades i preparades per a la revisió facultativa.

El sistema HemaCAM® (Fraunhofer-Gesellschaft, Alemanya) (Fraunhofer Institute for Integrated Circuits) és capaç d'analitzar mostres de SP i fer una preclassificació automàtica dels leucòcits en neutròfils, basòfils, eosinòfils, limfòcits i monòcits, així com ombres nuclears i plaquetes grans. Des de l'any 2010, HemaCAM® està disponible al mercat com a producte mèdic certificat.

El CellaVision®DM96 (CellaVisionAB, Lund, Suècia) (Fig. 2.13) és un dispositiu automàtic per a la preclassificació diferencial dels leucòcits, l'avaluació de la morfologia dels hematies, l'estimació del recompte de plaquetes i pot ser utilitzat per a analitzar líquids biològics.

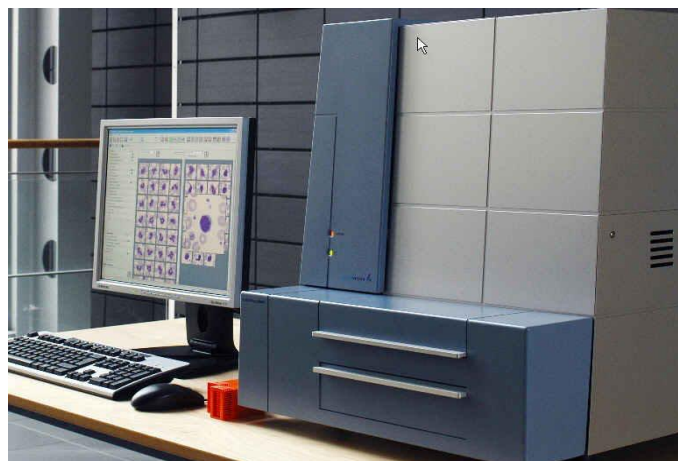


Figura 2.13: Visió externa del CellaVision® DM96 amb el monitor i l'aparell que conté a l'interior el microscopi robotitzat.

A l'interior, aquest instrument inclou un microscopi motoritzat (Fig. 2.14). El microscopi és la part verda i la platina robotitzada la groga. La zona taronja és d'entrada i sortida de les extensions i la part blava són els eixos que es van movent per anar canviant les extensions.

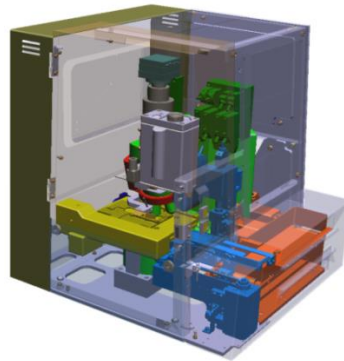


Figura 2.14: Interior del CellaVision®DM96.

A més del microscopi, aquest sistema inclou una càmera i un ordinador que conté el software d'adquisició i classificació (Fig. 2.15).

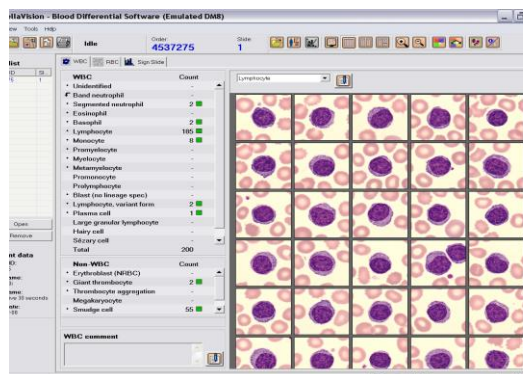


Figura 2.15: Software d'adquisició i preclassificació.

Després de la preclassificació, el facultatiu pot confirmar o reclassificar les cèl·lules (Fig. 2.16).

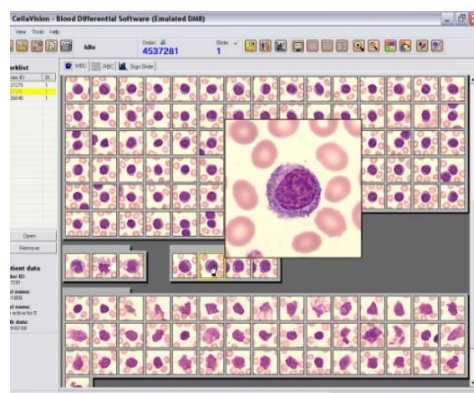


Figura 2.16: Supervisió facultativa de la preclassificació.

El CellaVision®DM96 ha estat estudiat en diferents treballs que avaluen la concordança entre l'anàlisi cel·lular automatitzada d'aquest instrument amb la manual realitzada pel facultatiu:

- Kratz *et al.* avaluen el CellaVision® DM96 obtenint uns valors de correlació entre la preclassificació automàtica i el recompte manual entre 0,67 i 0,96. La sensibilitat se situa entre el 95 i el 100 % i l'especificitat entre el 88 i el 97 %, depenent de l'anormalitat. Conclouen que aquest aparell obté un rendiment similar a l'obtingut de forma manual (49).
- Cornet *et al.* analitzen mostres de 440 pacients utilitzant el CellaVision®DM96. Obtenen un eficiència del 95 % per a la preclassificació automàtica, la qual augmenta fins a un 98 % després de la reclassificació manual per part del facultatiu. La correlació entre la metodologia manual i l'automàtica és excel·lent per als neutròfils, limfòcits i eosinòfils, acceptable per als granulòcits immadurs, eritroblasts i basòfils i baixa per als monòcits. En aquest article, algunes neoplàsies limfoides tipus B són avaluades, però el DM96 no és capaç de classificar correctament les mostres anòmales. La correcta classificació es realitza de forma manual per part del facultatiu mitjançant la pantalla i el software del DM96 (50).
- Briggs *et al.* comparen l'anàlisi morfològica cel·lular realitzada per diferents operadors utilitzant el CellaVision® DM96 amb l'anàlisi manual feta amb microscopi. L'exactitud de la preclassificació automàtica és del 89 %. Assenyalen que la classificació del DM96 en comparació amb el mètode de referència mostra diferències en el cas dels basòfils i els granulòcits immadurs. Això és degut al baix nombre de cèl·lules comptades, ja que aquests elements estan en proporcions molt petites a sang perifèrica. La correlació millora en el cas dels granulòcits immadurs quan tots ells s'analitzen en un mateix grup, ja que també és un error freqüent de l'observador humà equivocar-se entre les diferents formes maduratives dels elements de la sèrie granulocítica (51).
- Merino *et al.* avaluen la preclassificació automàtica del CellaVision® DM96 i la subsegüent reclassificació facultativa. Obtenen excel·lents valors de correlació entre aquest dispositiu i la microscòpia convencional per a neutròfils, limfòcits, monòcits i blasts, i acceptables valors per a eosinòfils, basòfils i cèl·lules plasmàtiques. A més a més, la correlació de la reclassificació és molt bona per a promielòcits i mielòcits, intermèdia per a limfòcits reactius i eritroblasts i baixa per a metamielòcits. Les cèl·lules limfoides anormals només són reclassificades pel facultatiu, ja que el DM96 no és capaç de fer-ho (46).

Les conclusions de tots els treballs publicats són la bona correlació i concordança del nou mètode d'anàlisi automatitzat d'imatges respecte al mètode tradicional al microscopi, tot i que destaquen la necessitat de la validació i revisió dels resultats per un facultatiu expert. Fan èmfasi en la reducció del temps d'anàlisi millorant el flux de treball, l'eficiència i la qualitat, amb els beneficis econòmics que tot això comporta. Alhora destaquen les avantatges que suposa poder crear bases d'imatges per al seguiment de pacients amb una malaltia hematològica i la seva resposta al tractament, així com per a l'activitat docent (46,50,51).

Per últim, un nou analitzador basat en la nova tecnologia Bloodhound®, el sistema Cobas® m 511 (Fig. 2.17), ha estat recentment aprovat per la *Food and Drug Administration* (març 2018) i combina un analitzador de morfologia digital, un comptador de cèl·lules i un classificador en un instrument aerodinàmic, que prepara, tenyeix i analitza el frotis de sang perifèrica.



Figura 2.17: Cobas ® m511 amb tecnologia Bloodhound.

Mitjançant el recompte, la identificació, l'aïllament i la classificació dels leucòcits, hematies i plaquetes, presenta les imatges digitals de tots aquests tipus de cèl·lules a l'estació de visualització (Fig. 2.18) incloent els paràmetres de l'hemograma i el recompte leucocitari diferencial de cinc grups (neutròfils, monòcits, limfòcits, basòfils i eosinòfils), així com les imatges de cèl·lules anormals del pacient per tal de la seva posterior identificació i classificació. A l'actualitat es troba disponible en alguns països, però a Espanya encara no (52).

In Process Analyzers Results Archives QC Configuration Maintenance Help Bloodhound 1 OK 10 Bloodhound 2 OK

Welcome Lab Administrator 03/26/15 12:52 PM Log Out Reche

Close Save Margaret M. Taylor Age: 70 years Received: 03/26/15 11:22 AM 11:58 AM (Bloodhound 1) Initial Review Autodiff Contact Action

Medical Record #: 4352313 Sex: Female Location: Slide was discarded Sample is awaiting review. (Discard these results) (Mark for Rerun) Print Release

Accession #: SL31987 Physician: Diagnosis:

Overview Results Report Reclassified

WBC

Parameter	Result	Units
WBC	17.40 (H)	$\times 10^9/\mu\text{L}$
RBC	3.24 (L)	$\times 10^6/\mu\text{L}$
HGB	9.1 (L)	g/dL
HCT	29.4 (L)	%
MCV	87.7	fL
MCH	28.1	pg
MCHC	32.0	g/dL
RDW	12.4	%
RDW-SD	39.2	fL
PLT	21 (L)	$\times 10^9/\mu\text{L}$
MPV	10.2	fL
%NRBC	2.4	/100 WBC
#NRBC	0.4	$\times 10^9/\mu\text{L}$

Cell Type % $\times 10^9/\mu\text{L}$

Unclassified	12.3	2.1
Neutrophil	35.2 (L)	6.1
Lymphocyte	4.5 (L)	0.8 (L)
Monocyte	48.0 (H)	8.4 (H)
Eosinophil	0.0	0.0
Basophil	0.0	0.0

WBC Messages:

- Bias Suspected
- Lymphopenia
- Monocytosis
- Unclassified Cells Present

Unclassified (12.3%)

Neutrophil (35.2%) Unsort

Figura 2.18: Estació de visualització del Cobas® m511.

2.6 Neoplàsies limfoides i classificació

Les neoplàsies limfoides són un conjunt de malalties neoplàsiques derivades dels limfòcits i els seus precursors que es desenvolupen en els òrgans hematopoètics i en el sistema limfàtic. La biologia i manifestacions clíniques d'aquest tipus de malalties són molt diverses degut a la gran complexitat cel·lular i funcional de les poblacions on s'originen i als diferents mecanismes patogènics que contribueixen al seu desenvolupament. Aquestes neoplàsies poden tenir lloc de dues maneres: com una leucèmia (cèl·lules tumorals al moll de l'os i a la sang) i/o un limfoma (massa sòlida de cèl·lules tumorals) (53). Són malalties que succeeixen a totes les etapes de la vida i la seva freqüència augmenta amb l'edat.

La seva detecció precoç i un tractament immediat són claus per a la supervivència dels pacients (34). Constitueixen un grup heterogeni de malalties amb característiques clíniques i biològiques diferenciades, tot i que en ocasions compartides, fet que dificulta el seu correcte diagnòstic. De cara al diagnòstic diferencial d'aquestes entitats és imprescindible l'anàlisi citològica, immunofenotípica, citogenètica i molecular, així com l'estudi histològic del moll de l'os i dels ganglis limfàtics (54,55), ja que no hi ha un únic marcador que sigui específic d'una neoplàsia.

Les neoplàsies limfoides de cèl·lules B madures representen el 90 % de les neoplàsies limfoides a tot el món, mentre que les T són menys freqüents. La classificació de les neoplàsies limfoides està basada en la utilització de tota la informació disponible per a definir les diferents malalties. L'aproximació multiparamètrica per a la classificació adoptada per l'Organització Mundial de la Salut ha estat validada en estudis internacionals ja que és altament reproduïble i reforça la interpretació dels estudis clínics. Considerant la cèl·lula d'origen (precursora o madura B o T), estableix els següents grups (38):

1. **Neoplàsies limfoides de precursors de cèl·lules B** (Cèl·lules B immadures): Leucèmies limfoides agudes B
2. **Neoplàsies limfoides de precursors de cèl·lules T** (Cèl·lules T immadures): Leucèmies limfoides agudes T
3. **Neoplàsies de cèl·lules B madures:** Leucèmia limfàtica crònica (LLC), Leucèmia prolimfocítica B (LPB), Limfoma de la zona marginal esplènica (LZME), Tricoleucèmia (HCL), Limfoma limfoplasmocític, Mieloma de cèl·lules plasmàtiques (CP), Limfoma MALT, Limfoma fol·licular (LF), Limfoma de cèl·lules del mantell (LCM), Limfoma difús de cèl·lules grans B, Limfoma de Burkitt.
4. **Neoplàsies de cèl·lules T madures i NK:** Leucèmia de limfòcits grans granulars T (LLGG), Leucèmia prolimfocítica T (LPT), Leucèmia o limfoma de cèl·lules T, síndrome de Sézary (SS).

Com s'ha comentat en apartats anteriors, la sang que circula als vasos sanguinis (sang perifèrica) conté diferents sèries cel·lulars: mieloide (basòfils, eosinòfils, neutròfils), monocítica (monòcits), limfoide (limfòcits), eritroide (eritròcits) i plaquetària (plaquetes). Totes elles es generen al moll de l'os a partir de la diferenciació i maduració de cèl·lules mare. En el cas dels limfòcits, hi ha una fase de maduració

adicional que té lloc als ganglis limfàtics, on aquestes cèl·lules adquireixen les seves capacitats defensives dins del sistema immunitari.

En condicions fisiològiques, hi ha un balanç autocontrolat d'aquest procés (hematopoesi) i a la sang circulen cèl·lules normals en quantitats adequades per a exercir les seves funcions. En situacions anormals, poden succeir alteracions en el procés, de manera que les cèl·lules pateixen mutacions i proliferacions descontrolades invasives (càncer). Quan aquest procés succeeix amb les cèl·lules immadures (blasts) al moll de l'os, aquestes passen massivament a sang perifèrica desplaçant a les cèl·lules normals, produint les malalties conegudes com a leucèmies. Quan el descontrol succeeix als limfòcits del sistema limfàtic, es produeixen els denominats limfomes, que també impliquen, en un elevat percentatge de casos, la presència de cèl·lules anormals a sang perifèrica (34).

Els limfòcits B són responsables de la immunitat humoral i poden fabricar anticossos o immunoglobulines després de la presentació antigènica per una cèl·lula presentadora d'antigen (macròfags, cèl·lules fol·liculars, cèl·lules dendrítiques). A més a més, posseeixen immunoglobulines de membrana que constitueixen el seu marcador fenotípic. La fabricació dels anticossos es realitza a nivell dels òrgans limfoïdes secundaris on els limfòcits es transformen en plasmòcits.

Els limfòcits T madurs expressen el receptor de membrana CD3 i participen a la resposta immune humoral estimulant o frenant la producció d'anticossos per part dels limfòcits B. A més a més, estan implicats a la immunitat cel·lular secretant citocines.

A la sèrie limfoïde, la cèl·lula més immadura és el limfoblast. S'origina a partir d'una cèl·lula mare al moll de l'os i la seva morfologia és la següent: té un diàmetre de 8 a 20 μm , un nucli de perfil rodó de cromatina poc condensada, en ocasions immadura, i amb nuclèols visibles. El citoplasma és molt escàs i basòfil i no té granulació. La seva diferenciació respecte a la cèl·lula més immadura de la sèrie mieloïde o mieloblast és, en ocasions, molt difícil. La detecció de cèl·lules blàstiques (limfoblasts o mieloblasts) a la sang indica una proliferació descontrolada d'aquestes cèl·lules al moll de l'os, fet que succeeix en els diferents tipus de leucèmies agudes (38).

Les anomalies morfològiques de la sèrie limfoïde B i T són molt variades: (1) augment de la mida, (2) immaduresa de la cromatina del nucli incloent la presència de nuclèol visible o la falta de condensació de la cromatina, (3) irregularitat del perfil del nucli o lobulació, (4) intensitat de la basofília (des de blau pàl·lid a molt basòfil), i (5) irregularitats en el perfil del citoplasma, que pot ser més o menys abundant. Els limfòcits anormals B o T acostumen a mostrar característiques morfològiques diferents en funció del tipus de limfoma/leucèmia que presenti el pacient.

En els següents paràgrafs, es descriuen les característiques morfològiques més rellevants dels limfòcits anormals presents a les diferents neoplàsies limfoides (imadures i madures) que han estat objecte d'estudi de la present *Tesi Doctoral*.

2.6.1 Leucèmia limfoblàstica aguda

La leucèmia limfoblàstica aguda (LLA) és una neoplàsia de cèl·lules limfoides precursors o imadures (limfoblasts) de tipus B o T. La LLA tipus B afecta a sang perifèrica i moll de l'os principalment en nens menors de 10 anys, però pot aparèixer en edats posteriors i a l'edat adulta. En SP s'observen cèl·lules blàstiques limfoides (BL) com la de la Figura 2.19, les quals es caracteritzen per tenir una mida petita-mitjana, un nucli de cromatina laxa i immadura amb algun nuclèol visible. El citoplasma és en general escàs. Habitualment no s'observa granulació al seu interior.

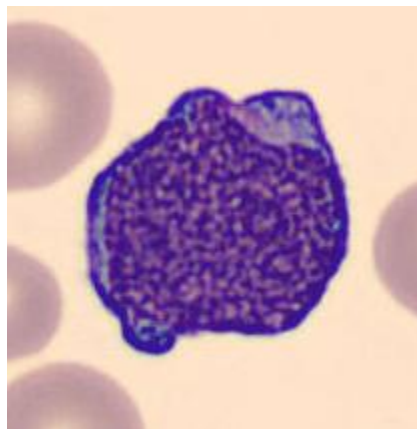


Figura 2.19: Cèl·lula blàstica limfoide.

2.6.2 Leucèmia limfàtica crònica

En el grup de limfomes/leucèmies de tipus B, la LLC és la malaltia més freqüent i apareix en persones grans (edat mitja de 65 anys). És una neoplàsia de baix grau caracteritzada per l'acumulació de cèl·lules B monoclonals amb un fenotip que coexpressa CD5, CD23 i CD43. Citològicament, els limfòcits són de mida petita, amb un citoplasma molt escàs i tenen una cromatina madura i condensada (“*grumelée*”) (Fig. 2.20). En aquesta malaltia pot veure's a més a més una petita proporció de prolimfòcits en sang. L'observació d'un major percentatge de prolimfòcits a la LLC està directament relacionat amb una possible transformació a leucèmia prolimfocítica, i amb determinades alteracions genètiques, com la trisomia del cromosoma 12 (56).

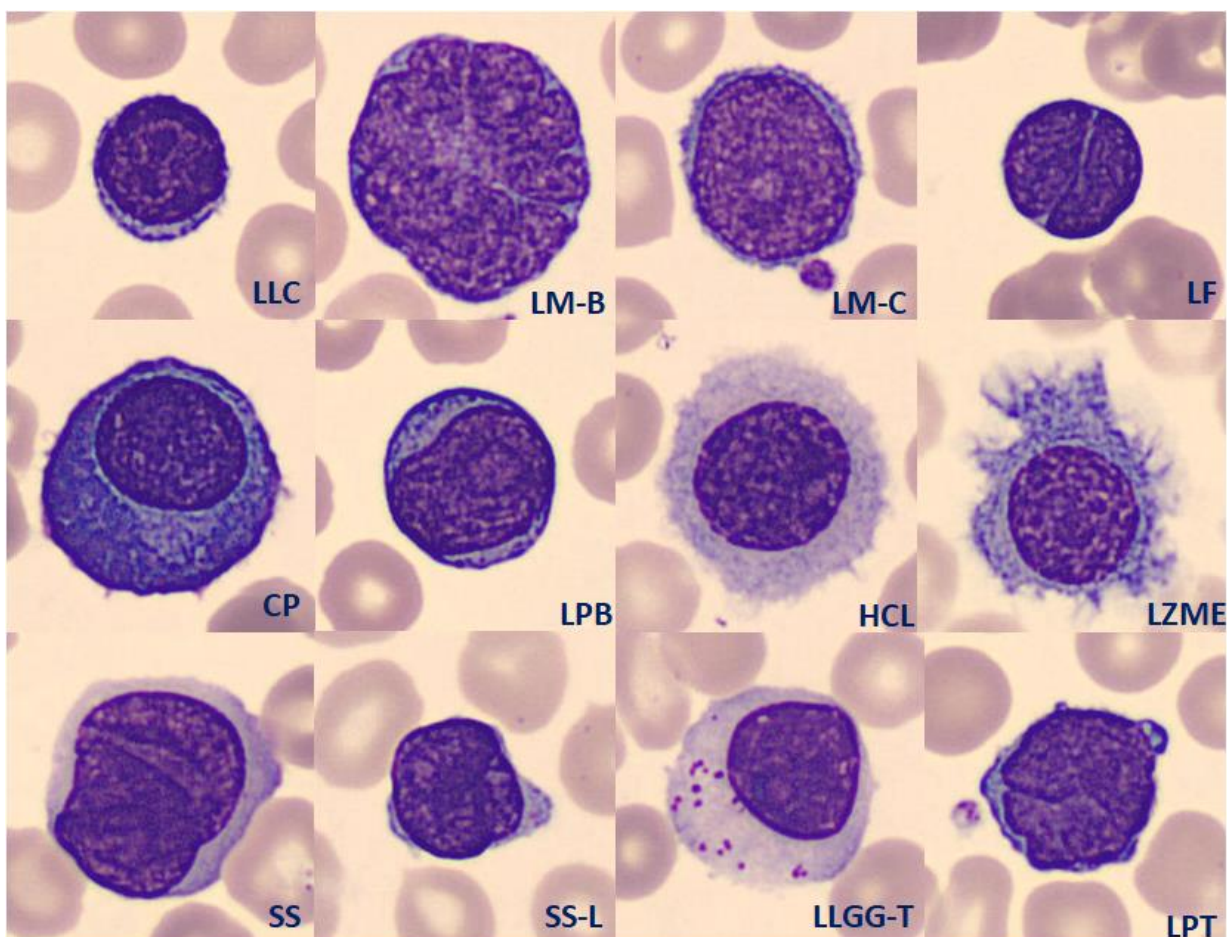


Figura 2.20: Cèl·lules limfoides anormals característiques de les Neoplàsies Limfoides B i T.

2.6.3 Leucèmia prolimfocítica B

El prolimfòcit B comparteix característiques morfològiques del limfoblast, ja que el seu nucli també té un perfil rodó amb un nuclèol central, en aquest cas poc aparent (Fig. 2.20). Conté menor citoplasma respecte al limfoblast i la cromatina és menys condensada. A la denominada LLC atípica, els limfòcits tenen una cromatina nuclear menys condensada i certa irregularitat nuclear. Si el percentatge de prolimfòcits en sang supera el 55 % es tracta d'una transformació a leucèmia prolimfocítica B (LPB).

2.6.4 Tricoleucèmia

A la tricoleucèmia o *hairy cell leukemia* (HCL), les cèl·lules limfoides tenen una major mida respecte als limfòcits normals, el nucli pot ser ovalat, rodó o parcial o completament bilobulat. El citoplasma és abundant, amb una tonalitat blava-gris o pàl·lida, d'aspecte hialí i prolongacions velloses a mode de "pèls" al voltant de tota la cèl·lula (Fig. 2.20). Aquesta neoplàsia de baix grau presenta afectació a sang, moll d'os i la polpa vermella de la melsa. El patró immunofenotípic característic és positivitat per CD103, CD20, CD22 i CD11c i negativitat per CD5, CD10 i CD23. Però no només a l'HCL es poden veure limfòcits d'aspecte vellós, també en el limfoma de la zona marginal esplènica que veurem a continuació.

2.6.5 **Limfoma de la zona marginal esplènica**

Aquest tipus de limfoma es considera de baix grau i s'ha relacionat amb l'existència d'algunes infeccions, com la del virus de l'hepatitis C i processos d'autoimmunitat. És una malaltia poc freqüent i pot afectar a la melsa, al moll de l'os i la sang. Les cèl·lules limfoides anormals també són velloses, però les prolongacions citoplasmàtiques es troben distribuïdes majoritàriament als extrems (Fig. 2.20), a diferència de l'HCL, on els limfòcits mostren prolongacions al voltant de tot el perímetre de la cèl·lula. El nucli pot ser rodó, ovalat, lleugerament o completament bilobulat.

2.6.6 **Limfoma fol·licular**

El limfoma fol·licular (LF) és de baix grau derivat de cèl·lules B madures que han adquirit, en la majoria de casos, la translocació $t(14;18)(q32;q21)$ que causa la sobreexpressió del gen BCL2, el qual bloqueja l'apoptosi de les cèl·lules B. Aquest limfoma és una de les formes més comuns de limfoma B als països occidentals. Citològicament, les cèl·lules limfoides són majoritàriament petites, amb poc citoplasma dèbilment basòfil i amb un nucli de perfil irregular (Fig. 2.20). A vegades les cèl·lules tenen una morfologia més variada (pleomòrfiques) i poden tenir un nuclèol visible. Fenotípicament expressen marcadors de línia B (com CD19, CD20 i CD22) i els marcadors centrefol·liculars, CD10 i BCL6.

2.6.7 **Limfoma de cèl·lules del mantell**

El limfoma de cèl·lules del mantell (LCM) és una neoplàsia molt agressiva i generalment té mal pronòstic en comparació amb altres leucèmies. Representa el 5 - 10 % dels limfomes de cèl·lula B madura. Genèticament es caracteritza per la translocació $t(11;14)(q13;q32)$ que comporta la desregulació del gen CCND1 que codifica per la ciclina D1. La detecció de la seva sobreexpressió constitueix una de les proves més importants en el diagnòstic actual de la malaltia.

Citològicament existeixen dues formes de LCM: la clàssica (LM-C) i la blàstica (LM-B). La clàssica es defineix per tenir cèl·lules de mida petita-mitjana amb poc citoplasma i un nucli irregular amb una cromatina immadura, semblant a la del limfoblast, amb un nuclèol visible. En canvi, la forma blàstica, es caracteritza per tenir una mida més gran i un curs clínic molt més agressiu (Fig. 2.20).

2.6.8 **Mieloma múltiple amb cèl·lules plasmàtiques circulants a sang perifèrica**

Les cèl·lules plasmàtiques (CP) anormals es poden veure a la sang dels pacients amb mieloma múltiple, o amb una leucèmia de cèl·lules plasmàtiques, i tenen una mida més gran que els limfòcits normals (Fig. 2.20). El seu nucli és excèntric, de perfil rodó i ovalat, de cromatina nuclear condensada i el citoplasma és molt basòfil, amb una zona més clara adjacent al nucli degut a que contenen un aparell de Golgi molt desenvolupat.

2.6.9 Leucèmia prolimfocítica T

Els pacients amb una leucèmia prolimfocítica T (LPT) mostren uns limfòcits molt semblants als prolimfòcits B ja descrits, però amb major basofília citoplasmàtica i major irregularitat en el perfil nuclear (Fig. 2.20). En un 25 % dels casos, el nuclèol no és visible al microscopi òptic (variant de cèl·lules petites) (57).

2.6.10 Síndrome de Sézary

La síndrome de Sézary (SS) és un limfoma T amb limfòcits en SP coneguts com a cèl·lules de Sézary que es caracteritzen per tenir una mida gran, excepte la variant petita o de Lutzner (SS-L), i un nucli de perfil irregular de cromatina madura que presenta plects similars a les circumvolucions cerebrals (“nucli d’aspecte cerebriforme”) (Fig. 2.20). Els autoanàlitzadors que s’utilitzen als laboratoris no donen alarmes quan existeixen aquestes cèl·lules en una mostra de sang, ja que per les seves característiques de maduresa, les confonen amb limfòcits normals.

2.6.11 Leucèmia de limfòcits grans granulars tipus T

La leucèmia de limfòcits grans granulars T (LLGG) es caracteritza per la presència d’aquests limfòcits a SP degut a una proliferació maligna dels mateixos. Tenen una morfologia com la descrita pels limfòcits normals de mida superior, però contenen al seu citoplasma grànuls petits i prominents (Fig. 2.20). Els limfòcits grans granulars comprenen un 10 – 20 % dels limfòcits d’una persona sana. Tot i així, si circulen a la SP amb percentatges superiors, s’han de realitzar altres estudis complementaris per a descartar clonalitat.

Capítol 3

3 Metodologia de recerca

3.1 Introducció

Aquest Capítol 3 detalla els materials i mètodes utilitzats a la present *Tesi Doctoral*, encaminats cap a una caracterització morfològica dels diferents tipus de cèl·lules limfoides de sang perifèrica a partir de l'obtenció i anàlisi de descriptors geomètrics, de color i de textura obtinguts mitjançant processament digital d'imatges.

En concret, el Capítol 3 descriu les diferents etapes del processament digital d'imatges necessàries per a l'obtenció dels descriptors quantitius així com l'anàlisi estadística realitzada per tal de valorar la seva capacitat discriminatòria entre els diferents grups cel·lulars estudiats en aquesta Tesi.

La secció 3.2 presenta el primer pas del processament digital d'imatges i és l'adquisició de les imatges obtingudes de frotis de pacients amb diferents malalties hematològiques mitjançant dos sistemes: automàtic (sistema DM96) i manual (sistema BX43). La secció 3.3 detalla l'algoritme de segmentació incloent el preprocessament, la metodologia de segmentació per a separar les regions d'interès de la cèl·lula i el postprocessament. La secció 3.4 detalla els diferents tipus de descriptors (geomètrics, de color i textura) extrets sobre el nucli, citoplasma i la cèl·lula completa, els quals són essencials per a la caracterització morfològica de les cèl·lules dels diferents grups d'interès. La secció 3.5 inclou l'últim pas del processament digital d'imatges i és la classificació dels diferents grups de cèl·lules limfoides mitjançant els descriptors prèviament seleccionats. Per últim, la secció 3.6 conté els mètodes estadístics utilitzats a l'anàlisi dels descriptors. La Figura 3.1 mostra un esquema de totes les etapes que s'han dut a terme en la present *Tesi Doctoral*.

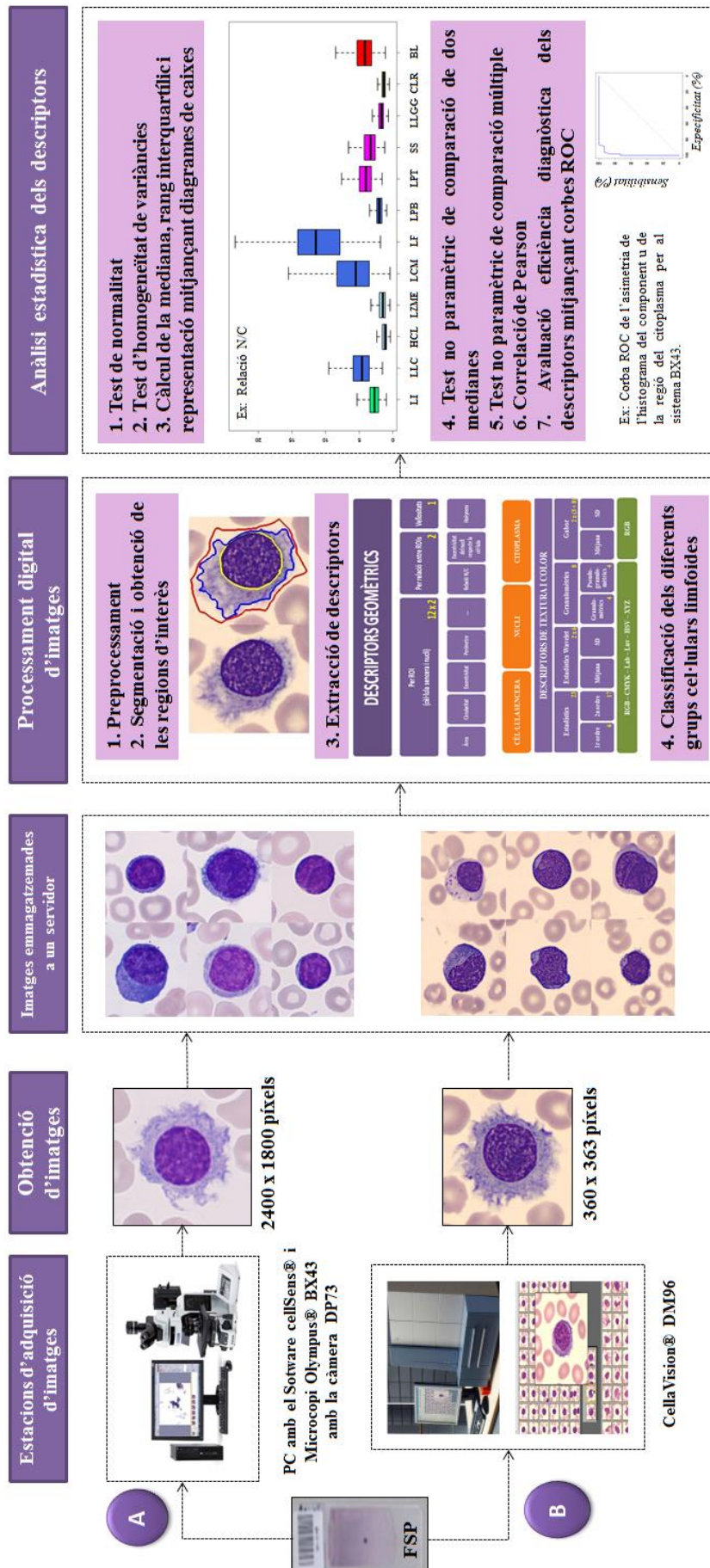


Figura 3.1: Esquema del sistema per a l'anàlisi morfològica dels frotis de sang perifèrica (FSP) a les dues plataformes.

3.2 Adquisició i emmagatzematge de les imatges

Es tracta de la primera etapa en el processament digital d'imatges (PDI) i és essencial per a obtenir bons resultats en etapes posteriors del PDI. En el cas de la present recerca, l'adquisició d'imatges digitals de cèl·lules individuals s'ha realitzat a partir de frotis de sang perifèrica obtinguts de la Unitat de Citologia del Laboratori Core de l'Hospital Clínic de Barcelona. Totes les imatges utilitzades durant la recerca corresponen a pacients d'ambdós sexes amb edats compreses entre els 16 i els 90 anys. Els diagnòstics dels pacients amb malalties hematològiques han estat confirmats pel Servei d'Hematologia Clínica de l'Hospital Clínic de Barcelona segons la classificació OMS 2016 (38).

Per al seu estudi morfològic, prèviament les mostres de sang, recollides en EDTAK₃ han estat analitzades a l'equip Advia 2120i® (Siemens Healthcare Diagnostics, Tarrytown, EEUU) obtenint així els paràmetres hematològics bàsics i els frotis s'han obtingut mitjançant l'extensor - tenyidor SP1000i (Sysmex, Kobe, Japan), el qual utilitza els colorants May Grünwald-Giemsa. Com s'ha explicat al Capítol anterior, la tinció de May Grünwald-Giemsa és la més emprada degut a les seves propietats, les quals ressalten els components basòfils de les cèl·lules sanguínies (34). En el cas dels limfòcits, aquests es veuen de color porpra.

La tinció és una de les tècniques de microscòpia òptica més emprades a l'hora de millorar el contrast de les cèl·lules ja que aquestes són incolores. Cal tenir en compte que la tinció també pot introduir artefactes, detalls estructurals que no representen les característiques legítimes de la mostra. Per tant, la tècnica de contrast, en aquest cas la tinció, ha de ser curosament escollida i ha de ser la que millor s'ajusti a les característiques a estudiar, ja que d'aquesta forma s'evitaran imatges degradades o amb excessiu soroll que comportarien errors en el procés de segmentació i classificació.

Els frotis de sang perifèrica han estat escollits per les seves característiques morfològiques per part del citòleg expert i corresponents a pacients diagnosticats d'una neoplàsia limfoide B o T, una leucèmia aguda o una infecció vírica, bacteriana o parasitària. Els diferents diagnòstics s'han confirmat mitjançant la integració dels resultats d'una sèrie de proves complementàries: citometria de flux, citogenètica i biologia molecular. A continuació es detallen tots els grups cel·lulars estudiats a la present *Tesi Doctoral*.

3.2.1 Grups de cèl·lules d'interès

Per a la present *Tesi Doctoral* s'han considerat els següents grups:

- 1) cèl·lules limfoides de pacients sense patologia (normals)
- 2) cèl·lules limfoides reactives de pacients diagnosticats d'una infecció vírica (virus de l'Epstein-Barr, Cytomegalovirus, virus del Dengue, virus de l'Hepatitis, etc), bacteriana o parasitària (*Plasmodium* sp.).
- 3) cèl·lules de pacients diagnosticats de limfomes B i T
- 4) cèl·lules blàstiques de pacients diagnosticats de leucèmies agudes limfoides

En concret, s'han analitzat les característiques morfològiques de les cèl·lules característiques de les següents entitats (Fig. 3.2):

- **Neoplàsies Limfoides B**
 - Leucèmia Limfàtica Crònica (LLC)
 - Limfoma Cèl·lules del Mantell (LCM): variants blàstica (LM-B) i clàssica (LM-C)
 - Limfoma Fol·licular (LF)
 - Mieloma Múltiple amb cèl·lules plasmàtiques (CP)
 - Leucèmia Prolimfocítica B (LPB)
 - Tricoleucèmia (HCL)
 - Limfoma de la Zona Marginal Esplènic (LZME)

- **Neoplàsies Limfoides T**
 - Síndrome de Sézary: variants petita (Lutzner) (SS-L) i clàssica (SS)
 - Leucèmia de Limfòcits Grans Granulars tipus T (LLGG)
 - Leucèmia Prolimfocítica T (LPT)

- **Leucèmies agudes**
 - Leucèmia limfoblàstica aguda (BL)

- **Infeccions**
 - Cèl·lules limfoides reactives (CLR)

- **Grup control**
 - Limfòcits normals (LI)

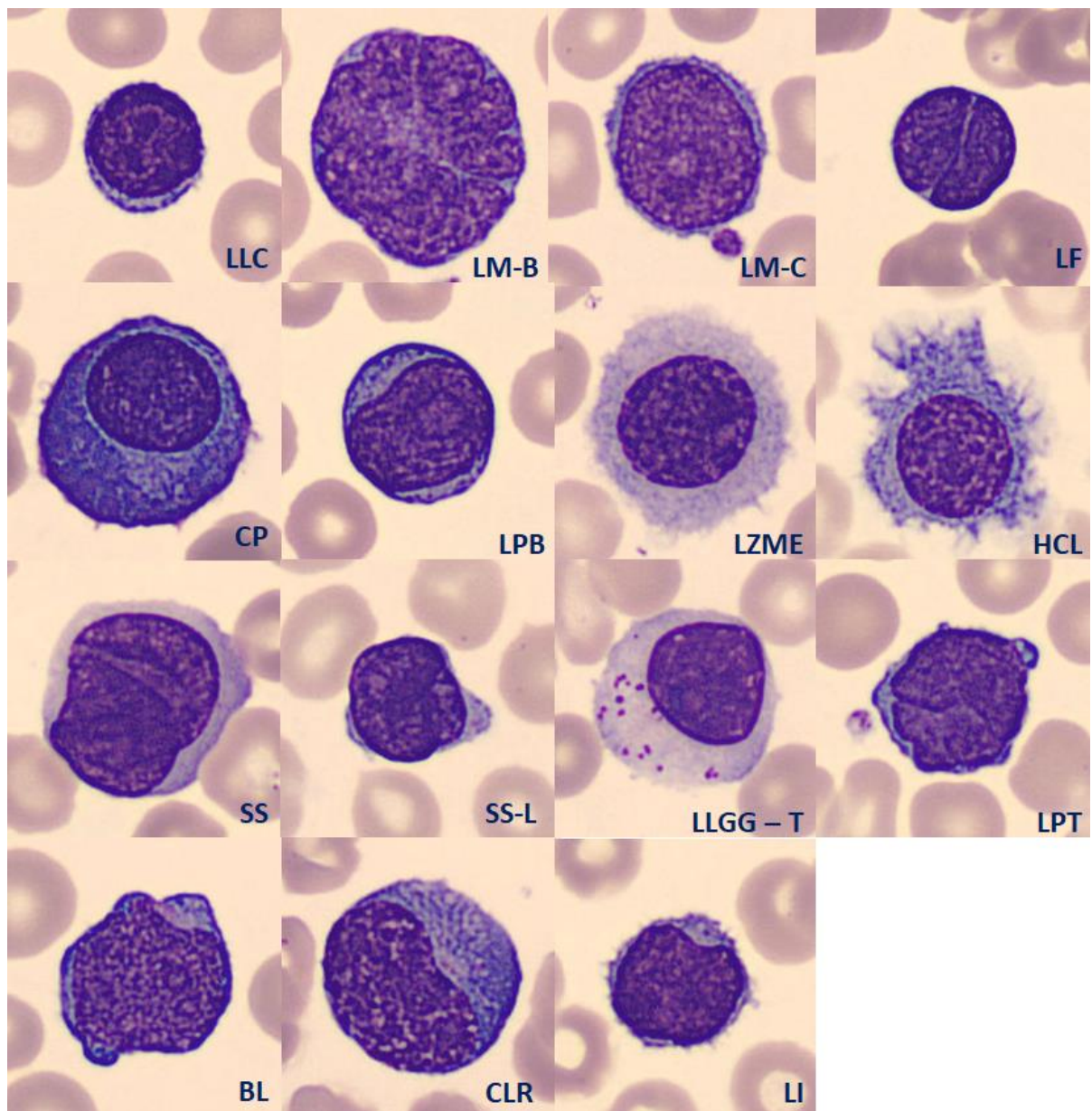


Figura 3.2: Conjunt de cèl·lules limfoides estudiades a la present *Tesi Doctoral*. Imatges obtingudes amb el CellaVision DM96 a x1.000 augments.

3.2.2 Sistemes d'adquisició de les imatges digitals

Anem a suposar que el frotis ha estat preparat adequadament seguint uns processos estandarditzats i ben coneguts, tal i com s'ha descrit al Capítol 2. Seguidament les imatges es capturen mitjançant un microscopi òptic. Mitjançant una combinació de lents (objectiu i ocular), el microscopi projecta una imatge en un pla en color visible per a l'ull humà. Si aquesta imatge es captura mitjançant una càmera, s'obté una fotografia digital. A la Figura 3.3 es mostra un microscopi òptic amb una càmera muntada a sobre.



Figura 3.3: Microscopi òptic Olympus BX43 amb càmera DP73 utilitzat en aquesta Tesi.

Una càmera digital disposa d'una malla quadriculada de sensors fotosensibles coincident amb el pla focal, els quals generen un senyal elèctric (voltatge) quantificable proporcional a la quantitat de llum que reben durant un interval de temps d'activació. L'espai ocupat per cada sensor es denomina píxel. El nombre total de sensors o píxels defineix la resolució de la càmera.

Per a l'adquisició de les imatges de sang s'utilitza majoritàriament la microscòpia òptica amb els objectius que utilitzen oli d'immersió de 50 i 100 magnificacions (500 i 1000 en total, respectivament) seguida d'una càmera digital amb sensors CCD acoblada (58). L'adquisició de les imatges es pot realitzar de forma manual amb un microscopi convencional o bé de forma automatitzada amb microscopis robotitzats com el CellaVision® DM96, el qual és utilitzat en aquesta Tesi.

Bàsicament, hi ha tres tipus d'imatges digitals:

- Imatges binàries: la intensitat de cada píxel només pot prendre el valor 0 (negre) o 1 (blanc).
- Imatges en escala de grisos: la imatge és monocromàtica, contenint ombres de diferent intensitat de gris, sense cap color. La intensitat de cada píxel pren qualsevol valor enter dins d'un rang definit.
- Imatges en color: cada píxel està representat en general per tres valors que descriuen la intensitat i la cromaticitat associada al color.

Les imatges digitals es poden expressar quantitativament d'una forma organitzada i pràctica. Una imatge en escala de grisos consisteix en una matriu rectangular formada per $M \times N$ elements quadrats de la mateixa mida on M és el nombre de columnes i N el nombre de files (Figura 3.4). Cada element de la matriu es correspon a un píxel. El nombre de píxels defineix la mida de la imatge. La matriu s'expressa en un sistema de coordenades cartesianes (X, Y) que permet identificar la posició de cada píxel a la imatge.

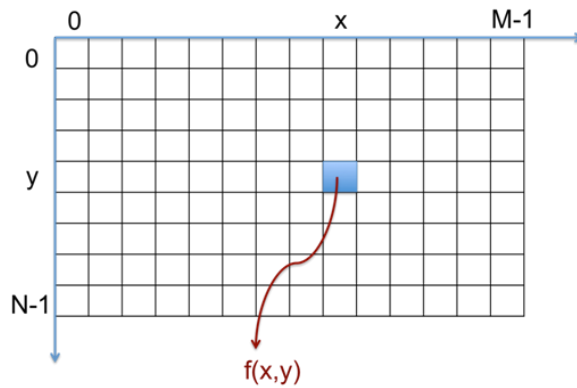


Figura 3.4: La imatge digital estructurada mitjançant una matriu formada per un nombre $M \times N$ de píxels quadrats de la mateixa mida.

Per quantificar la imatge digital, s'introdueix una funció $f(x, y)$, que anomenem intensitat de la imatge en cada píxel (x, y) , la qual assigna un valor. La intensitat depèn de la il·luminació incident en el frotis i de la il·luminació reflectida per les cèl·lules limfoides. La intensitat pren valors positius dins d'un rang, on el valor mínim correspon al negre i el valor màxim correspon al blanc. Tots els valors intermedis representen les possibles ombres grises. Per tant, a partir d'ara la imatge digital estarà representada per la seva funció intensitat $f(x, y)$. Hi ha dues formes molt habituals de representar aquesta funció:

1. Presentar els seus valors numèrics en una matriu:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, M-1) \\ f(1,0) & f(1,1) & \dots & f(1, M-1) \\ \vdots & \vdots & \dots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1, M-1) \end{bmatrix} \quad (3.1)$$

2. Visualització similar a com ho fa un monitor o una fotografia. La intensitat de cada punt és proporcional al valor de la funció en el píxel que conté al punt (Figura 3.5a).

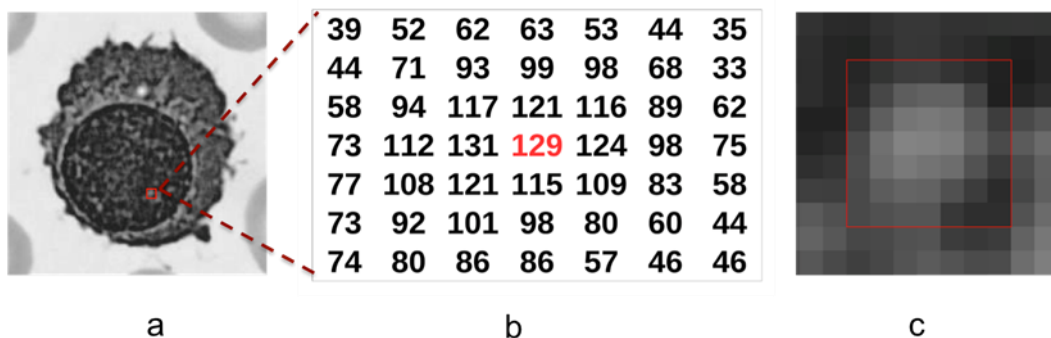


Figura 3.5: Part a: Imatge en escala de grisos d'una cèl·lula plasmàtica. Part b: valors de la intensitat en una escala entre 0 i 255 per a una zona petita del nucli. Part c: Visualització de la intensitat dels píxels corresponents d'aquesta zona.

La Figura 3.5b mostra la matriu d'intensitats per a una petita part del nucli, mentre que la part c mostra la visualització dels píxels amb els diferents nivells d'intensitat. Aquesta imatge conté una àmplia varietat de grisos, de manera que la matriu és força dispersa en els seus valors.

A la pràctica, habitualment es fa un canvi d'escala i una quantificació de la intensitat amb valors positius (per sentit físic). La quantificació agrupa els valors continus de la intensitat en determinats nivells (quants) caracteritzats per valors sencers en l'interval $[0, L-1]$, el qual s'anomena escala de grisos o rang dinàmic. Els sistemes físics d'emmagatzematge informàtic guarden els valors numèrics codificats utilitzant bits. Un bit és una variable binària caracteritzada únicament per dos dígits (0 i 1). D'aquesta manera, qualsevol nombre enter decimal pot expressar-se com una seqüència formada per un nombre determinat de bits. Per exemple, amb tres bits es pot expressar qualsevol valor enter entre 0 (0-0-0) i 7 (1-1-1), és a dir un total de $8 = 2^3$ valors. Amb aquest propòsit, el nombre de nivells d'intensitat de la imatge s'escull com una potència de 2, és a dir: $L = 2^k$.

Habitualment, es diu que la imatge té k bits de resolució en intensitat, o que té una profunditat de k bits. El nombre total de bits és $N \times M \times k$ per una imatge completa en escala de grisos. Un valor bastant típic en analitzadors d'imatges de cèl·lules sanguínies és $k = 8$ bits, el que significa un rang de 0 a 255. Com més gran és el nombre de bits, major és la capacitat de discernir petits canvis en la intensitat i per tant major és la resolució. Per contra, amb el nombre de bits augmenta considerablement la necessitat de memòria per a l'emmagatzematge de les imatges.

La Figura 3.6 mostra el protocol estandarditzat d'adquisició d'imatges utilitzat durant aquesta recerca, on a partir de les mostres dels pacients es prepara de forma automàtica el frotis tenyit amb May Grünwald-Giemsa i les imatges són obtingudes a partir dels dos sistemes següents:

- (1) Adquisició manual amb el microscopi Olympus BX43 (Olympus, Tòquio, Japó) amb la càmera DP73 i el software cellSens® d'Olympus. El microscopi compta per defecte amb un augment de 10 augments (10x). Per a l'adquisició de les imatges s'utilitza un objectiu de 100 augments, (100x), sent l'augment total aplicat sobre les cèl·lules de 1.000x.
- (2) Adquisició automàtica amb l'equip CellaVision® DM96 (CellaVision AB, Lund, Suècia), el qual incorpora un microscopi robotitzat. Per a l'adquisició de les imatges s'utilitza un objectiu de 100 augments, (100x), sent 1.000x l'augment total aplicat sobre les cèl·lules.

A continuació, s'exposen les principals diferències entre les imatges adquirides pel sistema DM96 i les obtingudes mitjançant el sistema BX43. En primer lloc, els dos sistemes proporcionen imatges amb diferent mida. Mentre que la imatge del DM96 té una mida de 360 x 363 píxels (en total 130.680 píxels), la imatge obtinguda per microscòpia òptica manual té una mida de 1.800 x 2.400 (en total 4.320.000 píxels). D'altra banda, la profunditat de bits també és diferent entre els dos tipus d'imatge, sent de 8 bits per a les imatges del DM96 i de 12 bits per a les imatges del sistema BX43. Donat que l'algoritme desenvolupat per Alférez *et al.* (59) treballa sobre les dimensions del DM96, s'han reajustat una sèrie de paràmetres perquè la segmentació es pugui dur a terme correctament amb el sistema BX43.

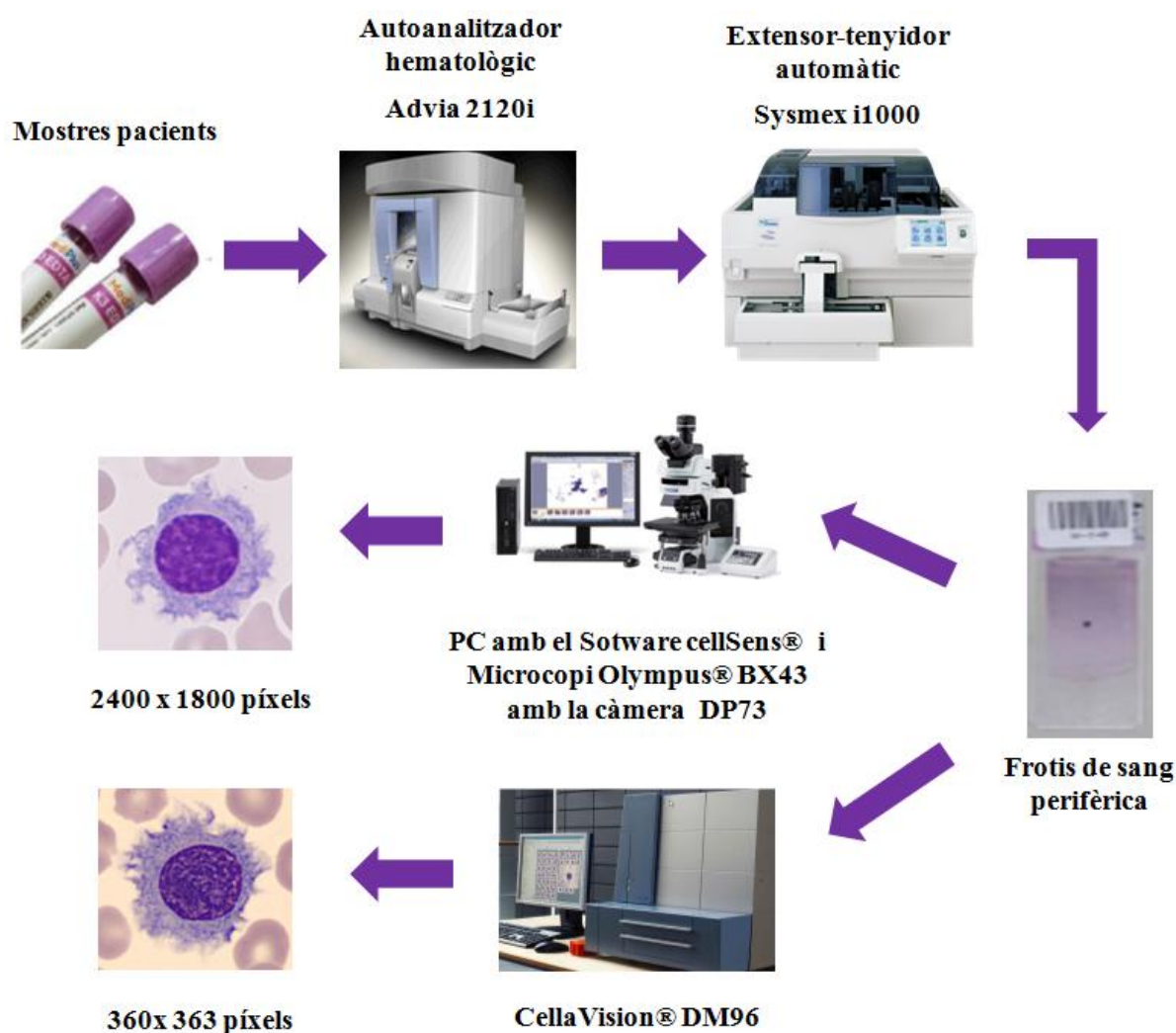


Figura 3.6: Protocol estandarditzat d'adquisició d'imatges utilitzat durant la recerca.

El format d'emmagatzematge de les imatges també difereix entre ambdós sistemes. Mentre que el sistema DM96 utilitza el format JPG (comprimit), les imatges de l'estudi es troben en format TIFF (sense compressió). A més, el sistema DM96 és completament automàtic. D'aquesta manera el procés és molt més àgil i el facultatiu només ha de visualitzar la imatge obtinguda en pantalla per tal de confirmar el diagnòstic. A més a més, el sistema DM96 centra totes les imatges en la cèl·lula detectada. Per tant, quan es troba amb una agrupació de cèl·lules, en comptes de fer una fotografia que les inclogui totes, adquireix imatges on al centre apareix cadascuna d'elles de forma individualitzada.

En canvi, les imatges del sistema BX43 han estat adquirides manualment a la present recerca, per la qual cosa la cèl·lula mai apareixerà perfectament centrada i en una única imatge podran observar-se agrupacions de cèl·lules. L'algorisme d'Alfárez *et al.* (59), segmenta únicament la cèl·lula central. Per tant, a les imatges de microscòpia preses de forma manual on apareguin diverses cèl·lules només se'n segmenta una. Finalment, el sistema DM96 representa un cost monetari molt més elevat que el microscopi BX43.

3.2.3 Creació de la base de dades

Un cop les imatges han estat adquirides, aquestes s'han emmagatzemat a un servidor de la Universitat Politècnica de Catalunya i la seva ruta a una base de dades. Per a l'automatització del processament de les imatges digitals s'ha creat una base de dades MySQL® per a organitzar la informació (dades demogràfiques, analítiques, resultats citogenètics, any del diagnòstic, entitat diagnosticada, número de petició i rutes de les imatges) (Fig. 3.7). Aquesta metodologia ens ha permès poder integrar totes les dades recollides de cada pacient i també per malaltia en paquets de càlcul i realitzar una posterior anàlisi de grans volums d'informació amb el paquet de software *R Commander* (60).

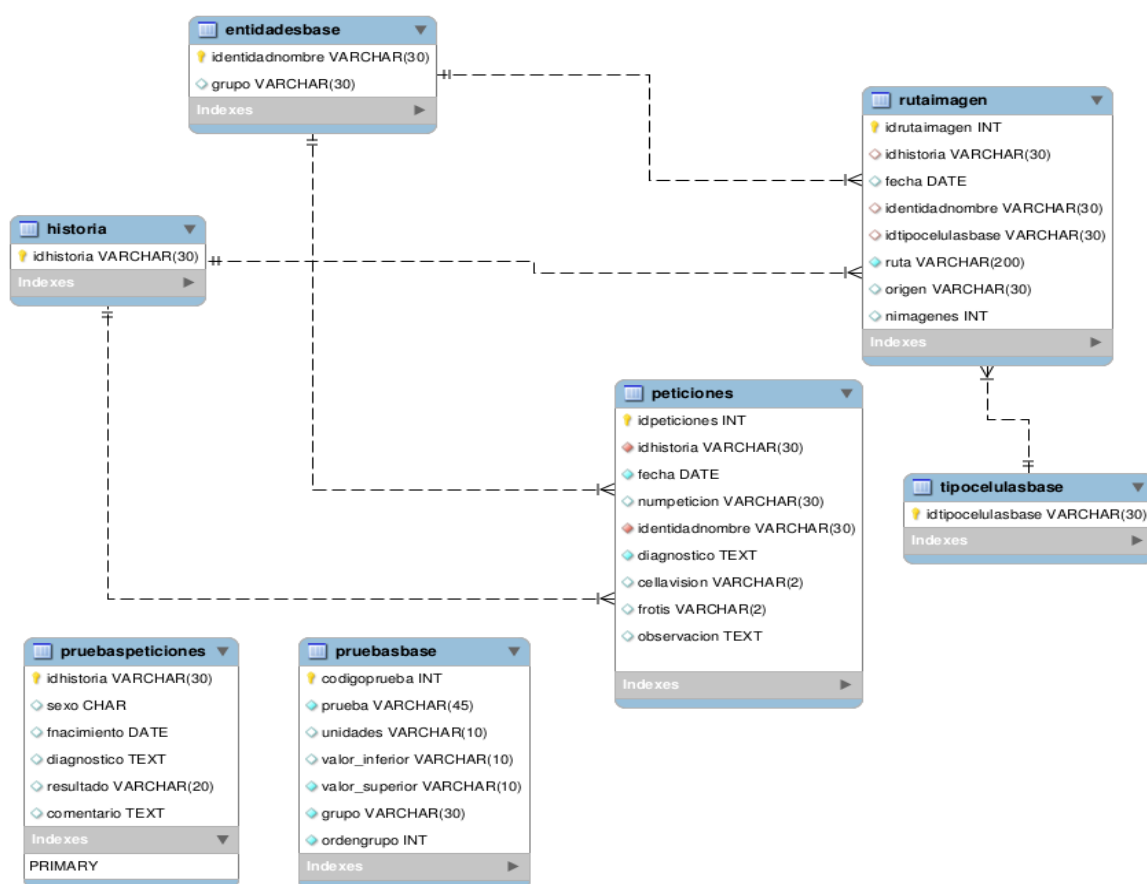


Figura 3.7: Diagrama de les diferents taules que conformen la base de dades.

La base de dades està formada per 7 taules: ‘entidadesbase’, ‘tipocelulasbase’, ‘història’, ‘peticions’, ‘rutaimagen’, ‘pruebaspeticiones’ i ‘pruebasbase’. Les taules ‘entidadesbase’ i ‘tipocelulasbase’ contenen informació sobre les entitats (LLC, LF, LCM, LZME, HCL, LPB, MM amb CP, CLR, Normal, Blasts limfoides, SS, LPT i LLGG-T) i tipus de cèl·lules que hi ha (limfòcits anormals, prolimfòcits, centrocits, limfòcits reactius, limfòcits normals, centroblasts, blasts, etc). Ambdues taules serveixen de referència per a les dades que s’emmagatzemen a les altres taules.

La taula ‘història’ conté tots els números de les històries clíniques codificades. És l’única dada d’aquesta taula i és la seva clau primària, és a dir, cada registre és únic i serveix de referència per a les altres taules. Com una història clínica pot tenir varies peticions en diferents dates, s’ha creat la taula ‘peticions’, la qual permet identificar de manera inequívoca un registre de peticions combinant el número de petició (‘numpeticion’) i la data. La clau primària d’aquesta taula és el camp ‘idpeticion’ i és de tipus enter incremental. D’altra banda, aquesta taula conté informació relativa a l’entitat diagnosticada, l’origen de les imatges (Microscòpia o CellaVision®DM96) i els marcadors de la citometria de flux, entre d’altres.

La taula ‘rutaimagen’ relaciona la petició i la data amb una ruta al directori on es troben les imatges de cada tipus de cèl·lula per a cada petició. Cal mencionar que els arxius de les imatges no s’emmagatzemen a la BD, sinó que es troben en una ruta específica en un servidor i és la ruta la que s’emmagatzema a la BD. S’ha decidit fer-ho d’aquesta manera perquè és més eficient perquè cada imatge adquirida pel microscopi té una mida de l’ordre de 25 MB. Aquesta taula també conté informació de l’origen de les imatges (Microscopi o CellaVision®DM96), de la quantitat d’imatges per ruta i l’entitat diagnosticada.

Per últim, les taules ‘pruebasbase’ i ‘pruebaspeticiones’ contenen informació sobre l’anàlisi del laboratori de cada petició. ‘Pruebasbase’ conté informació dels valors de referència de tots els paràmetres de l’anàlisi, les unitats i l’ordre en que s’han de presentar quan es consulti la informació. La taula ‘pruebaspeticiones’ conté l’anàlisi per a cada petició i dades demogràfiques com l’edat i el sexe, i també conté l’entitat diagnosticada. Aquesta taula pot relacionar-se amb la taula ‘peticiones’ a través del número de la petició i la data.

3.3 Algoritme de segmentació

La segmentació és la segona etapa del processament digital d'imatges i permet obtenir, en el cas de les cèl·lules, les tres regions d'interès: nucli, citoplasma i regió externa. L'algoritme de segmentació consta de tres etapes: el preprocessament, la segmentació de la cèl·lula i el postprocessament.

3.3.1 Preprocessament

El preprocessament pretén millorar la qualitat de la imatge disminuint els nivells de soroll i ressaltar els marges de les cèl·lules suavitzant la seva textura. Hi ha diferents factors que poden afectar la qualitat de l'adquisició d'imatges derivades del frotis: procés de tinció, variacions en la il·luminació, distorsions geomètriques òptiques a causa del tipus de lent i microscopi, format en què s'emmagatzemen les imatges (no es recomana formats comprimits ja que es perden dades), etc. D'aquesta manera podem millorar els resultats de la segmentació i així facilitar l'extracció de la informació i/o característiques de les cèl·lules.

S'han descrit algunes tècniques de preprocessament per millorar la qualitat de la imatge abans del procés de segmentació, tant per seleccionar l'àrea absoluta del frotis de sang (61) com per descriure les característiques individuals de cada cèl·lula: millora del contrast, transformacions de color, altres manipulacions de l'histograma o filtrat (62,63).

El filtrat de les imatges té com a objectiu reduir el nivell de soroll utilitzant diversos tipus de filtres. Eom *et al.* (64) utilitza un filtre gaussià de 3 x 3 per preparar imatges de leucòcits per al procés de segmentació. Scotti *et al.* (61) aplica un filtre gaussià *low-pass* per a reduir el soroll. Sabino *et al.* (65) utilitza també un filtre gaussià per al gradient de la imatge abans de realitzar una transformació per a la detecció automàtica del citoplasma. Angulo *et al.* (66) filtra el component de color verd de la imatge per eliminar el soroll i petites errades de digitalització i segmenta principalment els eritròcits del frotis. Ghosh *et al.* (67) utilitza un filtre Wiener per reduir el soroll de fons de les imatges dels leucòcits i millorar la metodologia de segmentació.

Com s'ha explicat al Capítol 2, les imatges tenyides amb May Grünwald-Giemsa són de color blau i porpra. Per tant, una transformació de color o un tractament de les imatges dels leucòcits que separi, ressalti o extregui els millors perfils es poden aplicar tant en el preprocessament com a la segmentació. Sinha i Ramakrishnan (68) apliquen la transformació de color des de l'espai de color RGB (vermell, verd i blau) a l'HSV (tonalitat, saturació i valor) abans del procés de segmentació. Würflinger *et al.* (62) implementa una transformació de l'espai de color mitjançant una anàlisi d'agrupament de l'histograma RGB i aplica la transformació de Fisher reduint la diferència entre tres procediments diferents de tinció.

A la present *Tesi Doctoral* s'ha utilitzat la metodologia i els algoritmes corresponents pel preprocessament i la segmentació desenvolupats per Alférez *et al.* (59) tant per a les imatges del sistema DM96 com del BX43.

La Figura 3.8 mostra un diagrama de blocs, on la primera part inclou les diferents tècniques de preprocessament:

- Transformacions de colors: RGB a XYZ i RGB a YCbCr
- Millora del contrast: *contrast stretching*
- Filtrat: *median filtering*, *average filtering*

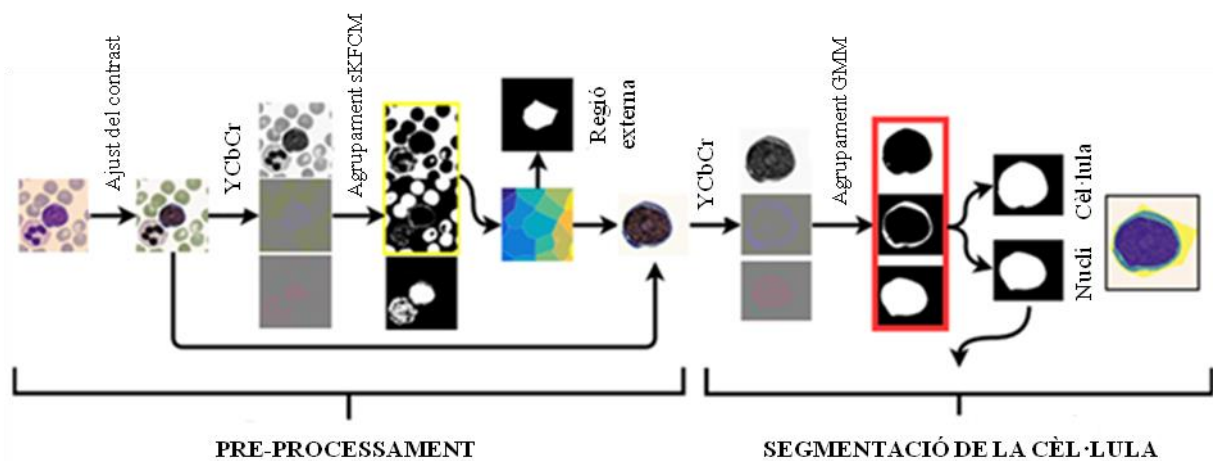


Figura 3.8: Diagrama que resumeix les etapes de l'algorisme de segmentació (Font: S.Alfárez et al. (59))

El primer pas és l'obtenció del fons i els hematies de la imatge. Per això, en primer lloc, es realitza un filtre de la imatge original i es defineix l'espai de color amb el qual es treballarà, sent aquest el YCrCb. Una vegada definits els primers paràmetres, s'inicia l'algorisme d'agrupament *spatial kernel fuzzy c-means* (*sKFCM*).

La tècnica *fuzzy c-means* (FCM) és un mètode no supervisat que divideix les dades en clústers, minimitzant la distància entre cada punt de dades i el seu respectiu centre de clústers. Tanmateix, sota certes restriccions, cada punt de dades pot pertànyer a diversos grups alhora de manera difusa (69). El FCM ha estat àmpliament utilitzat en aplicacions mèdiques, com ara l'anàlisi de les imatges de ressonància magnètica (70), microarrays (71), identificació de cèl·lules (72), entre d'altres.

L'algorisme emprat, el *sKFCM*, per a l'obtenció dels diferents grups de la imatge, manté el mateix principi que s'ha vist en FCM però modificant la funció objectiu mitjançant la introducció de funcions kernel i utilitzant informació espacial per modificar la partició matriu (3). Mitjançant aquest algorisme d'agrupació, a partir de la imatge en l'espai YCrCb, s'aconsegueixen separar els grups: la cèl·lula, els hematies i el fons.

Una vegada s'han separat els tres grups amb l'algorisme *sKFCM*, es procedeix primer a l'ús del mètode d'Otsu per a binaritzar el grup d'hematies obtinguts. Una binarització és una modificació en dos nivells i s'utilitza per separar els objectes del fons. El mètode d'Otsu (73) utilitza la variància, sent aquesta la mesura

de dispersió de valors (dispersió de nivells de gris). El seu objectiu és calcular el valor llindar de manera que la dispersió dins de cada classe sigui la mínima possible, però a la vegada la dispersió entre classes diferents és la màxima possible.

A continuació, es procedeix a la utilització d'una sèrie d'operacions morfològiques juntament amb la transformada *Watershed* (74). Les operacions morfològiques estan basades en la teoria de conjunts i es fan servir per simplificar imatges mantenint les principals característiques dels objectes. El valor de cada píxel de la imatge resultant depèn d'aquest píxel a la imatge original i la seva relació amb els veïns. Això permet identificar i reconstruir formes distorsionades o amb soroll.

Per últim, amb l'objectiu de poder determinar quantes cèl·lules es troben a la imatge, s'obté una aproximació dels nuclis. Sobre la imatge s'apliquen una sèrie d'operadors morfològics que tenen com a objectiu eliminar artefactes causats per la tinció. Una vegada han estat eliminats, es pot comptar a partir d'ells la quantitat de cèl·lules. Després s'associa el centroid de cada un dels nuclis amb la seva respectiva regió de la imatge i es procedeix a retallar la imatge a partir de la regió externa. Finalment, es calcula la mitjana del fons i es substitueixen els hematies pel valor obtingut. Així, l'única cèl·lula que queda a la imatge és la cèl·lula limfoide, tal i com es veu al final de l'etapa de preprocessament a la Figura 3.8.

3.3.2 Segmentació

L'etapa de segmentació separa els diferents objectes d'una imatge. Aquest procediment és essencial en el processament digital d'imatges perquè una segmentació robusta és primordial per tenir èxit en les etapes restants d'extracció de descriptors i classificació (75). La segmentació de les cèl·lules de sang perifèrica és un procediment complex a causa de la gran variabilitat cel·lular (tant de la morfologia com de la mida) i els problemes causats per les variacions de la metodologia de tinció. Hi ha molts estudis que han abordat aquest problema i existeixen un gran nombre de mètodes per a la segmentació d'imatges digitals (76).

A la literatura s'han descrit diverses metodologies sobre el procés de segmentació cel·lular, ja sigui per a les diferents subpoblacions de leucòcits, selecció de l'àrea de treball del frotis, descripció morfològica dels hematies o reconeixement de cèl·lules limfoïdes neoplàsiques.

La tècnica *thresholding* és la segmentació més simple, en la qual es tria un llindar de l'histograma (funció aproximada de distribució de probabilitat) i se separen dos objectes. Per la seva simplicitat i eficiència, aquesta tècnica s'ha utilitzat amb múltiples llindars (77), aplicant el mètode d'Otsu per a segmentar leucòcits (77), i una combinació de contrast automàtic de *stretching*, operacions aritmètiques d'imatges, filtrat mínim i tècniques de llindar globals per a localitzar el nucli dels leucòcits (78).

A més, Piuri i Scotti (79,80) utilitzen la detecció de vores Canny i morfologia matemàtica binària per segmentar la cèl·lula sencera i obtenir les regions d'interès així com l'ajust del contrast i el *thresholding* automàtic per separar el nucli dels leucòcits. Ghosh *et al.* (67) presenta una segmentació del nucli dels

leucòcits utilitzant divergència difusa a partir de funcions de pertinença de tipus Gamma, Gaussian i Cauchy dels píxels de la imatge i alguns mètodes *thresholding* modificats.

Una forma bàsica de tractar la segmentació del color és separar els components de l'espai de color RGB de la imatge. S'ha demostrat que un llindar de 100 (entre 0 i 255) aplicat sobre el component verd condueix a una bona separació del nucli cel·lular (81). A més, varis treballs utilitzen diverses transformacions a altres espais de color dins de l'aplicació de diferents mètodes de segmentació:

- *Scale space filtering* i transformada *Watershed* sobre un histograma 3D de l'espai de color de HSV per separar el citoplasma (82)
- Classificació de Bayes sobre CIELAB (65)
- *Thresholding* del component S per obtenir la regió del nucli de les cèl·lules blàstiques (83)
- Extracció i binarització dels components H i S de l'espai de color HSV aplicació d'operacions de morfologia matemàtica per tal de localitzar i segmentar limfoblasts (78)

La segmentació per clústers és un mètode no supervisat per a agrupar píxels d'acord amb algun criteri, generalment el color. En aquest cas, cada píxel és un vector de tres components (per exemple, a l'espai de color RGB: R, G i B) i el problema es resol mitjançant l'agrupació dels vectors similars segons un criteri. Foran *et al.* (84) segmenta les cèl·lules limfoides utilitzant la transformació de l'espai de color Luv de la imatge de la cèl·lula per implementar el mètode ràpid no paramètric de clusterització desenvolupat per Comaniciu i Meer (85), el qual és un algoritme estable i eficient.

Sinha i Ramakrishnan (68) segmenten el nucli i el citoplasma dels leucòcits a través de l'algoritme de maximització amb una estimació inicial que utilitza *k-means clustering* a l'espai de color HSV de la imatge de la cèl·lula. Ramoser *et al.* (86) aplica la tècnica de *k-means clustering* a l'espai HSV per a segmentar el nucli i segmentació probabilística per separar el leucòcit. González *et al.* (26) utilitza la tècnica de *k-means clustering* amb la distància euclidiana estàndard a l'espai de color CIELAB, però omet el component de lluminositat per separar la imatge en quatre grups: fons, citoplasma, nucli dels leucòcits i nucli dels hematies d'imatges de medul·la òssia. Angulo *et al.* (9,10) desenvolupa una metodologia per extreure els eritròcits i els leucòcits mitjançant un *thresholding* automàtic i un filtrat binari del component verd, i posteriorment la transformada *Watershed* s'aplica sobre un gradient morfològic dels components verds i saturació (RGB i HSV) per a segmentar el nucli i el citoplasma dels limfòcits.

En relació a la segmentació utilitzada en la present *Tesi Doctoral*, com s'ha introduït a l'apartat anterior, s'ha emprat l'algoritme de segmentació desenvolupat per Alférez *et al.* (59). A la segona part de la Figura 3.8 se'n pot veure l'esquema. A mode de resum, amb l'objectiu de realitzar la segmentació final, es fa servir la imatge retallada i preprocessada de la cèl·lula i es realitza un canvi d'espai de color de RGB a YCbCr. En primer lloc, es realitza una agrupació en tres conjunts utilitzant un model mixt Gaussià (GMM) als píxels de la imatge en YCbCr. El GMM és un model probabilístic d'agrupació que calcula la probabilitat de

pertinença de cada píxel a cada un dels grups (87). La finalització de la classificació es fa mitjançant l'assignació dels elements al grup més probable. A partir del grup nucli obtingut per GMM s'apliquen operadors morfològics per segmentar el nucli definitivament. Per a l'obtenció de la cèl·lula es sumen els dos grups de nucli i citoplasma (Figura 3.8), i s'apliquen també operadors morfològics per eliminar possibles àrees internes causades per soroll.

Com a exemple, a la Figura 3.9 es pot veure la segmentació d'una cèl·lula limfoide anormal. Tal i com s'aprecia a la figura, l'algoritme de segmentació ha estat capaç de separar de forma correcta les diferents parts de la cèl·lula i excloure els hematies de la zona externa de la cèl·lula.

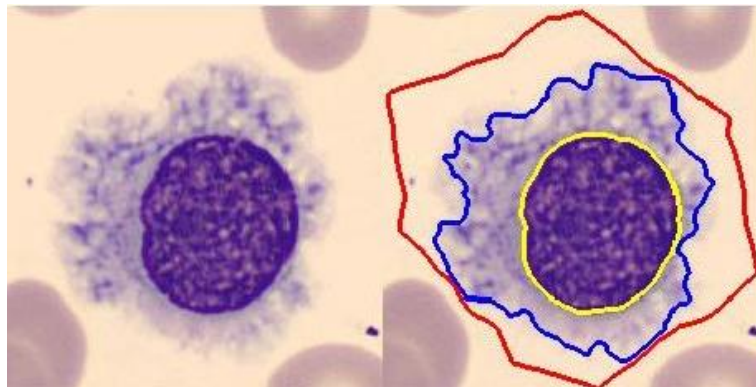


Figura 3.9: Segmentació d'una cèl·lula limfoide anormal (tricoleucòcit).

3.3.3 Postprocessament

L'etapa de postprocessament està destinada a la separació de les cèl·lules adjacents. Per decidir si n'hi ha, s'utilitza la variància circular de la regió d'interès obtinguda per a la cèl·lula i per al nucli. En cas afirmatiu, es processen de nou les tres regions d'interès de les cèl·lules mitjançant la utilització de la transformada de distància, la transformada H mínima i la Watershed per separar-les totes.

3.4 Extracció de descriptors

L'extracció de descriptors és l'element fonamental per al reconeixement automàtic i la indexació del contingut de la imatge digital. Una vegada se segmenten els objectes d'una imatge, és possible realitzar l'extracció de descriptors quantitius que permetin obtenir mesures per a la caracterització de la cèl·lula. Tal i com s'ha comentat anteriorment, la descripció qualitativa pot ésser subjecte a certa subjectivitat i requereix experiència.

Per tant, l'objectiu d'aquesta etapa és definir unes variables que tinguin una formulació matemàtica sòlida a partir de les imatges digitals processades prèviament, complint dues premisses:

- Ser discriminatòries entre els diferents tipus de cèl·lules sanguínies
- Ser interpretables en termes de morfologia cel·lular

Si la imatge és molt simple, l'extracció de descriptors es pot fer sense la necessitat de filtrar i /o segmentar, per exemple, si la imatge només conté un tipus de textura. Fins i tot si la imatge és relativament complexa, també es poden extreure descriptors útils sense segmentació. No obstant això, s'ha de minimitzar al màxim possible l'entropia de la informació d'entrada a l'algoritme de descripció.

Per exemple, en el cas dels leucòcits, la imatge d'entrada prové d'un algoritme de localització i segmentació de la imatge que ens assegura que només hi ha un leucòcit a la imatge. A partir d'ara, entendrem per extracció de descriptors com qualsevol algoritme que busca obtenir paràmetres quantitius dels objectes de la imatge.

Les variables (descriptors cel·lulars morfològics) que ens permetran obtenir mesures per a la caracterització de la cèl·lula, es basaran en variables morfològiques per a la seva descripció, com la mida, la relació nucli/citoplasma, el contorn nuclear, l'excentricitat nuclear, característiques de la cromatina, presència de nuclèols, basofília del citoplasma i contorn d'aquest, presència de granulació o vacuoles, textura i altres d'interès citològic.

Hi ha un elevat nombre de descriptors que es poden extreure d'un objecte d'una imatge i aquests es classifiquen en tres categories (76):

1. *Descriptors geomètrics*: àrea, relació nucli/citoplasma, circularitat, entre d'altres.
2. *Descriptors de color o basats en l'histograma*: altrament coneguts com estadístics de primer ordre i inclou a sis paràmetres: mitjana, desviació estàndard, asimetria, curtosi, energia i entropia.
3. *Descriptors de textura o basats en la intensitat*: n'hi ha de diferents tipus i destaquem
 - Estadístics de segon ordre (basats en la matriu de co-ocurrència)
 - Granulomètrics
 - *Wavelet*
 - Gabor

Donat que l'objectiu de la present *Tesi Doctoral* és identificar un conjunt de descriptors per a la caracterització de la forma, mida, color, textura i altres particularitats de les cèl·lules limfoides normals, reactives, anormals presents a les neoplàsies limfoides i blàstiques, aquest apartat conté una descripció detallada de cada tipus de descriptor. S'han extret tant descriptors geomètrics com de color i textura de les diferents regions d'interès i alguns d'ells representen característiques qualitatives morfològiques habitualment emprades pel citòleg (34) o paràmetres quantitius abstractes (75).

3.4.1 Descriptors geomètrics

Els descriptors geomètrics són un grup de valors numèrics obtinguts per descriure propietats geomètriques d'una forma concreta. Mesuren característiques morfològiques bàsiques com la mida, la forma, el perímetre, l'àrea de la cèl·lula, del nucli o del citoplasma. Els descriptors geomètrics tracten de quantificar les interpretacions geomètriques de les cèl·lules que es correlacionen amb la percepció i la interpretació que fa el citòleg expert.

A més, també inclouen característiques relacionades amb l'orientació relativa i sobre la forma. La representació dels descriptors geomètrics i les seves tècniques de descripció es divideixen en dues categories principals:

- Les representacions basades en l'aproximació *Boundary*, les quals extreuen descriptors només de les zones perifèriques, és a dir, del contorn de les regions d'interès seleccionades, ignorant per complet el que es troba al seu interior. Un exemple d'aquest tipus de descriptors són el perímetre, la circularitat i l'excentricitat.
- Les representacions basades en l'aproximació de *Region*, les quals tenen en compte tant els detalls interiors com els detalls que delimiten la regió d'interès. Alguns exemples són l'àrea o la longitud de l'eix major i menor.

Per a un major coneixement dels descriptors geomètrics que es poden extreure de cada cèl·lula individual, es mostra a continuació la seva classificació segons la regió d'interès de la qual han estat calculats:

- Nucli: àrea, longitud de l'eix major, longitud de l'eix menor, excentricitat, perímetre, circularitat, diàmetre.
- Citoplasma: *hairiness*, relació nucli/citoplasma, àrea, perímetre.
- Cèl·lula: àrea, relació nucli/citoplasma, circularitat.

Aquest tipus de descriptors són senzills però tot i així són considerats com a característiques essencials i fonamentals per al reconeixement d'objectes dins d'una imatge. És important destacar que, tot i que aquests descriptors són rarament decisius per a propòsits discriminatoris, són molt útils per a diferenciar entre diferents tipus de cèl·lules.

A continuació, explicarem alguns dels descriptors geomètrics que es poden utilitzar a la caracterització de cèl·lules en sang perifèrica:

3.4.1.1 Àrea

Considerem àrea a aquella superfície delimitada o marcada per uns límits. A matemàtiques, una àrea és l'extensió que pot presentar una figura geomètrica, la qual pot ser mesurada quantitativament. A la imatge d'una cèl·lula sanguínia, l'àrea és una mesura calculada a partir del nombre de píxels de la regió d'interès concreta que es vulgui analitzar. Aquest descriptor és usat per quantificar la mida de la regió d'interès, ja sigui del nucli, de la cèl·lula o del citoplasma. En alguns casos, pot ser d'interès convertir el valor numèric expressat com nombre de píxels quadrats a una unitat de mesura determinada (mm^2 , μm^2 , etc).

3.4.1.2 Rati nucli/citoplasma

El rati és definit com el quocient entre nombres o quantitats comparables, és a dir, és el quocient entre dues magnituds relacionades entre si. En el cas del rati nucli/citoplasma, aquest es calcula amb la següent fórmula:

$$\text{Rati nucli/citoplasma} = \frac{\text{Àrea nucli (píxels)}}{\text{Àrea cèl·lula} - \text{àrea nucli (píxels)}} \quad (3.2)$$

És d'esperar que aquelles cèl·lules on el nucli ocupi la major part del citoplasma, presentin valors alts per a aquest descriptor geomètric, mentre que altres tipus cel·lulars amb nuclis de menor mida i més extensió de citoplasma mostrin valors baixos.

3.4.1.3 Rati nucli/cèl·lula

El descriptor geomètric del rati nucli/cèl·lula és calculat com la relació entre la mida de l'àrea del nucli i la mida de l'àrea de tota la cèl·lula a estudi. La fórmula matemàtica utilitzada per calcular aquest descriptor es mostra a continuació:

$$\text{Rati nucli/cèl·lula} = \frac{\text{Àrea nucli (píxels)}}{\text{Àrea cèl·lula (píxels)}} \quad (3.3)$$

3.4.1.4 Perímetre

El perímetre és la distància al voltant d'una figura de dues dimensions, és a dir, la suma de les longituds dels costats d'una figura geomètrica plana. És una mesura fonamental en la determinació de la mida d'un objecte. El valor numèric per a aquest descriptor ve donat com nombre de píxels i, per tant, també pot ser transformat a escala mètrica.

3.4.1.5 Hairiness

Aquest descriptor geomètric va ser proposat per Alférez *et al.* (29) i és útil per a caracteritzar el perfil citoplasmàtic. La Figura 3.10 (adaptada de la Tesi Doctoral de S. Alférez) mostra un exemple de com es calcula el descriptor *hairiness* en la imatge d'un tricoleucòcit.

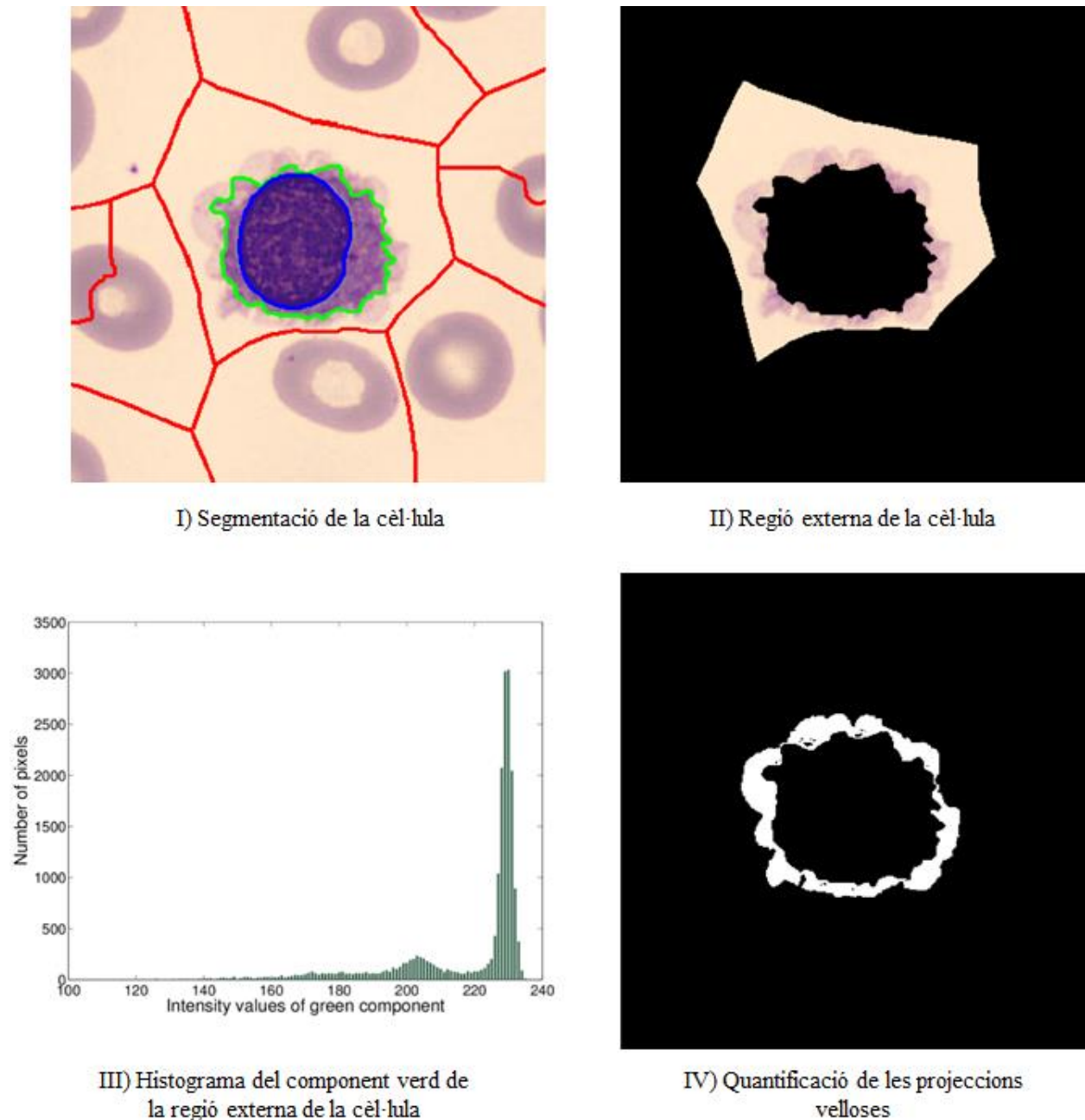


Figura 3.10: Etapes per calcular el descriptor citoplasmàtic *hairiness* a la imatge d'un tricoleucòcit. Després de la segmentació de la cèl·lula (I), es selecciona la regió externa de la cèl·lula (II). L'histograma del component verd d'aquesta regió mostra una distribució bimodal (III), sent la primera (70-215 d'intensitat) qui conté la major part de les projeccions citoplasmàtiques. Per últim, s'estableix la presència d'aquestes projeccions (IV) i es quantifica l'àrea. Font: Tesi Doctoral S. Alférez (3).

En primer lloc es realitza la segmentació en tres regions d'interès (part I): la cèl·lula completa (contorn verd), el nucli (contorn blau) i la regió externa delimitada pel contorn extern (vermell). A continuació treballarem amb la màscara de la regió externa (part II) i l'histograma del component verd d'aquesta regió externa (part III), on a l'eix d'ordenades es representa el nombre de píxels i a l'eix d'abscisses els valors de les intensitats del component verd. L'histograma d'aquesta regió mostra una distribució bimodal per tots els tipus de cèl·lules estudiades. L'experiència ens ha indicat que un interval d'intensitat de verd entre 70 i 215 conté la major part de les projeccions citoplasmàtiques. Valors inferiors a 70 podrien correspondre a la presència de plaquetes i valors superiors a 215 corresponen al fons de la regió externa. Assignant el valor 1 a tots els píxels d'aquest interval i 0 a la resta, s'obté la màscara final (part IV), on la regió blanca correspon a les vellositats citoplasmàtiques. La suma de tots els píxels quantifica l'àrea i defineix el descriptor *hairiness*.

El descriptor *hairiness* es pot expressar també en escala mètrica perquè es tracta d'una àrea. Partint del coneixement que 1 píxel d'una imatge del sistema DM96 correspon a 99,3227 nanòmetres i que un píxel és un quadrat, per convertir els píxels a escala mètrica, cal elevar al quadrat aquest factor. Per tant, per convertir els píxels a micròmetres al quadrat es pot seguir la següent formulació:

$$1 \text{ píxel} = 99,3227 \text{ nm} \times 99,3227 \text{ nm} = 9865 \text{ nm}^2 \times \frac{1 \mu\text{m}^2}{10^6 \text{nm}^2} = 0,009865 \mu\text{m}^2 \quad (3.4)$$

Durant el transcurs de la Tesi Doctoral, es fan servir ambdues unitats a conveniència.

3.4.1.6 *Diàmetre equivalent*

Es tracta d'una paràmetre que especifica el diàmetre d'un cercle amb la mateixa àrea que la regió d'interès.

3.4.1.7 *Circularitat*

La circularitat és una condició d'una superfície d'un objecte, que s'utilitza per controlar la rodonesa de les parts o característiques circulars. Aquestes característiques circulars poden ser definides per cilindres, esferes i cons. Aquest descriptor descriu com de similar és una regió d'interès a un cercle, reflectint així la complexitat del contorn de l'objecte. La fórmula matemàtica utilitzada per calcular aquest descriptor és la següent:

$$\text{Circularitat} = \frac{(\text{Perímetre regió d'interès})^2}{\text{Àrea regió d'interès}} \quad (3.5)$$

3.4.1.8 *Àrea convexa*

Especifica el nombre de píxels de la imatge convexa. S'entén com a polígon convex a aquella figura plana composta per segments rectes no alineats, les diagonals són sempre interiors i els angles interns no superen els 180°. Aquest descriptor és útil quan es mesura l'àrea de la regió d'interès sense considerar les irregularitats del contorn d'aquesta regió.

3.4.1.9 Excentricitat cònica

L'excentricitat és un paràmetre que determina el grau de desviació d'una secció cònica respecte a una circumferència. És definida com la relació de la distància de qualsevol punt de la corba d'una el·lipse i un dels seus focus. El cercle té una excentricitat de zero, una el·lipse té una excentricitat menor a un, una paràbola té una excentricitat d'un i una hipèrbola té una excentricitat major a un. El descriptor geomètric de l'excentricitat cònica és molt útil per calcular l'excentricitat nuclear d'una cèl·lula, és a dir, la distància entre el centre de la cèl·lula i el centre del nucli.

3.4.1.10 Elongació

L'elongació o allargament és el resultat d'augmentar la longitud d'un objecte. Descriu el grau d'allargament d'un objecte, de manera que el seu valor és zero per a un cercle i arriba fins a un per una línia. A mesura que aquest valor augmenta allunyant-se de zero, indica que l'objecte té una forma més allargada. L'elongació pot ser calculada a partir de la següent fórmula matemàtica:

$$\text{Elongació} = 1 - \frac{\text{Longitud eix menor el·lipse (píxels)}}{\text{Longitud eix major el·lipse (píxels)}} \quad (3.6)$$

3.4.1.11 Solidesa

Entenem per convexitat d'una corba o una superfície com la zona que s'assembla a l'exterior d'una circumferència o una superfície esfèrica, és a dir, que té la seva part externa dirigida a l'observador. Concavitat, en canvi, és la part que s'assembla a la zona interior d'una circumferència o una esfera, és a dir, que té la seva part enfonsada dirigida a l'observador. Tots dos conceptes són complementaris entre ells.

El descriptor de la solidesa també pot ser calculat com el quocient entre l'àrea i l'àrea convexa de la regió d'interès, on la solidesa d'una regió d'interès convexa sempre és 1. La fórmula matemàtica que dóna lloc a aquesta expressió és la següent:

$$\text{Solidesa} = \frac{\text{Àrea regió d'interès (píxels)}}{\text{Àrea convexa regió d'interès (píxels)}} \quad (3.7)$$

3.4.1.12 Bounding box

El descriptor *bounding box*, també conegut com a *minimum bounding rectangle*, correspon al rectangle més petit que conté la regió d'interès.

3.4.1.13 Extensió

Aquest descriptor especifica la proporció de píxels de la regió en relació als píxels de la *bounding box*.

$$\text{Extensió} = \frac{\text{Àrea regió d'interès (píxels)}}{\text{Àrea bounding box regió d'interès (píxels)}} \quad (3.8)$$

3.4.1.14 Convex hull

El *convex hull* o càpsula convexa d'un conjunt de punts X de dimensió n es defineix com la intersecció de tots els conjunts convexes que contenen a X . Donats k punts, la seva càpsula convexa C ve donada per la següent expressió:

$$\text{Convex hull} = \left\{ \sum_{i=1}^k \alpha_i x_i \mid x_i \in X, \alpha_i \in \mathbb{R}, \alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1 \right\} \quad (3.9)$$

3.4.2 Descriptors de color

Quan a la secció 3.2.2 s'ha introduït l'estructura bàsica d'una imatge en escala de grisos, no s'ha explicat que el color juga un paper fonamental en la descripció morfològica de les cèl·lules sanguínies, concretament, els descriptors de color.

El microscopi proporciona imatges amb color que són producte de tincions especialment creades per a ressaltar-ho. El sistema visual humà és capaç de discernir entre molts matisos i intensitats de color, mentre que és poc hàbil per reconèixer canvis subtils en escenes que contenen grisos. D'altra banda, en el processament digital d'imatges l'anàlisi del color genera descriptors molt útils per al reconeixement d'objectes. En aquest apartat es presenten conceptes bàsics sobre com representar el color i sobre l'estructura d'una imatge a color.

Isaac Newton (1642 – 1727) va ser un físic, matemàtic, filòsof anglès i autor dels *Philosophiae Naturalis Principia Mathematica* (1687), on va descriure la llei de la gravitació universal i les tres lleis del moviment (lleis de la inèrcia), base de la mecànica clàssica. En òptica, va construir el primer telescopi reflector pràctic i va desenvolupar a l'any 1666 una teoria sobre el color basada en l'observació que un prisma descompon un raig de llum blanca en els colors de l'espectre visible (colors de l'arc de Sant Martí). Posteriorment, nous desenvolupaments físics van emmarcar aquest descobriment com la part visible a l'ull humà que s'anomena espectre electromagnètic, com il·lustra la Figura 3.11.

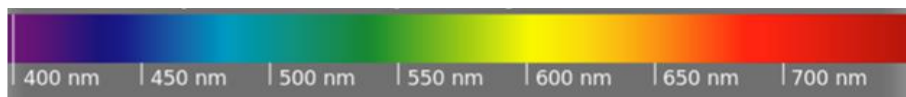


Figura 3.11: Part visible de l'espectre electromagnètic.

L'espectre visible es divideix en sis regions de color: violeta, blau, verd, groc, taronja i vermell, corresponents a longituds d'ona entre 400 i 700 nanòmetres (nm). Els colors que percep l'humà en un objecte venen determinats per la naturalesa de la llum reflectida per l'objecte. Un cos que es veu blanc reflecteix la

llum de forma equilibrada en totes les zones de l'espectre visible. Un objecte que es veu vermell reflecteix la llum principalment amb longituds d'ona per sobre dels 700 nm i absorbeix la major part de l'energia per a les altres. L'espectre visible és continu, de manera que no hi ha una transició clara entre colors. Per exemple, un objecte que reflecteix llum en un interval aproximat entre 600 i 650 nm es veu d'un color groc però és difícil matisar-ho.

Hi ha tres colors que es denominen primaris: el blau, el verd i el vermell. Amb efecte d'estandardització, la Comissió Internacional de la Il·luminació (CIE) al 1931 va fixar les següents longituds d'ona per als colors primaris: 435,8 nm (blau), 546,1 nm (verd) i 700 nm (vermell). Sovint es diu que una combinació apropiada d'aquests tres colors pot produir qualsevol color de l'espectre visible i una barreja dels tres produeix color blanc. Combinant dos colors primaris qualsevol s'obtenen els anomenats colors secundaris: magenta (blau i vermell), cian (blau i verd) i groc (verd i vermell).

S'anomena color complementari o oposat a aquell color primari que no intervé en la definició d'un color secundari. Així, el verd és complementari del magenta, el color vermell ho és del cian i el blau és complementari del groc. Un color secundari barrejat amb el seu complementari produeix també color blanc. En aquest sentit, els colors secundaris actuen com a filtres dels primaris corresponents. Així, un cos de color magenta absorbeix, i per tant no transmet, el color verd. Un cos de color cian filtra el vermell i un cos groc filtra el blau. D'aquesta manera, una combinació dels colors secundaris en les proporcions adequades produeix color negre com a resultat d'una absorció total de la llum blanca. D'acord amb aquesta propietat, es diu que els colors secundaris són sostractius. Per contra, els colors primaris són additius.

Els colors secundaris són particularment importants per als pigments i les tincions. Es tracta de materials amb la propietat d'absorbir de forma selectiva una certa longitud d'ona (color) i transmetre les altres. És la base de la creació de pintures, colorants i tintes per impressió, afegint pigments fins a aconseguir la coloració desitjada. Els colors secundaris són particularment importants en aquesta Tesi, ja que són la base de les tincions utilitzades en la microscòpia òptica per millorar el contrast en les imatges i ressaltar parts de les cèl·lules. Lògicament la tinció és rellevant tant per a l'observació visual de la morfologia al microscopi com per al processament i el reconeixement automàtic.

D'altra banda, és important conèixer les característiques que permeten diferenciar entre colors:

- **Brillantor**: és la intensitat de la llum transmesa per l'objecte. En el cas de les imatges en escala de grisos vistes anteriorment, és l'única característica utilitzada. De forma anàloga a aquest cas, en presència de color, una major intensitat correspon a color més clar (més proper al blanc), mentre que menys intensitat indica color més fosc (proximitat al negre).
- **Matís (*hue*)**: representa el color dominant en si mateix, tal com és percebut i descrit per un observador. És el que fa que una persona, de manera intuïtiva, digui verd, taronja o qualsevol altre color.

- **Saturació:** indica la puresa relativa d'un color, o la quantitat de blanc que es barreja amb un color pur. Els colors purs de l'espectre estan saturats completament. El nivell de saturació és inversament proporcional a la quantitat de blanc afegit. Un color saturat dóna sensació de viva, mentre que la manca de saturació pot donar sensació de color apagat o brut.

El matís i la saturació defineixen l'anomenada cromaticitat i són característiques que tenen un fons més qualitatiu. La brillantor és una mesura clarament quantitativa, per això sol anomenar-se també com el valor del color.

3.4.2.1 *Espais de color*

Un model o espai de color és un mètode per especificar els colors d'una manera estandarditzada i acceptada de forma general. Essencialment consisteix en un sistema de coordenades en tres dimensions (75). Hi ha una àmplia varietat de models avui en dia, ja que la tecnologia del color està present en moltes àrees d'aplicació.

En la present *Tesi Doctoral* s'ha utilitzat la descomposició en sis espais de color per a l'extracció dels descriptors i obtenir informació rellevant. A la Figura 3.12 es mostren els diferents espais de color utilitzats i la seva representació en sistemes de coordenades.

CIE RGB

L'espai de color CIE 1931 RGB i l'espai de color CIE 1931 XYZ van ser creats per la CIE al 1931. Va ser el resultat d'una sèrie d'experiments realitzats a finals de la dècada de 1920 per William David Wright i John Guild. Els resultats experimentals es van combinar en l'especificació de l'espai de color CIE RGB, del qual es deriva l'espai de color CIE XYZ.

El model RGB és molt utilitzat per a la representació, la transmissió i l'emmagatzematge d'imatges en televisió, ordinadors i càmeres digitals, com succeeix en els analitzadors d'imatges i els microscopis dels laboratoris. L'espai RGB (Red, Green, Blue) descriu un color general com una combinació dels tres colors primaris (vermell, verd i blau). Per a això s'estableix un sistema de coordenades cartesianes on cada eix correspon a un dels tres colors primaris i l'origen es troba en el color negre (K).

Qualsevol color possible, C_i , es representa com un punt amb coordenades: $C_i = (R_i, G_i, B_i)$, on R_i , G_i , B_i són els valors de cada component de color. Aquests valors, de forma anàloga al que s'ha vist anteriorment per a les imatges en escala de grisos, representen una intensitat i prenen valors discrets entre 0 i $L-1$. Quan es fan servir 8 bits, els valors es troben entre 0 i 255. Per conveniència, els valors dels components es normalitzen considerant que els seus valors estan en un determinat rang $[0,1]$, de manera que el sistema de coordenades forma un cub on cada costat té la longitud d'una unitat (Fig. 3.12).

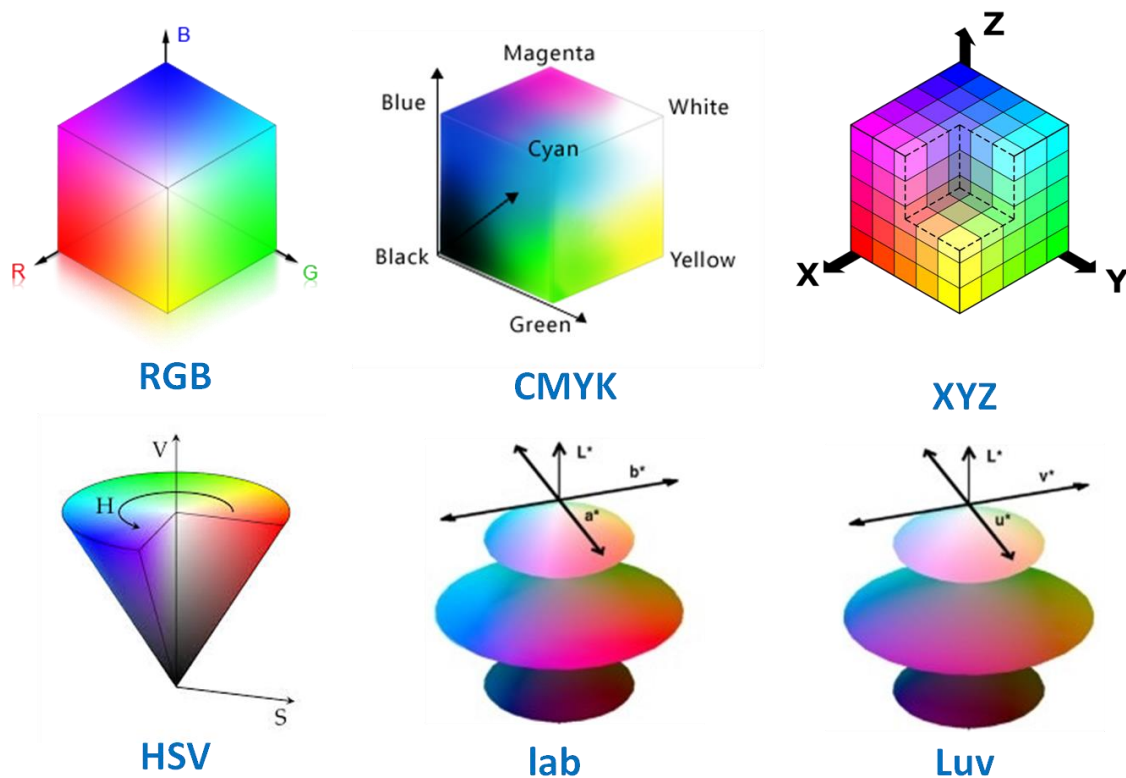


Figura 3.12: Representació dels diferents espais de colors utilitzats a la *Tesi Doctoral*.

El color negre (origen de coordenades) té components (0,0,0), mentre que el color blanc ocupa el vèrtex oposat de coordenades (1,1,1). Això indica que el negre no té gens de color, mentre que el blanc és una combinació perfecta de tots els colors primaris. Els colors purs vermell, verd i blau ocupen els vèrtexs del cub en els seus eixos corresponents: $R = (0,0,1)$, $G = (0,1,0)$ i $B = (0,0,1)$. Els colors secundaris purs ocupen la resta de vèrtexs del cub. Per exemple, el magenta té components (1,0,1) ja que és una combinació del vermell i del blau en les mateixes proporcions.

Segons el model de color RGB, una imatge està formada per tres imatges superposades, una per a cada color primari. Cadascuna d'elles és una imatge en escala de grisos i podem aplicar el que s'ha exposat anteriorment per a aquest tipus d'imatges. Així tindrem per a cadascuna d'elles una matriu de M columnes i N files de píxels. Si es treballa, per exemple, amb un rang d'intensitats de 8 bits, llavors la imatge completa en color té un rang de colors de $3 \times 8 = 24$ bits. Això vol dir que una imatge RGB de 24 bits té un nombre total de colors igual a $(2^8)^3 = 16.777.216$.

Per a l'emmagatzematge i el processament de les imatges, cada component té una matriu d'intensitats en escala de grisos diferent. Aquesta separació és molt important ja que el tractament diferenciat dels components permet obtenir descriptors molt interessants per al reconeixement d'objectes a les imatges.

CIE XYZ

Com s'ha comentat anteriorment, aquest espai de color es deriva del CIE RGB i engloba totes les sensacions de color que són visibles per a una persona amb vista mitjana (Figura 3.12). És per això que CIE XYZ és una representació de color invariable del dispositiu. Serveix com a referència estàndard enfront de la qual es defineixen molts altres espais de color. En aquest model, Y vol dir lluminositat, Z és aproximadament igual a l'estímul de blau i X és una barreja tendent a la corba de sensibilitat del vermell al verd.

CMYK o CMGN

Com s'ha indicat anteriorment, cian, magenta i groc són els colors secundaris i defineixen un sistema sostractiu i complementari al sistema RGB. En efecte, el color cian (C) sostreu llum vermella (R) de la llum transmesa per un objecte, i el mateix fa el magenta (M) amb el verd (G) i el groc (Y) amb el blau (B). D'aquesta manera, suposant que els valors del color estan normalitzats entre 0 i 1, es poden obtenir els valors CMY en la forma següent:

$$C = 1 - R$$

$$M = 1 - G$$

$$Y = 1 - B$$

En teoria, quantitats iguals de coloracions C, M i Y haurien de produir color negre. A la pràctica, com pot ésser el procés de tinció en les impressores en color, els tres colors C, M i Y no produeixen un negre perfecte, de manera que als tres colors secundaris s'afegeix el color negre com un quart component (K) per constituir així el model denominat CMYK o en català, CMGN (Fig. 3.12).

HSV

El model RGB és molt pràctic per la seva relació directa amb l'ull humà pel que fa a la percepció dels colors vermell, verd i blau i per la seva facilitat per a la captura i emmagatzematge per dispositius físics. Una cosa similar podem dir del model CMYK en relació a la impressió. L'inconvenient d'ambdós models és que no descriuen els colors en termes dels atributs que són més intuïtius per a la interpretació per part de les persones. Per exemple, és rar que algú descriuï el color d'un objecte mitjançant percentatges de vermell, blau o verd. En canvi és més natural descriure-ho en termes de matís (hue), saturació i intensitat (valor), com fa el model HSV.

Aquest model és molt útil per desenvolupar tècniques de reconeixement d'imatges basades en característiques de color naturals i intuïtives per a l'humà. El model HSV se sol representar visualment mitjançant un con o un cilindre, com es mostra a la Figura 3.12. A l'eix longitudinal del cilindre hi ha el negre a la part inferior (valor 0) i el blanc a la superior (valor 1). Movent-nos per l'eix, un pla perpendicular a l'eix caracteritza qualsevol valor de la intensitat V. És a dir, tots els punts situats en aquest pla tenen la mateixa intensitat V amb un valor entre 0 i 1.

La saturació indica el grau de puresa del color. La saturació augmenta quan augmenta la distància dels punts respecte a l'eix negre-blanc del cilindre, ja que en aquest eix tots els punts són grisos, és a dir, no tenen color. Així doncs la saturació es mesura com aquesta distància en una escala entre 0 (eix) i 1 (superfície exterior del cilindre). El valor de la saturació (S) fixa el radi d'un cilindre tal que, qualsevol que sigui el color, tots els punts tenen la mateixa puresa. Per fixar el hue (H), independentment de quins siguin la intensitat (V) i la saturació (S), considerem un cercle perpendicular a l'eix longitudinal, com s'il·lustra a la Figura 3.13. Independentment de quin sigui el radi (saturació S), situem un punt per al color vermell que serveix de referència per situar els altres colors. El hue (H) es defineix com l'angle mesurat (en una escala entre 0 i 1) al voltant del cercle.

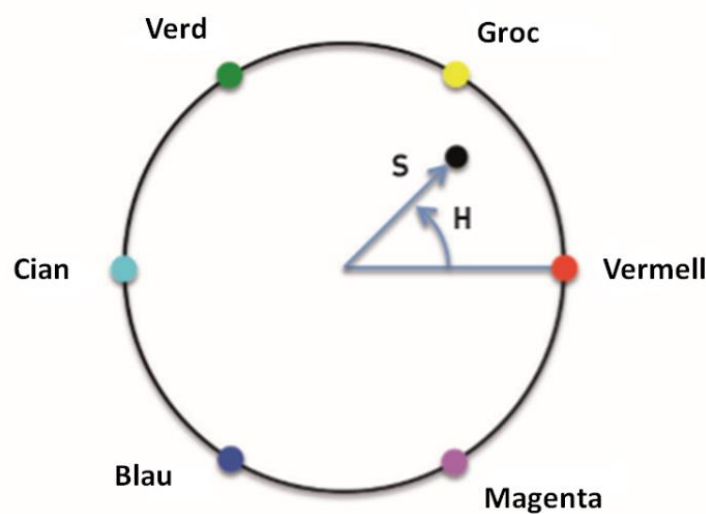


Figura 3.13: Cercle representatiu del model de color HSV, on es representa la tonalitat o hue (H) com l'angle mesurat (en una escala entre 0 i 1) al voltant del cercle amb referència al color vermell (hue = 0). La saturació (S) es representa com la distància radial al centre del cercle que té un valor de hue = 0.

Els valors per als colors primaris i secundaris estan espaiats regularment i són els següents:

Vermell (0), groc (1/6), verd (2/6), cian (3/6), blau (4/6) i magenta (5/6).

CIELAB

Es tracta del model d'elecció per a molts sistemes de gestió del color anomenat CIELAB (1977) i és un espai de color colorimètric (els colors percebuts com a coincidència estan codificats de manera idèntica), perceptualment uniforme (les diferències de color entre diferents matisos es perceben uniformement) i independent del dispositiu utilitzat.

Aquest model està format per tres components: component L^* és la lluminositat, que va de 0 (negre) a 100 (blanc), component a^* representa la gamma d'eixos vermell (valor positiu) fins al verd (negatiu) i component b^* representa la gamma d'eixos groc (valor positiu) fins al blau (negatiu) (Fig. 3.12).

Tot i no ser un format que es pugui obtenir directament (es requereix la conversió a un altre espai de color), la seva gamma engloba tot l'espectre visible i pot representar amb precisió els colors de qualsevol dispositiu de visualització o impressió. Igual que el model HSV, el model CIELAB és un excel·lent dissociador d'intensitat (representat per la luminància) i del color (representat pels dos components de crominància), fent-lo útil tant en la manipulació d'imatges (tonalitat i edició de contrast) com en aplicacions de compressió d'imatges.

CIE Luv

L'espai de color CIELUV (1976) és una modificació de l'espai estàndard CIE XYZ per tal de mostrar les diferències de color més adequadament. Aquest model també és colorimètric, uniforme, independent del dispositiu utilitzat i està format per tres components (Fig. 3.12):

- Component L^* : correspon aproximadament a la il·luminació o la brillantor. Si ambdues u^* i v^* s'estableixen a 0, a continuació, variant L^* s'obté l'escala de grisos.
- Component u^* : imita els canvis de verd a vermell (amb u^* creixent).
- Component v^* : imita colors de tipus blau i violeta majoritàriament.

L'espai CIELUV és especialment útil per a barreges additives de llums, a causa de les seves propietats d'addició lineals.

3.4.2.2 Descriptors de primer ordre basats en l'histograma

L'histograma és una forma molt pràctica de descriure quantitativament els píxels d'una imatge i és la base del càlcul dels descriptors de color. Com s'ha vist anteriorment, una imatge digital està formada per un nombre concret de píxels i cada un d'ells es caracteritza pel nivell d'intensitat, el qual pren valors enters dins d'un determinat rang.

L'histograma és una funció discreta que mostra el nombre de píxels $H(i)$ a la imatge que té un valor d'intensitat de píxels i freqüències:

$$p(i) = H(i) / L, \quad i = 0, 1, \dots, L-1 \quad (3.10)$$

L'histograma es representa mitjançant un gràfic de dos eixos en què l'eix horitzontal representa la intensitat de gris dins d'un rang de 0 (negre) a un valor màxim normalitzat (blanc), dividit en un nombre discret d'interval. L'eix vertical dona el nombre de píxels en la imatge amb les intensitats corresponents a cada interval. L'histograma també es pot interpretar com una densitat de probabilitat d'ocurrència dels valors d'intensitat si les freqüències es divideixen pel nombre de píxels de la imatge. La suma de tots els valors d'un histograma normalitzat és igual a 1.

A la Figura 3.14 es mostra una cèl·lula limfoide anormal en escala de grisos i el seu corresponent histograma de les intensitats. Es pot observar que el rang d'intensitat va de 0 a 255, per tant, és una imatge de 8 bits.

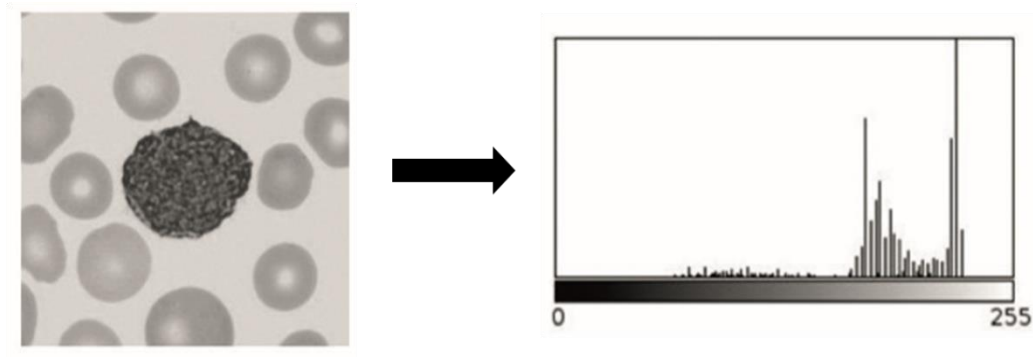


Figura 3.14: Imatge en escala de grisos d'un limfòcit anormal de sang perifèrica (esquerra) i el seu corresponent histograma (dreta).

Per tant, l'histograma no conté informació sobre la distribució espacial dels píxels en la imatge (a diferència de la matriu de co-ocurrència, que es veurà més endavant) i és possible que dues imatges diferents tinguin histogrames idèntics.

Els histogrames de les imatges a color poden ser principalment de dos tipus:

1. D'intensitat o lluminositat de la imatge: és l'histograma de la imatge en escala de grisos (Figura 3.14). La imatge original a color és convertida a escala de grisos calculant la intensitat de cada canal de la imatge. Si és una imatge RGB, es calcula la intensitat dels canals R, G i B i després es realitza una suma ponderada per obtenir la intensitat total.
2. De cada canal de color: s'obté l'histograma per a cada canal de la imatge per separat (Figura 3.15). Aquest histograma és de gran utilitat perquè mostra característiques de la imatge que l'histograma d'intensitats no mostra.

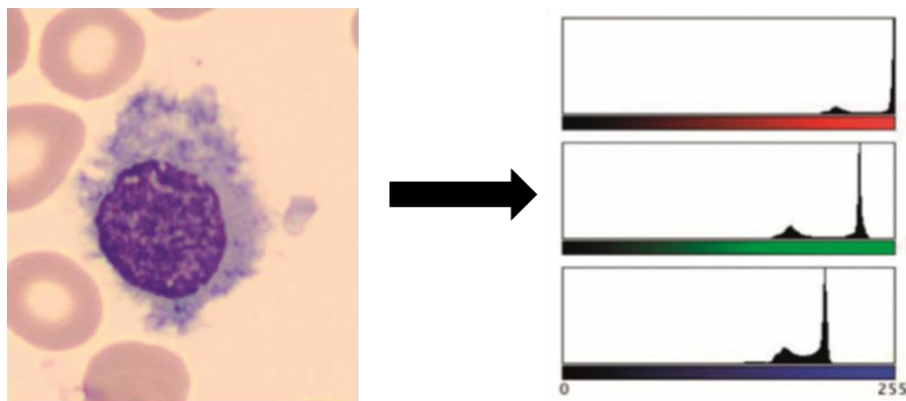


Figura 3.15: Histogrames de cada canal de color de la imatge del tricoleucòcit.

A continuació es detallen els sis descriptors de color, també coneguts com estadístics de primer ordre, que es poden calcular a partir de l'histograma.

Mitjana(μ)

La mitjana proporciona una mesura de la intensitat mitjana de la imatge ja que és la mitjana de tots els valors d'intensitat dels píxels de la imatge o d'una regió d'interès en concret (cèl·lula, nucli o citoplasma) i es calcula a partir de la següent fórmula:

$$\mu = \sum_{i=0}^{L-1} i p(i) \quad (3.11)$$

Desviació estàndard (DE)

La desviació estàndard és una mesura que proporciona una comprensió de la propagació d'intensitats a través de la imatge i alhora és un indicador de contrast en la imatge. Es calcula amb la següent fórmula:

$$DE = \sum_{i=0}^{L-1} (i - \mu)^2 p(i) \quad (3.12)$$

Asimetria (A) o *skewness*

L'asimetria o *skewness* descriu el grau de simetria de la distribució d'intensitats de la imatge. Si l'asimetria és negativa, les intensitats dels píxels s'estenen més a l'esquerra de la mitjana (intensitats baixes) que a la dreta (intensitats brillants). Si l'asimetria és positiva, les intensitats dels píxels es troben més cap a la dreta. Davant d'una distribució perfectament simètrica, l'asimetria és zero.

L'asimetria d'una distribució es defineix com:

$$A = \sigma^{-3} \sum_{i=0}^{L-1} (i - \mu)^3 p(i) \quad (3.13)$$

Curtosi (Curt)

La curtosi és una mesura del grau d'aplanament de l'histograma d'intensitats. En una distribució normal, la curtosi pren un valor de zero. La fórmula de la curtosi es detalla a continuació:

$$Curt = \sigma^{-4} \sum_{i=0}^{L-1} (i - \mu)^4 p(i) - 3 \quad (3.14)$$

Energia (E₁)

L'energia és una mesura que mostra com es distribueixen els valors de nivell gris dins de la imatge. Té una relació inversa amb l'entropia, en què l'energia d'una imatge serà la més alta si només té un valor de nivell gris. Com més nivells de grisos estiguin presents en un objecte, menor serà la seva energia. L'energia de l'histograma dona informació de la uniformitat i es calcula mitjançant la següent fórmula:

$$E_1 = \sum_{i=0}^{L-1} [p(i)]^2 \quad (3.15)$$

Entropia (H_1)

L'entropia proporciona una mesura de la suavitat d'una imatge en termes de nivell de gris. Com més alta sigui l'entropia, més nivells de grisos estan presents a la imatge. L'entropia descriu la variabilitat de l'histograma segons l'expressió:

$$H_1 = \sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (3.16)$$

3.4.3 Descriptors de textura

En aquest apartat es descriu en primer lloc el concepte de textura i a continuació s'expliquen quatre tipus de descriptors de textura: estadístics de segon ordre o basats en la matriu de co-ocurrència, granulomètrics, *Wavelet* i Gabor. Tots ells són d'aplicació en el camp de la citologia de sang perifèrica i s'utilitzen en aquest treball.

3.4.3.1 El concepte de textura

La textura en una imatge digital es defineix per la uniformitat, la densitat, el to dels píxels i les seves relacions espacials. A mode d'exemple, si observem la imatge A de la Figura 3.16, que correspon a una cèl·lula limfoide reactiva, direm que tant la densitat de la cromatina del nucli (en aquest cas laxa), com la basofília del citoplasma (coloració blavosa difusa), són característiques de la cèl·lula que ens aporten informació sobre la textura del nucli i del citoplasma, respectivament. Per contra, a la imatge B es mostra un limfòcit normal, el qual es caracteritza per una textura diferent al nucli, ja que la densitat de la cromatina del nucli és madura. El citoplasma també és basòfil però més escàs.

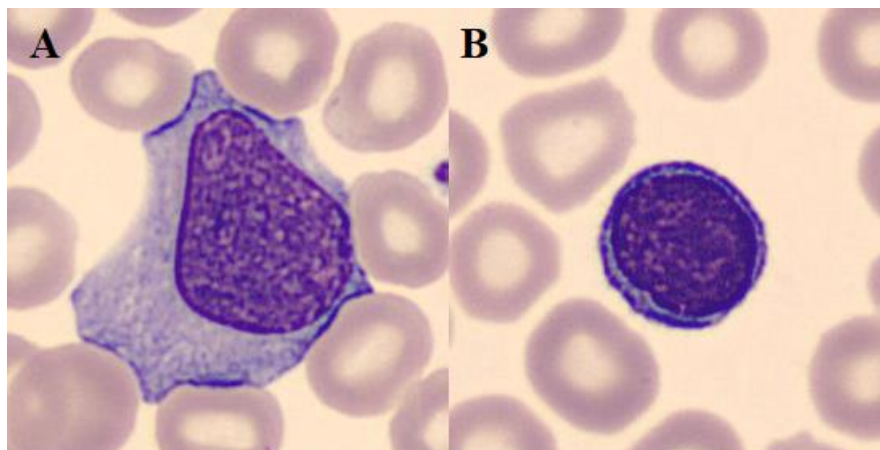


Figura 3.16: Cèl·lules limfoides de sang perifèrica amb nuclis de diferent textura.

A continuació s'expliquen quatre tipus de descriptors de textura aplicats a les cèl·lules de sang perifèrica objecte del nostre estudi. Tots ells es poden aplicar sobre els diferents components de color dels diversos espais de color que hem vist a l'apartat anterior. Hi ha diferents softwares per al càlcul de descriptors, els quals permeten extreure una gran quantitat d'informació que posteriorment és analitzada i permet diferenciar entre els diferents tipus de cèl·lules. En la present *Tesi Doctoral* s'ha utilitzat com a software el programa MATLAB®.

3.4.3.2 Descriptors de segon ordre basats en la matriu de co-ocurrència

La informació textural en una imatge està continguda en la relació espacial que els tons de gris tenen entre ells. Els descriptors estadístics de segon ordre consideren la relació espacial entre dos píxels veïns i es basen en la matriu de co-ocurrència del nivell de gris (*Gray Level Co-occurrence Matrix* o GLCM) d'una imatge digital (88).

La GLCM és un histograma de dues dimensions que especifica la freqüència en què dos nivells de grisos es troben en parelles de píxels separats per una distància i direcció particular. Per tant, considera la probabilitat conjunta $P(i, j)$ que un parell de píxels amb valors d'intensitat i i j , es trobin a una distància d en una direcció particular θ , expressada per un angle, per exemple: 0° , 45° , 90° i 135° . Aquesta probabilitat es pot calcular com el recompte de freqüències d'ocurrències dividit pel nombre total de píxels veïns.

Com a exemple il·lustratiu, la Figura 3.17 mostra com es calcula la GLCM per a una imatge corresponent a un limfòcit anormal de la leucèmia limfàtica crònica obtinguda amb el sistema DM96 (I). En primer lloc, es considera un component concret per un model de color determinat. En aquest exemple, es fa servir el component cian del model CMGN, el qual es transforma a escala de grisos mitjançant MATLAB (II). Les intensitats dels píxels s'agrupen en un nombre de nivells que es decideix *a priori*. En aquest cas el nombre és vuit (III). El nivell més baix (1) correspon amb la intensitat mínima (negre) i el més alt correspon al més brillant (blanc). D'aquesta forma, la matriu GLCM té una mida de 8×8 elements o nivells d'intensitat.

Per últim, la GLCM s'ha construït a partir de la regió del nucli tenint en compte el píxel veí ($d = 1$) i la direcció nord-oest ($\theta = 135^\circ$). Si ens fixem en el valor de la posició (7,7) de la GLCM, que és 1.445, aquest indica que el píxel amb un nivell d'intensitat set es troba al costat en direcció nord-oest d'una altre d'intensitat set 1.445 vegades. D'altra banda, la posició (1,6) indica el nombre de vegades que un píxel amb un nivell d'intensitat 1 es troba al costat en la direcció nord-oest d'un píxel amb nivell d'intensitat sis. Com es pot observar, només es produeix en una ocasió.

Aquest mètode permet extraure una gran quantitat d'informació de textura de la imatge per la gran varietat de descriptors que és possible obtenir d'aquesta matriu, els quals fan possible caracteritzar amb un conjunt de valors quantificables cada imatge analitzada. En aquesta *Tesi Doctoral*, s'han utilitzat GLCM calculades en quatre direccions ($\theta = 0^\circ$, 45° , 90° i 135°) considerant el píxel adjacent ($d = 1$). Per a fer que aquest

descriptor sigui invariant respecte a la rotació, s'ha fet la mitjana de les GLCM obtingudes a partir de les quatre direccions.

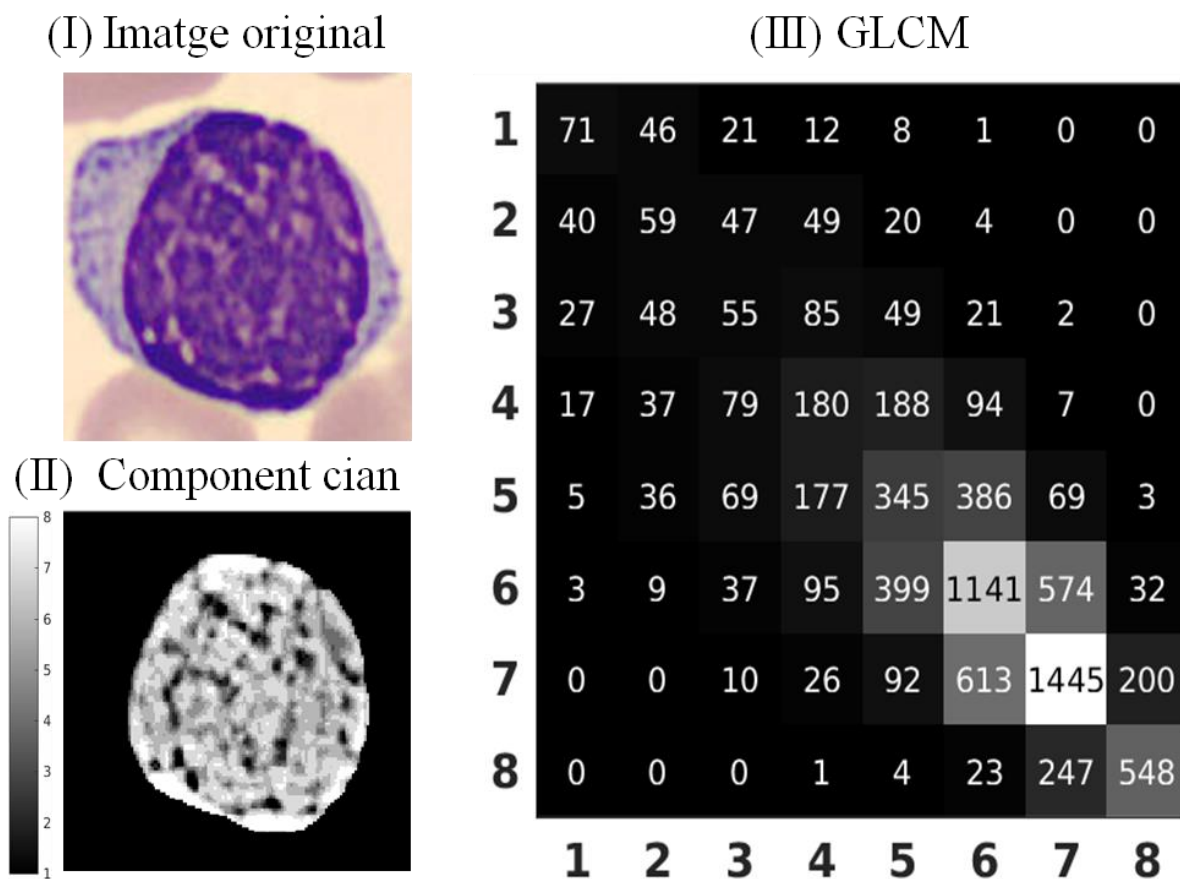


Figura 3.17: Càlcul la matriu de co-ocurrència del nivell de gris (*Gray Level Co-occurrence Matrix* o GLCM).

Tenint en compte que la imatge en escala de grisos té N nivells de grisos i els seus píxels poden prendre valors entre el rang de $[1, N]$, un total de 17 paràmetres estadístics poden ser calculats a partir de la GLCM, \mathbf{P}_d , com l'energia, entropia, probabilitat màxima, contrast, homogeneïtat, correlació, variància i *cluster shade*, entre d'altres (89,90). A continuació se'n detallen alguns.

Energia o moment angular de segon ordre (E_2)

L'energia és una mesura que mostra com es distribueixen els valors de nivell de gris dins de la imatge. És la suma dels quadrats de les probabilitats per a totes les parelles de píxels:

$$E_2 = \sum_{ij} [P_d(i, j)]^2 \quad (3.17)$$

Proporciona informació sobre la uniformitat de la imatge. L'energia d'una imatge és més alta quan només hi ha un valor de nivell de gris. Com més nivells de grisos estiguin presents en un objecte, menor serà la seva energia. Aquesta mesura dóna valors alts quan en la matriu de co-ocurrència té poques entrades de gran magnitud, i és baixa quan totes les entrades són similars. Si tots els píxels són iguals, l'energia és mínima.

A la Figura 3.18 es presenta com es calcula l'energia d'una GLCM agafant d'exemple el component de la saturació (model de color HSV) de la regió del nucli. Es consideren dues imatges de limfòcits anormals amb diferents valors d'energia (I):

- cèl·lula del mantell variant blàstica (LM-B): energia baixa (0,0655)
- cèl·lula limfoide anormal de la leucèmia limfàtica crònica (LLC): energia alta (0,0842)

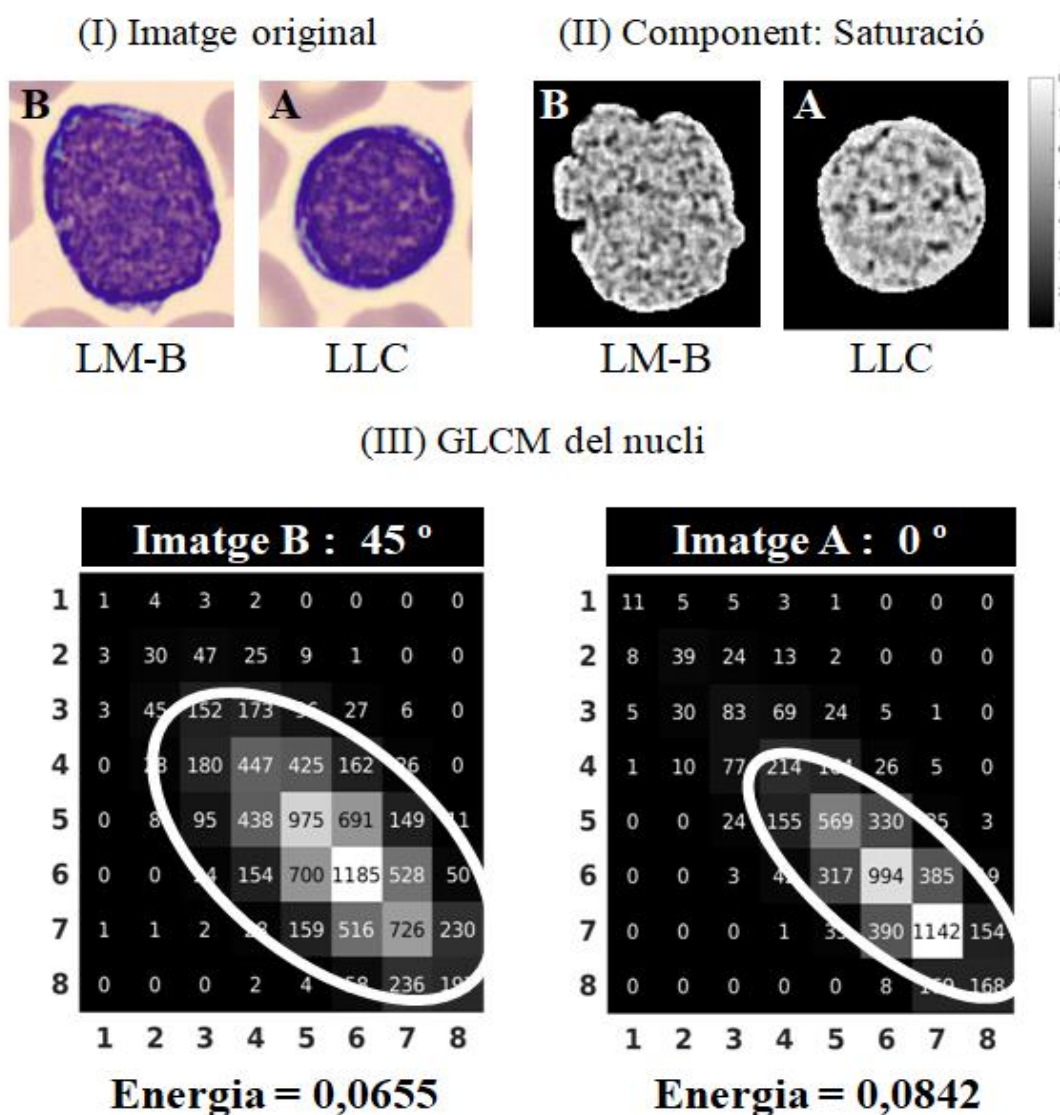


Figura 3.18: Càlcul de l'energia a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de la saturació del model de color HSV per la regió del nucli.

En primer lloc, s'obtenen les corresponents imatges en escala de grisos per al component de la saturació (II). A continuació la GLCM es calcula per a la regió del nucli amb $d = 1$ (es tria el píxel adjacent) i $\theta = 0^\circ$, 45° , 90° i 135° per a les dues imatges. Per motius d'espai només es mostra una GLCM per imatge (III).

Tal com es veu a la cèl·lula limfoide anormal de la LLC (imatge A), el nucli mostra una cromatina condensada i presenta un valor d'energia més alt a causa de la seva major uniformitat ja que hi ha menys dispersió en els nivells d'intensitat de la seva GLCM (el·lipse més petita), major energia. Per contra, en el cas del nucli del LM-B (imatge B) s'obté una major dispersió dels nivells d'intensitat de la GLCM, que significa una menor uniformitat de la cromatina i un menor valor d'energia (el·lipse més gran, els valors es troben menys concentrats).

Entropia (H_2)

Aquest descriptor mesura l'aleatorietat de la imatge, aconseguint el seu màxim quan tots els elements de la matriu de co-ocurrència són iguals. Es calcula a partir de la fórmula següent:

$$H_2 = - \sum_{i,j} P_d(i,j) \log(P_d(i,j)) \quad (3.18)$$

L'entropia té una relació inversa amb l'energia.

Contrast o inèrcia (I)

És una mesura de la variació local en una imatge d'acord amb l'expressió:

$$I = \sum_{i,j} (i - j)^2 P_d(i,j) \quad (3.19)$$

Arriba a prendre un valor alt quan la imatge té molt contrast i un valor baix quan els valors alts de la matriu GLCM són a prop de la diagonal principal. La forma més senzilla de calcular el contrast de la textura utilitza les intensitats de nivell de grisos i la matriu de co-ocurrència.

Homogeneïtat o *inverse differential moment* (IDM)

L'homogeneïtat és alta quan la matriu de co-ocurrència es concentra al llarg de la diagonal. Això passa quan la imatge és localment homogènia. És l'oposat al contrast i es calcula mitjançant la fórmula:

$$IDM = \sum_{i,j(i \neq j)} \frac{P_d(i,j)}{(i-j)^2} \quad (3.20)$$

Correlació (C)

Mesura la correlació d'un píxel amb el seu veí. Es calcula mitjançant la fórmula següent:

$$C = \frac{1}{\sigma_i \sigma_j} \sum_{i,j} (i - \mu_i)(j - \mu_j) P_d(i,j) \quad (3.21)$$

On μ_i i μ_j són els valors mitjans de les probabilitats marginals del píxel i del seu veí.

A continuació es mostra un exemple d'aquest descriptor. En concret, la Figura 3.19 presenta la correlació de la GLCM del component blau de la cèl·lula per a una cèl·lula limfoide reactiva i una cèl·lula blàstica. En primer lloc es mostren (I) dues imatges originals corresponents a una cèl·lula blàstica (imatge 1) i una reactiva (imatge 2) i, (II) el component blau en escala de grisos i (III) les seves GLCM amb $d = 1$ (es tria el píxel adjacent) i $\theta = 45^\circ$.

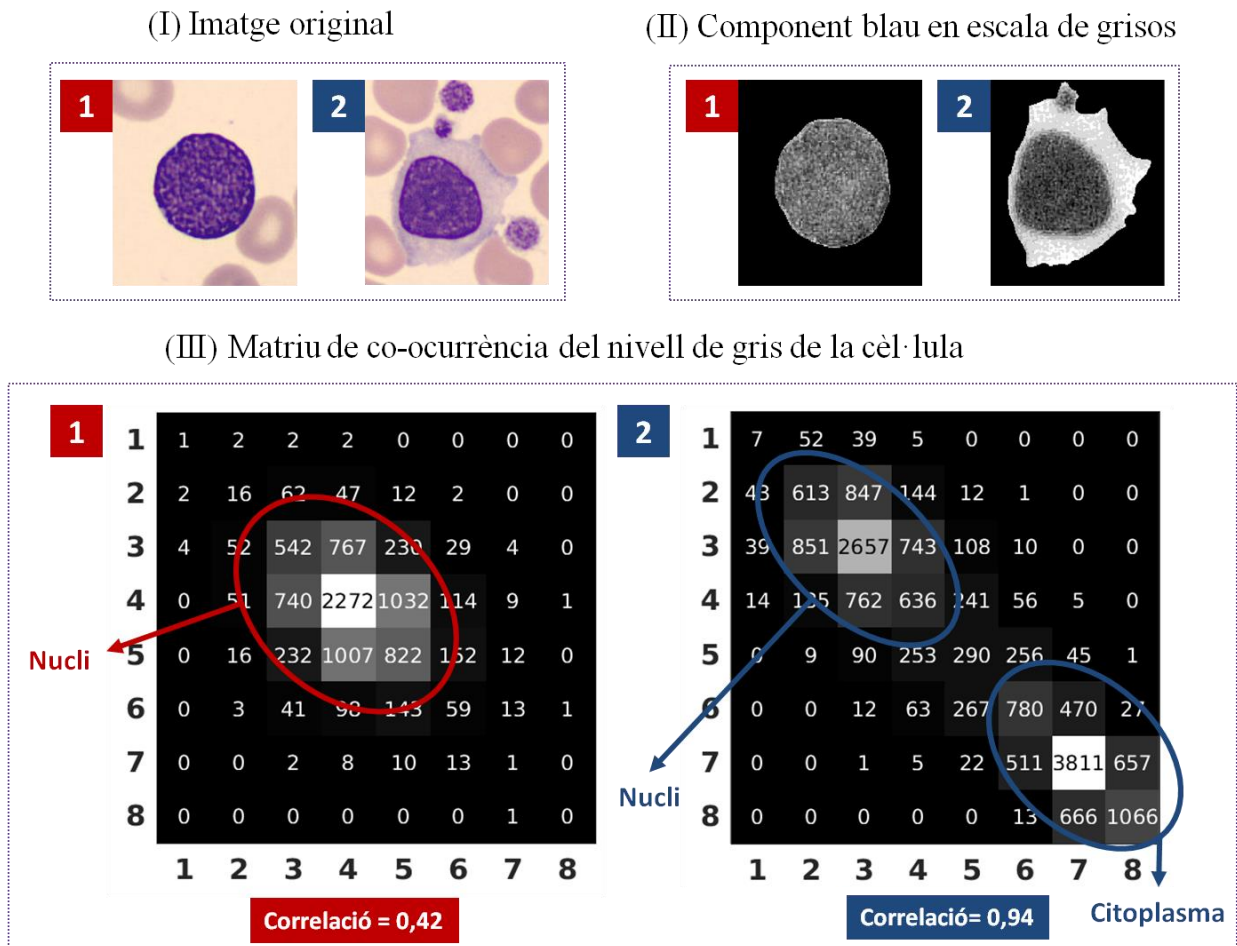


Figura 3.19: Càlcul de la correlació a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de blau del model de color RGB per la regió de la cèl·lula.

En el cas de la cèl·lula limfoide reactiva (imatge 2), la qual presenta una relació nucli/citoplasma baixa, la correcció dels píxels és elevada com a conseqüència de que els píxels d'intensitats brillants es troben propers entre si (corresponents al citoplasma), així com els d'intensitats fosques (corresponents al nucli). Per tant, la majoria dels valors es concentren a la diagonal principal de la GLCM, el que resulta en un major valor de correlació. En contraposició, la cèl·lula blàstica (imatge 1), caracteritzada per tenir una relació nucli/citoplasma elevada, presenta molts píxels amb diferents nivells d'intensitat de grisos i, per tant, hi ha diferències d'intensitat més altes entre píxels veïns. Això vol dir que la correlació és menor.

Probabilitat màxima (P)

La màxima probabilitat és la d'aquella parella d'intensitats de píxels que més es repeteix en la imatge, és a dir, la parella més probable de la GLCM. Si la probabilitat màxima és alta, indica que el valor de la mitjana de les probabilitats màximes és elevat. A continuació es mostra la seva fórmula:

$$P = \max_{i,j} P_d(i, j) \quad (3.22)$$

En aquest treball, la probabilitat màxima es basa en la mitjana aritmètica de les quatre GLCM calculades de cada direcció. A la Figura 3.20 es presenta un exemple d'aquest descriptor aplicat per a la diferenciació de cèl·lules blàstiques.

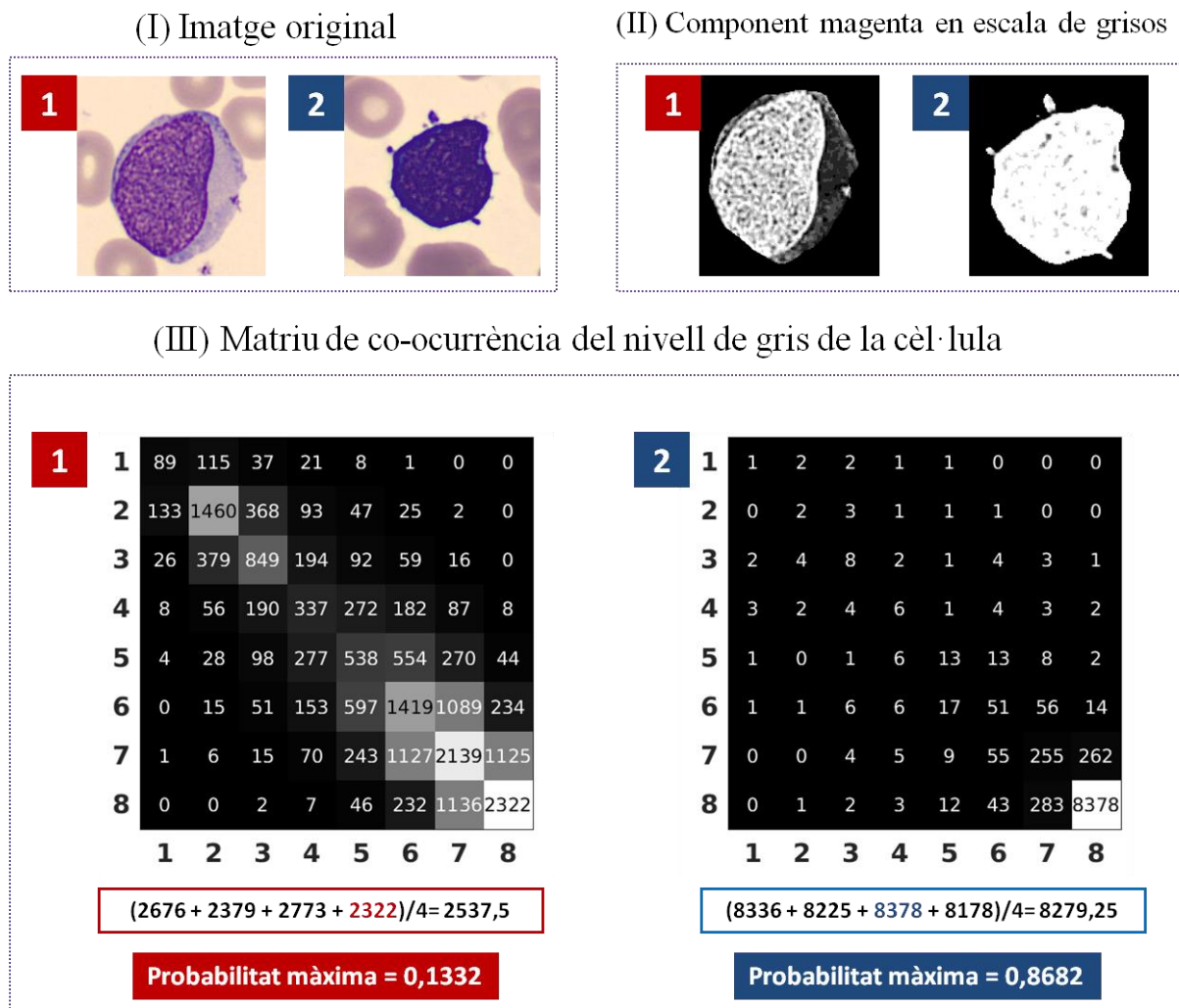


Figura 3.20: Càlcul de la probabilitat màxima a partir de la matriu de co-ocurrència del nivell de gris (GLCM) obtinguda pel component de magenta del model de color CMGN per la regió de la cèl·lula.

En concret, es mostra la probabilitat màxima del component magenta de la cèl·lula. La Figura 3.20 mostra:

- (I) dues imatges originals amb valors baixos d'aquest descriptor (imatge 1, blast mieloide) i alts (imatge 2, blast limfoide).
- (II) el component magenta en escala de grisos de cada imatge.
- (III) les seves GLCM amb $d = 1$ (es tria el píxel adjacent) i $\theta = 135^\circ$ (imatge 1) i $\theta = 90^\circ$ (imatge 2).

Si observem la imatge 2, el blast limfoide presenta una gran quantitat de component magenta al seu interior, ja que la majoria de píxels de la cèl·lula són d'intensitat brillant. Aquest fet es correlaciona amb que el parell més probable de píxels veïns sigui el més brillant, és a dir, amb intensitats de vuit en ambdós píxels, l'element (8, 8).

A la GLCM de la imatge 2 es tradueix en el fet que en 8.378 ocasions en la direcció $\theta = 90^\circ$ es troben parells de píxels d'intensitat vuit (8,8). A l'ésser un nombre tant elevat, resulta en una mitjana més elevada de probabilitat màxima. A la part inferior de la GLCM de la imatge 2 es pot veure que la mitjana de la probabilitat màxima en les quatre direccions és de 0,8682.

En canvi, si ens fixem en la imatge en escala de grisos del component magenta de la cèl·lula blàstica mieloide (imatge 1), veiem que presenta parells de píxels amb intensitats molt diferents, el que fa disminuir el valor màxim. A la GLCM, prenent la direcció $\theta = 135^\circ$, el valor màxim és de 2.322 per a l'element (8,8) seguit de 2.139 per al (7,7). Aquesta és la raó per la qual la mitjana de totes les probabilitats màximes és menor i el descriptor pren un valor inferior. A la part inferior de la GLCM de la imatge 1 es troba la mitjana de la probabilitat màxima en les quatre direccions que és de 0,1332.

3.4.3.3 Descriptors granulomètrics

Pel que fa a la granulometria, aquesta calcula la distribució de mida dels píxels foscos i brillants a la imatge (29). En una imatge en escala de grisos, generalment es consideren habitualment dos tipus de partícules: les brillants i les que són més fosques respecte al fons, les quals poden ser identificades mitjançant operacions de morfologia matemàtica, com l'erosió, la dilatació, l'obertura (*opening*) i el tancament (*closing*) (75). Aquestes operacions consisteixen en processar la imatge original amb una imatge auxiliar de mida petita (anomenada element estructural), el qual es mou sobre la imatge per extreure determinats patrons d'interès localitzats.

L'element estructural més comú és un disc amb un cert radi mesurat en píxels. L'operació d'obertura/erosió elimina de la imatge els detalls brillants que són més petits en àrea que l'element estructural. El procediment complet consisteix en aplicacions successives de l'operació morfològica utilitzant un disc com a element estructural amb una mida variable i creixent, per després sumar les intensitats de tots els píxels (a la regió) de la imatge resultant. Ja que l'obertura/erosió redueix les intensitats de les partícules brillants, aquesta

suma és decreixent i finalitza després d'un cert nombre d'iteracions. Per emfatitzar les variacions és millor calcular les diferències entre els valors obtinguts en successives iteracions de l'operació morfològica. En canvi, l'operació tancament/dilatació busca eliminar els detalls foscos menors que l'element estructural, eliminant les intensitats més baixes a través d'un procediment iteratiu similar al descrit anteriorment (veure Figura 3.21).

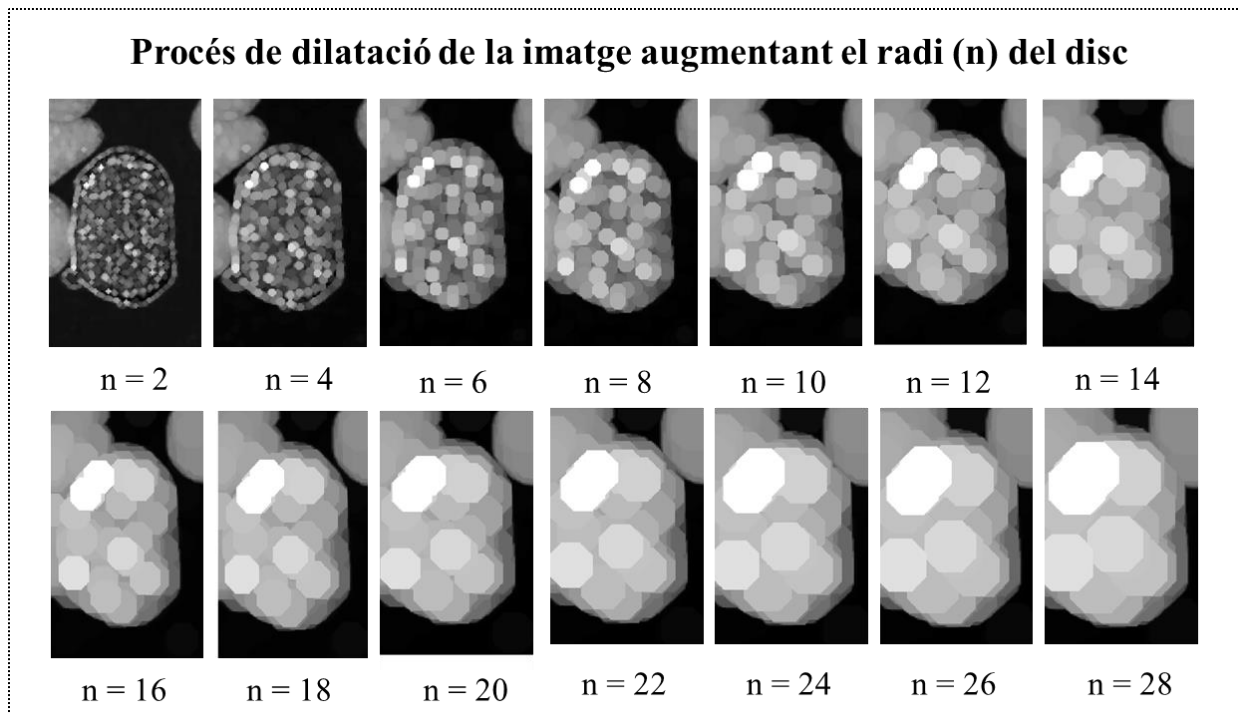


Figura 3.21: Procés de dilatació (part esquerra de la corba pseudo-granulomètrica)

La unió de tots dos processos es visualitza mitjançant la corba granulomètrica, representant a l'eix negatiu l'operació de tancament i en l'eix positiu l'operació d'obertura. De la mateixa manera, la corba pseudogranulomètrica representa la unió de l'operació de dilatació en l'eix negatiu juntament amb l'operació d'erosió. En altres paraules, en ambdues corbes es representen la granulometria de les partícules fosques a l'esquerra i la granulometria de les partícules brillants a la dreta.

La Figura 3.22 mostra la corba pseudo-granulomètrica obtinguda de la regió del citoplasma d'una cèl·lula blàstica limfoide per al component de color negre de l'espai de color CMGN a partir d'una imatge original mitjançant MATLAB®. L'eix horitzontal representa el radi de l'element estructural (disc) i l'eix vertical representa la diferència de les sumes de les intensitats de la imatge resultant després d'aplicar l'operació morfològica respectiva. La part dreta (amb valors positius de mida) representa la densitat dels elements brillants i correspon al procés d'erosió, mentre que l'esquerra (amb valors negatius de mida) correspon als elements foscos, per tant, al procés de dilatació.

Com s'observa a la Figura 3.21 i a la corba de la Figura 3.22, a mesura que augmenta el radi del disc (2, 4, 6, 8, etc), les diferències d'intensitat del citoplasma són cada vegada menors. La informació més significativa ve donada pel pic de la gràfica que correspon al que s'assenyala amb el cercle vermell, el qual indica on les diferències són més grans i és entre els discs de menor radi ($n=2, n=4$).

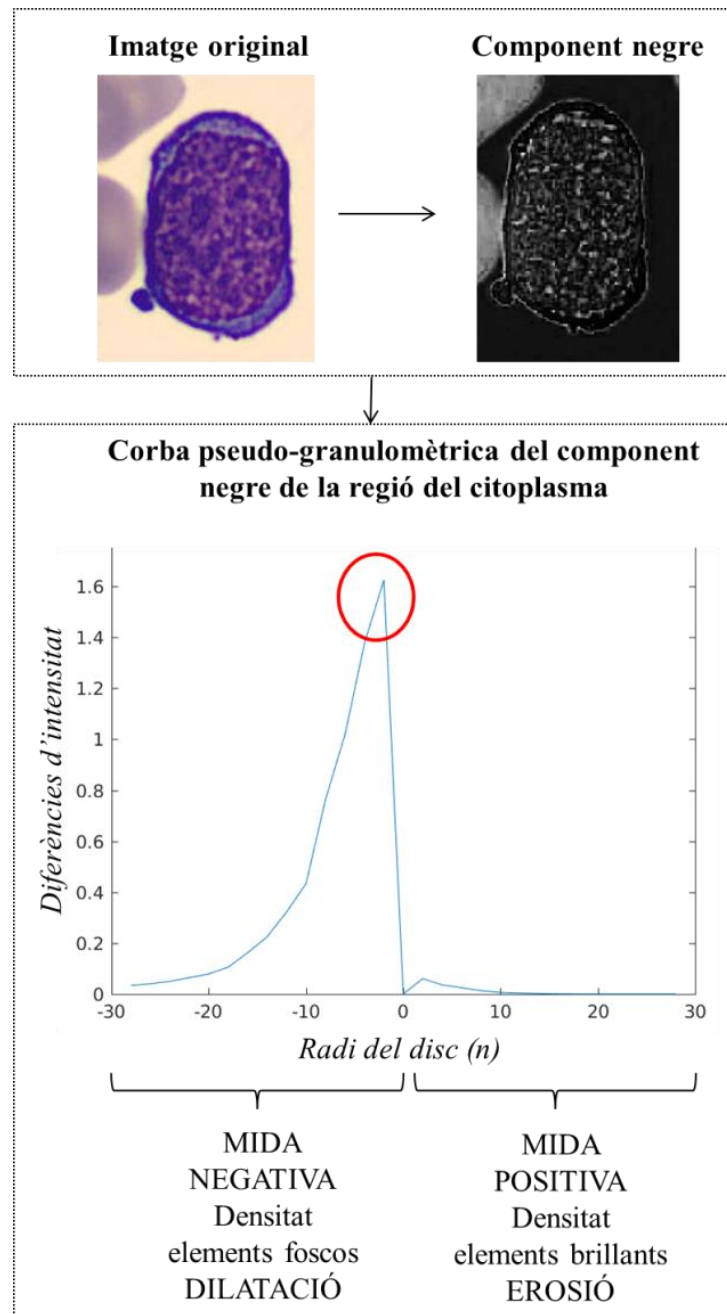


Figura 3.22: Corba pseudo-granulomètrica per a la regió del citoplasma d'una cèl·lula blàstica limfoide considerant el negre com a component de color.

Les corbes granulomètriques es fan servir per extreure informació sobre la textura. Particularment en aquesta Tesi, s'han calculat quatre paràmetres sobre cada una de les corbes: la mitjana, la desviació estàndard, l'asimetria i la curtosi. Cal mencionar que tots els paràmetres es veuran influenciats per la mida de la regió a estudi. Per exemple, si el descriptor fa referència al citoplasma, quan més gran sigui el citoplasma (més píxels tingui), major valor prendrà el descriptor.

A la Figura 3.23 s'ha calculat el descriptor de la mitjana de la corba granulomètrica i pseudogranulomètrica de dues imatges de cèl·lules limfoides anormals amb diferents graus de granulació: un limfòcit gran granular amb granulació abundant (imatge 1) i un limfòcit vellós del limfoma de la zona marginal esplènic sense granulació (imatge 2). D'ambdues imatges originals s'ha obtingut el component 'u' de l'espai de color Luv en escala de grisos per a l'obtenció de les corbes.

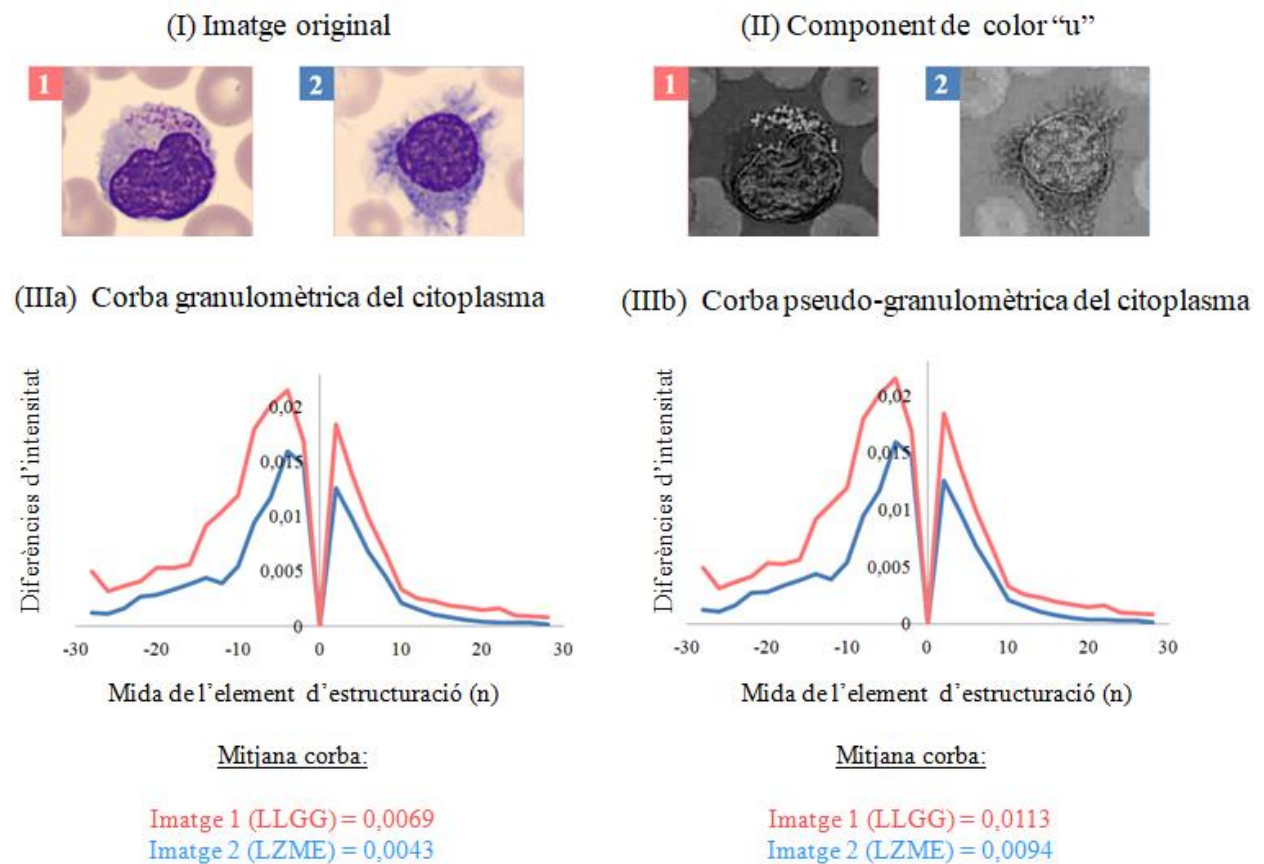


Figura 3.23: Corbes granulomètriques i pseudo-granulomètriques per a la regió del citoplasma d'un limfòcit gran granular (imatge 1) i un limfòcit anormal vellós d'un limfoma de la zona marginal esplènic considerant el "u" com a component de color.

En primer lloc, si comparem la mida dels citoplasmes d'ambdues cèl·lules, observem que a la imatge 1 el citoplasma és major. Per tant, el descriptor de la mitjana prendrà valors més elevats en comparació amb la imatge 2. A més a més, el fet que el citoplasma de la imatge 1 contingui més grànuls brillants, també fa augmentar el valor del descriptor. Això vol dir que hi ha major contribució del component "u" a la imatge. En canvi, a la imatge 2, ambdues corbes presenten mitjanes més baixes perquè hi ha menys grànuls brillants i foscos.

D'altra banda, com succeïa a la cèl·lula blàstica limfoide de la Figura 3.22, en totes les corbes (imatge 1 i 2) hi ha major proporció d'elements foscos (corba esquerra) en comparació amb els elements brillants (corba dreta) a la regió del citoplasma. Això es tradueix en pics més elevats a la part esquerra de cada corba.

Els descriptors granulomètrics així com els geomètrics i estadístics de primer i segon ordre es veuran en més detall en els capítols posteriors, en els resultats de la present *Tesi Doctoral*.

3.4.3.4 *Altres descriptors*

Descriptors estadístics *Wavelet*

Aquests descriptors de textura es calculen utilitzant la transformada *Wavelet* per obtenir una descomposició de segon nivell (una descomposició de primer nivell sobre la imatge i una altra descomposició successiva de primer nivell sobre la sub-banda d'aproximació) (91). D'aquesta manera s'obtenen sis descomposicions de la imatge en detalls (freqüències altes) i una d'aproximació (freqüències baixes). Per als dos nivells de descomposició, hi ha tres sub-bandes corresponents als detalls horitzontals (H1 i H2), diagonals (D1 i D2) i verticals (V1 i V2). A partir de les sub-bandes, es poden calcular la mitjana i la desviació estàndard. Per tant, es poden arribar a obtenir un total de 12 descriptors. A mode d'exemple, un descriptor *Wavelet* podria ser la mitjana del nucli en el component de color vermell filtrat pels detalls diagonals de la descomposició *Wavelet* de segon nivell.

Descriptors Gabor

Aquests descriptors es calculen sobre versions de la imatge processades mitjançant filtres Gabor (92). Les funcions Gabor depenen de cinc paràmetres, tres dels quals es poden deixar constants, i els dos restants, la longitud d'ona i l'orientació, són variables. Per a cada versió invariant del component de color en particular es calculen dos paràmetres: la mitjana i la desviació estàndard. A la present Tesi s'ha utilitzat un banc de 28 filtres corresponents a tres longituds d'ona $\lambda = \{8, 12, 15\}$ i a vuit orientacions entre 0 i 180 °.

3.5 Classificació

En sentit genèric, un sistema de reconeixement (o classificació) automàtic té el següent objectiu: donat un objecte individual que pertany a una població formada per diferents classes mútuament excloents, determinar a quina d'aquestes classes pertany. Per això el sistema utilitza com a entrada el conjunt de descriptors que es trien convenientment per la seva capacitat per descriure els objectes i discriminar entre les diferents classes.

Quan s'introdueix el concepte d'aprenentatge automàtic, el subjecte que aprèn és una màquina, d'aquí el terme *machine learning*, i més específicament, un ordinador. Es tracta doncs de dotar-lo d'elements que permetin que aquest millori les seves capacitats a partir de l'experiència.

Com hem comentat anteriorment, el nostre objecte és la cèl·lula limfoide de la qual s'extreuen un determinat nombre de descriptors geomètrics, de color i de textura. Els passos a seguir per a realitzar una classificació adient són els següents:

1. *Entrenament*: s'estudia el problema i a partir del coneixement de la morfologia de les cèl·lules limfoides, s'identifiquen característiques típiques (definició de descriptors) tant per al cas de limfòcits madurs com, especialment, per al cas de la presència d'una neoplàsia limfoide (limfòcits anormals). A continuació s'escriu l'algoritme d'aprenentatge automàtic per a identificar cadascuna de les característiques triades i es fixen uns criteris per separar els limfòcits anormals i donar possibles alarmes.
2. *Validació*: en aquesta etapa l'algoritme s'entrena amb un conjunt diferent d'imatges fins que es considera, d'acord amb algun criteri quantitatiu de rendiment, que està preparat per a la seva implementació definitiva. Es pot dir que l'algoritme aprèn amb l'experiència obtinguda a partir d'imatges sense una programació explícita prèvia.

Avui dia existeixen un bon nombre de metodologies d'aprenentatge automàtic que es poden utilitzar. Entre les més conegudes hi ha les màquines de suport vectorial o SVM (*support vector machines*) i les xarxes neuronals profundes (*deep neural networks*) i combinacions d'ambdues.

L'èmfasi d'aquesta tesi es centra en l'etapa posterior, que consisteix en l'anàlisi dels descriptors que són útils per al reconeixement de cèl·lules limfoides anormals circulants en sang perifèrica. Més concretament, els Capítols 4 i 6 analitzen i interpreten els 20 descriptors que han estat seleccionats en base a la teoria de la informació, utilitzant el criteri de 'màxima rellevància i mínima redundància', per tal de diferenciar entre 12 tipus de cèl·lules limfoides mitjançant dos sistemes d'adquisició d'imatges. El criteri de 'màxima rellevància i mínima redundància' calcula l'entropia condicional dels descriptors pel que fa a la presència d'una classe i els descriptors queden ordenats des del més rellevant i menys redundant al menys rellevant i més redundant.

D'altra banda, els Capítols 5 i 7 utilitzen les mateixes fonts d'adquisició d'imatges (el sistema DM96 i el BX43) però el punt de partida no és el processament digital d'imatges, sinó algunes anormalitats cel·lulars establertes *a priori* pels citòlegs, les quals estan associades a cèl·lules i malalties específiques i de particular interès clínic. L'objectiu serà identificar descriptors quantitius específics que descriguin objectivament característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius. El resultat és l'obtenció de nous descriptors quantitius citològics per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, la tricoleucèmia i el limfoma de la zona marginal esplènica, la síndrome de Sézary i la leucèmia de limfòcits grans granulars mitjançant l'anàlisi d'imatges.

3.6 Anàlisi estadística dels descriptors

Aquest apartat presenta els diferents tests estadístics aplicats als descriptors objecte de la present recerca amb l'objectiu d'assolir una caracterització morfològica de les diferents cèl·lules limfoides estudiades. Tots els tests descrits han estat executats mitjançant la interfície gràfica *R-Commander* (60) del software R (93). Abans de descriure els tests estadístics es presenten uns conceptes bàsics d'estadística (94).

3.6.1 Conceptes bàsics

- *Individu*: unitat experimental (siguin elements, persones o objectes) sobre la qual l'observador realitza les mesures i valora, ja que conté certa informació del fet que es vol estudiar.
- *Població*: conjunt d'individus que formen el col·lectiu d'estudi.
- *Mostra*: subconjunt de la població escollit per realitzar l'estudi. El nombre d'individus que formen la mostra és denominada mida de la mostra.
- *Variable*: mesura (quantitativa o qualitativa) que es vol estudiar en una mostra d'individus. La variable independent és la que mesurem per a un estudi, mentre que el valor de la variable dependent depèn dels resultats de la variable independent.

Hi ha diferents tipus de variables:

o Qualitatives o nominals: aquelles que no poden ser mesurades numèricament. Adopten valors categòrics que poden ser binaris (únicament hi ha dues categories, ex. el gènere o sexe) o no dicotòmics (hi ha més de dues categories per a definir la variable, ex. el color dels ulls).

o Ordinals: aquelles que poden ser mesurades utilitzant una escala ordinal.

o Quantitatives: aquelles que es poden mesurar numèricament, es poden quantificar. Van acompanyades d'unitats de mesura i, poden ser:

- Discretes: en les quals entre dos valors consecutius no es troba cap altre valor (ex. nombre de fills).
- Contínues: en què entre dos valors consecutius poden trobar infinits valors (ex. el pes d'un individu).

En el cas de la present recerca, *l'individu* és el frotis de sang perifèrica i la *mostra* són tots els pacients de cada una de les diferents neoplàsies limfoides estudiades, així com individus sans i pacients amb infeccions víriques. La *població* de l'estudi està formada per tots els pacients a qui se'ls faci un frotis de sang perifèrica.

En relació a les variables, principalment n'hi ha dues:

- 1) Descriptor (geomètric, de color i de textura): variable quantitativa contínua.
- 2) Grup cel·lular limfoide: variable qualitativa.

3.6.2 Estadística descriptiva

L'estadística descriptiva té com a principal objectiu descriure una població a partir de les dades d'estudi obtingudes sobre una o diverses variables. La descripció de les variables sol ser un resum representatiu d'aquestes, en forma de resum numèric (taules que representin les dades de l'estudi), o bé en forma de resum gràfic (gràfics que representin les dades de l'estudi).

Tant els resums numèrics com els gràfics poden fer-se sobre tota la població o una mostra d'aquesta. La població és el total, on s'inclouen totes les dades, mentre que a la mostra només s'inclouen les dades d'un col·lectiu concret de la població, que pot ser representatiu o no. Quan es treballa amb la població i la mostra hi ha dos mètodes de treballar les dades d'estudi:

- Inferència: passar dels casos particulars (mostra) als generals (població)
- Deducció: passar del cas general (població) als particulars (mostra)

Normalment les variables qualitatives es descriuen calculant la freqüència (absoluta, relativa o acumulativa) amb la qual es presenta cada categoria de la variable. Per descriure les variables quantitatives és més comú usar les mitjanes de tendència central (mitjana, mediana) o mesures de dispersió (desviació estàndard, variància, coeficient de variació). Tant les dades de les variables qualitatives com les de les variables quantitatives poden ser representades en forma de taules o gràfics.

3.6.2.1 Variables quantitatives

Les variables quantitatives s'agrupen en dos tipus de paràmetres: els paràmetres de localització i els paràmetres de dispersió. En primer lloc definirem els paràmetres de localització:

- Moda: és el valor més freqüent, el que té una major freqüència a la distribució de dades. Normalment s'utilitza per a variables categòriques.
- Mitjana aritmètica: és la suma de tots els valors dividit pel nombre d'observacions.

$$\bar{X}_n = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.21)$$

- Mediana: valor de la variable que ocupa el lloc central en una sèrie ordenada de dades.

En una distribució normal, la mitjana aritmètica i la mediana són iguals. La mitjana és un estadístic en què influeix molt l'homogeneïtat de les dades i, sobretot els valors atípics (o *outliers*).

Els percentils ens donen una idea dels valors on s'acumula un percentatge concret de les dades. Els quartils divideixen la distribució en quatre parts (25 %, 50 %, 75 % i 100 %). Els decils, en canvi, divideixen la distribució en deu parts i, corresponen als percentils 10 %, 20 %, 30 %, etc. Per exemple, el percentil 25 representa el nombre de valors d'una distribució que es troben per sota del 25% dels valors, mentre que el percentil 50 representa el nombre de dades que hi ha per sota del 50 % dels valors de la distribució. Es coneix com a rang interquartílic (RIQ) el nombre de dades compreses entre el percentil 25 i 75.

Els paràmetres de dispersió més habituals són:

- Variància: el sumatori de la diferència al quadrat dels valors mitjans menys la mitjana, dividit pel nombre total de dades:

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X}_n)^2 \quad (3.22)$$

- Desviació estàndard (DE): arrel quadrada positiva de la variància

$$\sigma = \sqrt{\sigma_n^2} \quad (3.23)$$

- Coefficient de variació (CV): quocient entre la DE i la mitjana i multiplicat per 100. Serveix per comparar les diferents dispersions per diferents conjunts de valors en diferents unitats de mesura, ja que el CV és independent de les unitats.

$$CV (\%) = \frac{\sigma}{\bar{x}} \times 100 \quad (3.24)$$

- Rang: diferència entre el valor màxim (x_n) i el valor mínim (x_1).

$$Rang = x_n - x_1 \quad (3.25)$$

Les variables quantitatives es poden representar mitjançant taules o gràfics per tal de visualitzar millor les dades. En el cas de la present recerca, els descriptors (variables quantitatives) s'han representat mitjançant taules (mediana i rang interquartílic) i diagrames de caixes o *box plots*. El diagrama de caixes o *box plots* permet observar d'una forma clara la distribució de les dades i les seves principals característiques, permetent, a més, comparar diversos conjunts de dades simultàniament com el que es mostra a la Figura 3.24.

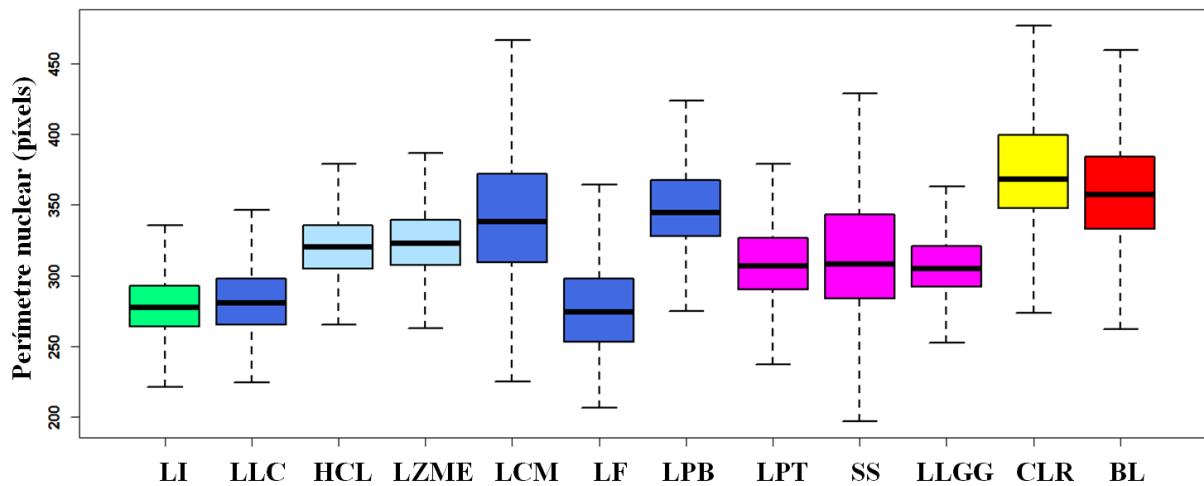


Figura 3.24: Diagrama de caixes del perímetre nuclear de 12 tipus diferents de cèl·lules limfoides.

A la Figura 3.24 es mostra una representació d'aquest tipus de gràfic, que conté una caixa o rectangle orientat en el sistema de coordenades on es representa el perímetre nuclear (en píxels) de 12 tipus diferents de cèl·lules limfoides: limfòcit normal (LI), leucèmia limfàtica crònica (LLC), tricoleucèmia (HCL), limfoma de la zona marginal esplènica (LZME), leucèmia de cèl·lules del mantell (LCM), limfoma fol·licular (LF), leucèmia prolimfocítica B (LPB), leucèmia prolimfocítica T (LPT), síndrome de Sézary (SS), leucèmia de limfòcits grans granulars (LLGG), cèl·lules limfoides reactives (CLR) i blasts limfoides (BL).

La part inferior i superior de la caixa coincideix amb el primer i tercer quartil (Q1 i Q3), respectivament, que conté el 50 % de les dades de la distribució. La línia negra horitzontal que divideix la caixa representa la mitjana o Q2. Les línies verticals discontinües (els bigotis o *whiskers*) defineixen els límits on els valors per sobre o per sota d'aquests són considerats atípics o *outliers*. Els bigotis són calculats en funció del rang interquartílic, el qual s'obté de restar el Q1 a Q3.

3.6.2.2 Variables qualitatives

Les variables qualitatives s'acostumen a agrupar mitjançant taules de freqüència. Els tres tipus de freqüència són:

- Freqüència absoluta: nombre total de vegades que es repeteix una mateixa característica, és a dir, nombre total d'individus d'una mostra que comparteixen una mateixa característica.
- Freqüència relativa: percentatge respecte al total de nombre de vegades (o individus) que es repeteix una característica.
- Freqüència acumulativa: és la suma de la freqüència estudiada més totes les anteriors.

3.6.3 Contrast d'hipòtesi

El mètode científic consisteix a validar hipòtesis sobre un esdeveniment. En primer lloc, es formula una teoria sobre un fet, es dissenya un experiment per intentar corroborar aquesta teoria i si els resultats concorden amb la teoria, es valida provisionalment. En cas contrari, es modifica o s'elabora una nova teoria. Per tant, un dels principals objectius del mètode científic es basa en contrastar diferents teories, per això l'eina que s'utilitza en estadística és el contrast d'hipòtesi (95).

Es contrasten dues hipòtesis:

- Hipòtesi nul·la (H_0): correspon a la situació estàndard o normal, normalment és la més simple.
- Hipòtesi alternativa (H_1): correspon a la situació diferent a la normal, la majoria de vegades és la negació d' H_0 .

El contrast d'hipòtesi verifica si la mostra obtinguda és compatible amb una suposició paramètrica. Són contrastos paramètrics, ja que les hipòtesis es fan sobre els paràmetres de les distribucions de la variable aleatòria que s'està estudiant. Per tant, el contrast d'hipòtesi no deixa de ser una regla que permet decidir entre dues hipòtesis a partir d'una mostra aleatòria.

En el contrast d'hipòtesi hi ha dos tipus d'error en funció de quina hipòtesi s'esculli:

- Error tipus I (α): acceptar H_1 quan en realitat la hipòtesi certa és H_0 .
- Error tipus II (β): acceptar H_0 quan en realitat la hipòtesi certa és H_1 .

Cada tipus d'error té unes conseqüències diferents: l'error tipus I és el nivell de significació que s'ha de suposar *a priori* de qualsevol experiment, ja que és la probabilitat de rebutjar l' H_0 quan aquesta és certa (normalment en els tests estadístics els criteris òptims són que $\alpha = 0,05$, és a dir, un 5 % de possibilitats de cometre un error α). En canvi, l'error tipus II és la probabilitat de rebutjar l' H_0 quan aquesta és falsa i la seva probabilitat associada s'anomena β o potència del test (normalment en els tests estadístics els criteris òptims són que $\beta = 0,2$, és a dir un 20 % de possibilitats de cometre un error β).

Sempre s'ha d'intentar minimitzar els dos tipus d'error. Per reduir simultàniament els dos errors en un contrast d'hipòtesi hi ha la possibilitat d'incrementar la mida mostral. Incrementant la mida de la mostra disminueix la dispersió de les mitjanes mostrals i així, la probabilitat de confondre-les és menor.

Altres conceptes que s'han d'introduir abans d'entrar en l'explicació de cada un dels tests estadístics són els conceptes de *p-valor* i el tipus de mostra.

El *p-valor* es defineix com el nivell de significació més petit possible que es pot escollir amb el qual s'accepta la hipòtesi alternativa. Permet resumir objectivament el resultat d'un experiment i es coneix un cop acabat aquest, ja que depèn de les dades i la mostra. Valors petits del *p-valor* indiquen que és molt poc

freqüent trobar mostres com la que s'està estudiant, mentre que valors grans del *p-valor* signifiquen que hi ha moltes mostres com la de l'estudi.

La decisió de les hipòtesis es pren en funció del *p-valor* i el nivell de significació α :

Si $p\text{-valor} \leq \alpha \rightarrow$ acceptar H_1

Si $p\text{-valor} \geq \alpha \rightarrow$ acceptar H_0

Pel que fa al tipus de mostra, les comparacions de dos grups es basen en recollir una mostra aleatòria simple de cada grup, sent així les dades independents dins del grup. És important distingir si les dades entre els dos grups tenen dependència entre elles o són independents, ja que depenent del tipus de dades es faran servir uns mètodes o uns altres.

- Les mostres independents són mesures realitzades en dos conjunts d'elements diferents.
- Les mostres aparellades són mesures dependents en un mateix conjunt d'elements.

Per entendre els dos tipus de mostres que existeixen es presenten dos exemples. En un primer cas, es vol comparar l'eficàcia de dos fàrmacs per aprimar en una mostra de 20 individus, on 10 d'aquests individus són assignats a l'atzar a cada un dels dos tractaments. Es mesura el pes dels individus abans i després dels dos mesos de tractament, per comparar la pèrdua de pes entre els dos grups. Es tracta de dues mostres independents, per què les mesures corresponen a persones diferents. En un segon cas, es trien 10 individus aleatòriament, es mesura el seu pes i se'ls subministra durant dos mesos un tractament farmacològic i, al final de l'estudi es torna a mesurar el seu pes. Es tracta de mostres aparellades, ja que la comparació es fa dins el mateix conjunt d'individus.

3.6.4 Normalitat

Una distribució de probabilitat especifica la probabilitat d'obtenir una observació en un rang determinat de valors i una distribució normal és una corba en forma de campana, amb una alta probabilitat d'obtenir una observació propera a la mitjana i les probabilitats més baixes a mesura que s'avança en l'eix. Una distribució normal es pot descriure completament amb només dos paràmetres, la mitjana i la desviació estàndard. Els tests estadístics que assumeixen que les dades es poden descriure mitjançant aquests dos paràmetres, s'anomenen proves paramètriques (ex. ANOVA).

Moltes variables biològiques encaixen bastant bé amb la distribució normal. Aquest fet s'explica pel teorema del límit central, que diu que quan es pren un gran nombre de nombres aleatoris, les mitjanes d'aquests nombres es distribueixen aproximadament de manera normal. Però quan les dades no segueixen una distribució normal, s'han d'aplicar proves no paramètriques (ex. Kruskal-Wallis).

Hi ha dos tipus de tècniques per a verificar la normalitat (96):

- Tècniques gràfiques: es poden representar les dades en un histograma i veure si recorden a una distribució normal, és a dir, a la típica campana de Gauss.
- Tècniques d'inferència: es basen en proves estadístiques, que segons el tipus de mostra i variables s'aplicarà una o una altra:
 - *Test de Shapiro-Wilk*: s'utilitza per a mostres petites (menys de 50 individus).
 - *Test de Kolmogorov-Smirnov*: s'utilitza per a mostres grans, on les dades són contínues.
 - *Test Chi-quadrat*: per a variables categòriques.

El contrast d'hipòtesis per a verificar la normalitat és:

- H_0 = hi ha normalitat dels residus (la variable segueix una distribució normal).
- H_1 = no hi ha normalitat dels residus (la variable no segueix una distribució normal).

Cal remarcar però que en casos de no normalitat, el primer que s'ha de fer és una transformació de les dades. Transformar les dades no deixa de ser una operació matemàtica (logarítmica, exponencial, inversa, etc) que es fa en cadascuna de les observacions, per utilitzar aquests valors com a dades per al test estadístic. Tot i així, si les dades no segueixen una distribució normal, l'única opció vàlida serà l'aplicació de proves no paramètriques.

En relació a la present *Tesi Doctoral*, donat l'elevat nombre d'imatges utilitzat per cada descriptor i la naturalesa contínua dels descriptors, s'ha aplicat el test de *Kolmogorov-Smirnov* per a verificar la normalitat.

3.6.5 Homoscedasticitat

Un cop hem verificat si la nostra variable és normal, s'ha de comprovar la igualtat de les variàncies en tots els grups. Parlem d'homoscedasticitat quan la desviació estàndard dels diferents grups és la mateixa mentre que quan és diferent, parlem d'heteroscedasticitat (96). Tots els nivells d'un factor haurien de tenir la mateixa variabilitat, ja que l'òptim seria treballar amb homoscedasticitat i dissenys balancejats.

A continuació es detallen els diferents tests estadístics que s'utilitzen per a comprovar que les variàncies siguin iguals:

○ *Test de Bartlett*: s'utilitza quan es té una variable quantitativa i una variable nominal. La hipòtesi nul·la comprova que les desviacions estàndard de la variable quantitativa són iguals per als diferents grups. No és adient en casos de distribucions no normals i heteroscedasticitat.

○ *Test de Levene*: és un altre test estadístic que es pot usar per a verificar l'homoscedasticitat, que no és tan robust davant desviacions de la normalitat. No obstant això, si les dades presenten una distribució bastant normal, aquest test perd potència respecte al test de Bartlett.

o *Test de Fligner-Killeen*: és un test no paramètric per estudiar l'homogeneïtat de les variàncies dels grups, i ho fa basant-se en rangs. És molt útil quan les dades de l'estudi no segueixen una distribució normal o presenten valors atípics o *outliers*. Acostuma a ser molt robust davant desviacions de la normalitat.

El contrast d'hipòtesi per a verificar l'homoscedasticitat és:

- H_0 = hi ha igualtat de variàncies entre grups.
- H_1 = no hi ha igualtat de variàncies entre grups.

En casos d'heteroscedasticitat, el primer que s'ha de fer és comparar les desviacions estàndard dels diferents grups i veure si són diferents entre elles. Malgrat tots els estudis que s'han fet sobre aquest tema, sembla ser que no existeix un consens establert sobre què fer quan l'heteroscedasticitat suposa un problema greu que no pot ser solucionat amb un test que assumeix homoscedasticitat. En cas d'haver-hi una gran diferència de les desviacions estàndard entre grups, la primera solució és fer una transformació de les dades. Si malgrat la transformació segueixen sent molt heterogènies, es pot realitzar el test de Kruskal-Wallis.

Donat que tots els descriptors estudiats a la present recerca no segueixen la distribució normal, per tal de verificar l'homogeneïtat de les variàncies s'ha aplicat el test de *Fligner-Killeen*.

3.6.6 Comparació de dues mostres

Un cop avaluada la normalitat i la homogeneïtat de les variàncies, es procedeix a realitzar una anàlisi estadística que permeti conèixer si existeixen diferències significatives d'un descriptor entre els diferents tipus de cèl·lules limfoides estudiades.

Existeixen dues alternatives no paramètriques que no són tant sensibles a desviacions de la normalitat. Es detallen a continuació (96):

o *Test de Kruskal-Wallis*: és l'alternativa no paramètrica a l'ANOVA d'un factor, és a dir, no assumeix que les dades segueixin una distribució normal. Consisteix en convertir els valors de les observacions en els seus rangs, el valor més baix adopta el rang 1, el segon més baix el rang 2, i així successivament. A més, assumeix que els diferents grups presenten la mateixa distribució.

El contrast d'hipòtesi per a verificar les diferències entre els grups d'estudi és:

- H_0 = no hi ha diferències estadísticament significatives entre els grups.
- H_1 = hi ha diferències estadísticament significatives entre els grups.

o *Test de Mann-Whitney-Wilcoxon*: és l'alternativa no paramètrica al test t de Student per a mostres aparellades. Permet decidir si les distribucions en els diferents grups d'estudi són iguals sense assumir que les dades segueixen una distribució normal. S'usa per a variables nominals dicotòmiques.

En relació a la Tesi Doctoral, els descriptors geomètrics, de color i de textura analitzats en imatges de diferents tipus de cèl·lules limfoides són variables independents, per tant, s'ha aplicat el test no paramètric de *Kruskal-Wallis* per verificar si un determinat descriptor presenta diferències estadísticament significatives entre els diferents grups cel·lulars limfoides.

3.6.7 Comparació múltiple

Un cop el test de *Kruskal-Wallis* ha estat significatiu i s'accepta la hipòtesi alternativa que implica que hi ha diferències entre els diferents grups d'estudi, l'objectiu és esbrinar quins grups cel·lulars limfoides presenten les diferències.

Hi ha diferents tipus de tests estadístics per fer comparacions múltiples, els quals es diferencien entre si per com garanteixen el nivell de significació global, és a dir, en com imposen restriccions a nivell individual de cada grup. Existeixen al voltant de 30 proves estadístiques per fer comparacions múltiples (94). En aquest apartat només es citen les més usades:

- *Test de Bonferroni*: és un dels test més conservadors i generals que hi ha. Realitza la prova t de Student, però amb els nivells de significació individual que equivalen al global, és a dir, dividit pel nombre de comparacions de parelles entre grups que s'hagi realitzat. Es basa en fer comparacions 2x2 entre els grups d'estudi.
- *Test de Tukey*: és un test una mica més liberal que l'anterior i es basa en comparar el valor absolut de la diferència de les mitjanes aritmètiques amb un valor crític (obtingut de la taula del rang t de Student). Si la diferència entre mitjanes és superior a aquest valor crític tabulat, es rebutja la hipòtesi nul·la. També fa comparacions dos a dos entre grups.
- *Test de Dunnet*: en aquest cas no interessa fer comparacions dos a dos entre grups d'estudi. Es pot donar el cas que una de les poblacions sigui un control i interessa comparar tota la resta de poblacions respecte al grup control, o bé comparar totes elles sense tenir en compte el grup control.

Tots aquests tests estadístics permeten conèixer entre quins grups d'estudi hi ha diferències significatives i entre quins no. A la present recerca (Capítols 4-7) s'ha utilitzat el test de *Kruskal-Wallis* juntament amb el *test de Dunn aplicant l'ajust de Bonferroni* per a la comparació múltiple utilitzant el paquet 'PMCMR' (20). Per a minimitzar la influència dels valors atípics, s'han pres els valors de les medianes (quartil 50) en comptes dels valors de la mitjana com a valors representatius per a la comparació múltiple de tots els descriptors. *P* valors inferiors a 0,05 s'han considerat estadísticament significatius.

3.6.8 Correlació

La correlació és una anàlisi bivariant que permet obtenir tres tipus de resultats:

1. *Conèixer si hi ha una associació entre les dues variables:* a mesura que augmenta la variable X, la variable Y tendeix a canviar (augmentant o disminuint). Mitjançant un test d'hipòtesi, es pot conèixer si existeix associació entre variables:
 - H_0 = no hi ha associació entre les dues variables.
 - H_1 = hi ha associació entre les dues variables.
2. *Mesurar la força de l'associació:* descriure quant d'associades es troben les dues variables. Normalment el valor del coeficient de correlació varia entre +1 i -1 (o r^2 que oscil·la entre 0 i 1). Un valor de ± 1 indica un grau d'associació perfecte entre les dues variables. En canvi, quan el valor del coeficient de correlació va cap a 0, la relació entre les dues variables és més feble. Quan hi ha una relació molt estreta, r^2 serà proper a 1, això vol dir que si es coneix el valor d'una variable, es pot predir la segona variable. En canvi, quan r^2 pren un valor baix, tot i existir una relació significativa entre les variables, el fet de conèixer una variable no permet predir amb molta precisió la segona variable. La direcció de la relació s'indica amb el signe del coeficient; un signe (+) indica una relació positiva i un signe (-) indica una relació negativa.
3. *Obtenir una equació que descriu la relació i es pugui utilitzar per a predir valors desconeguts.* Aquesta correlació es representa mitjançant un dels gràfics més comuns en la ciència: una variable a l'eix X (abscisses) enfront d'una altra a l'eix Y (ordenades). L'objectiu és determinar l'equació d'una línia que travessa el núvol de punts. L'equació d'una línia es dona en la forma:

$$\hat{Y} = a + bX \quad (3.26)$$

on \hat{Y} és el valor de Y predit per un valor determinat de X, a és la intercepció Y (el valor de Y quan X és zero), i b és el pendent de la recta (el canvi en \hat{Y} per un canvi en X d'una unitat).

Normalment, a estadística es mesuren quatre tipus de correlacions: la correlació de Pearson, la correlació de Kendall, la correlació de Spearman i la correlació punt biserial (97).

- *Correlació de Pearson:* és la més àmpliament utilitzada per mesurar el grau de relació entre variables linealment relacionades. Es basa en el càlcul del coeficient de correlació r i un dels requisits és que ambdues variables s'han de distribuir normalment (les variables distribuïdes normalment tenen una corba en forma de campana) i presentar homoscedasticitat.
- *Correlació de Spearman:* el coeficient de correlació de Spearman, ρ (rho), és una mesura de la correlació (l'associació o interdependència) entre dues variables aleatòries (tant contínues com discretes). S'utilitza quan les variables no segueixen una distribució normal. Per calcular ρ , les

dades són ordenades i canviades pel seu respectiu número d'ordre. La interpretació del coeficient de Spearman és igual que la del coeficient de correlació de Pearson. Oscil·la entre -1 i +1, indicant associacions negatives o positives respectivament. El zero significa no correlació.

- *Correlació de Kendall*: la tau de Kendall és un coeficient de correlació per rangs, inversions entre dues ordenacions d'una distribució normal bivariant.
- *Correlació punt biserial*: indica la relació entre una variable quantitativa (X) i una variable dicotòmica (Y).

Com hem descrit doncs, la correlació de Spearman s'utilitza com una alternativa no paramètrica a la regressió lineal i la correlació quan es tenen una o dues variables que no es distribueixen normalment. Tot i així, nombrosos estudis de simulació han demostrat que la regressió lineal i la correlació no són sensibles a la no normalitat: una o ambdues variables de mesura poden ser no normals i la probabilitat d'un fals positiu ($p < 0,05$, quan la hipòtesi nul·la és vertadera) encara és de 0,05 (98).

En relació a la Tesi Doctoral, al Capítol 4 s'utilitza la correlació de Pearson per veure l'associació entre el descriptor de la relació nucli/citoplasma i diferents descriptors de color i textura. D'altra banda, als Capítols 6 i 7 també s'utilitza per a veure si els descriptors obtinguts mitjançant els dos mètodes (sistema DM96 i BX43) són lineals per a cada tipus de cèl·lula limfoide estudiada així com per trobar el punt de tall de certs descriptors amb el sistema BX43 que han estat prèviament obtinguts amb el sistema DM96.

3.6.9 Avaluació de l'eficiència diagnòstica d'un test

Molts dels estudis diagnòstics que realitzem al laboratori són estudis quantitius en els quals volem establir un punt de tall. En funció on fixem aquest punt de tall, tindrem més falsos positius (FP) o més falsos negatius (FN). En tot estudi és important conèixer els falsos negatius i els falsos positius per interpretar correctament el resultat.

De tota prova diagnòstica, s'estudien les següents característiques estadístiques:

- *Sensibilitat*: probabilitat que el test sigui positiu quan l'individu està malalt, és a dir, aquells individus en què el test dóna positiu és perquè realment són individus que presenten la malaltia.

$$\text{Sensibilitat} = \frac{VP}{VP+FN} \quad (3.27)$$

És el quocient entre els veritables positius (VP) dividit per la suma d'aquests i els que són negatius per al test, però realment estan malalts (FN).

- *Especificitat*: probabilitat que el test sigui negatiu quan és un individu sa, és a dir, aquells individus en què el test dona negatiu és perquè realment són individus sans.

$$Especificitat = \frac{VN}{VN+FP} \quad (3.28)$$

És el quocient entre els veritables negatius (VN) dividit per la suma d'aquests i els que són sans amb test positiu (FP).

- Valor predictiu positiu (VPP): probabilitat de tenir la malaltia tenint un test positiu.

$$VPP = \frac{VP}{VP+FP} \quad (3.29)$$

- Valor predictiu negatiu (VPN): probabilitat de no tenir la malaltia (estar sa) tenint un test negatiu.

$$VPN = \frac{VN}{VN+FN} \quad (3.30)$$

Per a finalitzar l'avaluació de la prova diagnòstica és molt comú, en aquest tipus d'estudis, la realització de la corba ROC, una gràfica que mostra totes les parelles sensibilitat / especificitat resultants de la variació contínua dels punts de tall en tot el rang de resultats observats.

El nom prové de l'anglès "*Receiver Operating Characteristic*". Aquestes corbes relacionen l'especificitat i la sensibilitat. Aquesta última es representa a l'eix de la "Y" i en l'eix de la "X" es representa 1 menys l'especificitat. El que ens interessa és l'àrea que se situa sota la corba, que pot prendre valors entre 0 i 1. Si l'àrea sota la corba és 1, el test discrimina perfectament, i la sensibilitat i especificitat tindran valors d'1 (o 100%). Si el valor és de 0,5, no hi ha discriminació o, el que és el mateix, la probabilitat de tenir un test positiu estant malalt és la mateixa de tenir el test positiu estant sa. A la Figura 3.25 es mostra una corba ROC. Ja hem comentat que en l'eix d'abscisses es representa 1 menys l'especificitat. No obstant això en la corba només figura l'especificitat, però cal fixar-se que l'eix va de 100 a 0.

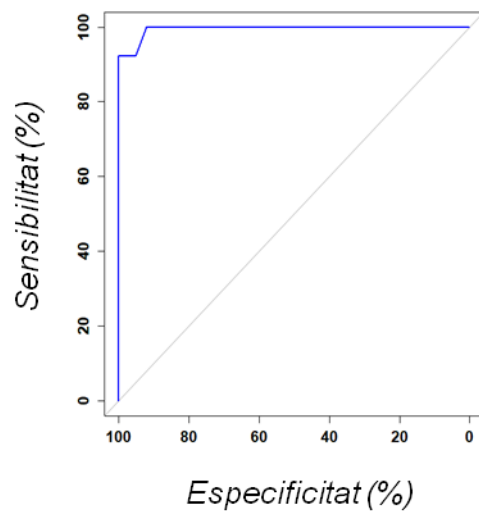


Figura 3.25: Corba ROC.

En relació a la present recerca, les corbes ROC s'han utilitzat a la segona part dels Capítols 5 i 7, quan s'ha avaluat el rendiment de descriptors quantitius específics d'un subtipus cel·lular limfoide utilitzant el paquet R 'pROC' (99). En primer lloc, les corbes ROC han permès obtenir un *cut-off* o valor de tall per a cada descriptor (si el valor és superior o inferior al valor de tall, la cèl·lula de la imatge mostra o no la característica citològica corresponent). En segon lloc, les corbes ROC també han permès obtenir un valor llindar per decidir quan un frotis de sang perifèrica és considerat positiu o negatiu per la característica citològica a estudi.

Capítol 4

4 Descriptors quantitius per a la diferenciació morfològica de cèl·lules limfoides anormals a partir d'imatges de sang perifèrica adquirides amb el sistema DM96

Basat en: Puigví L, Merino A, Alférez S *et al.* *New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood.* J Clin Pathol. 2017; 70; 1038-1048. doi: 10.1136/jclinpath-2017-204389.

Resum

L'objectiu d'aquest capítol és analitzar i realitzar una interpretació citològica d'un conjunt de descriptors quantitatius obtinguts per un procediment de *machine learning* per a la diferenciació morfològica de 12 grups de cèl·lules limfoides incloent cèl·lules limfoides reactives, anormals i blàstiques de sang perifèrica. Pel que fa a les cèl·lules limfoides anormals incloses, aquestes provenen de nou neoplàsies limfoides: leucèmia limfàtica crònica, leucèmia prolimfocítica B, tricoleucèmia, limfoma de la zona marginal esplènica, limfoma de cèl·lules del mantell, limfoma fol·licular, leucèmia prolimfocítica T, leucèmia de limfòcits grans granulars T i la síndrome de Sézary. S'han utilitzat 12.574 imatges adquirides mitjançant el sistema DM96 de cèl·lules procedents de 325 pacients i de cada una d'elles s'han calculat i analitzat un total de 2.676 descriptors: 27 geomètrics i 2.649 de color i textura. Posteriorment, s'han analitzat els 20 descriptors més rellevants per a la diferenciació morfològica dels 12 grups de cèl·lules limfoides objecte d'estudi. La majoria de descriptors presenten diferències significatives: 19/20 si comparem el limfoma fol·licular amb el de les cèl·lules del mantell, 18/20 per cèl·lules blàstiques i limfoides reactives, 17/20 per les cèl·lules de Sézary i els prolimfòcits T i 16/20 si comparem els prolimfòcits B amb el T i entre els limfòcits vellosos. Per tant, s'han establert noves variables citològiques basades en descriptors geomètrics i de color i textura que han contribuït a una valoració morfològica més precisa. L'associació d'aquestes variables amb altres mètodes de citometria de flux pot resultar interessant a explorar en el futur immediat.

4.1 Introducció

L'observació al microscopi del frotis de sang perifèrica és necessària i clínicament rellevant en un gran nombre d'ocasions i per diverses raons (39,100–102). Tot i així, està sotmesa a variacions inter- i intra-observadors (103) perquè a l'actualitat no existeixen valors objectius per a definir variables citològiques (2).

El frotis de sang perifèrica és essencial per a diferenciar entre les cèl·lules limfoides reactives (CLR), els limfòcits anormals (LA) i les cèl·lules blàstiques perquè comparteixen algunes característiques morfològiques subtils. D'acord amb les recents recomanacions per a l'estandardització de la nomenclatura de l'*International Council of Standardization in Hematology*, en el text es fa servir el terme “anormal” per incloure a les cèl·lules malignes de les neoplàsies limfoides (37).

Els LA són les cèl·lules més difícils d'identificar pels participants dels programes externs de garantia de qualitat d'Hematologia (4). La interpretació morfològica dels limfòcits de sang perifèrica és desafiant i depèn de l'experiència del citòleg així com de la informació clínica disponible (5,104).

Les cèl·lules limfoides de la leucèmia limfàtica crònica (LLC) i el limfoma fol·licular (LF) són sovint petites, mentre que en el limfoma de cèl·lules del mantell (LCM) són pleomòrfiques ja que varien tant en la mida com en la relació nucli/citoplasma (N/C) i algunes cèl·lules poden semblar blàstiques amb un nuclèol prominent (variant blàstica). A la tricoleucèmia (HCL), els tricoleucòcits són més grans que els limfòcits normals (LI) i tenen un abundant citoplasma blau gris amb vellositats citoplasmàtiques al voltant.

Les cèl·lules del limfoma de la zona marginal esplènica (LZME) contenen característicament prolongacions velloses en disposició polar. En el cas dels prolimfòcits B, es tracta de cèl·lules que tenen una mida de gairebé el doble de la d'un limfòcit normal, tenen una cromatina moderadament condensada i un nuclèol central poc aparent així com un citoplasma escàs i dèbilment basòfil. Els nuclis dels prolimfòcits T tenen un contorn irregular. Els limfòcits grans granulars (LGG) presenten un citoplasma abundant que conté una granulació fina o marcada azuròfila. Les cèl·lules de la síndrome de Sézary (SS) poden ser de mida petita (variant Lutzner) o gran (variant clàssica) però la morfologia de la cromatina nuclear, clàssicament descrita com cerebriforme, és el tret diferencial d'aquest grup de cèl·lules (38).

Durant els anys 90, es van realitzar estudis encaminats a quantificar paràmetres citològics en talls histològics (12,105,106), mentre que a la següent dècada, els estudis en sang perifèrica van prendre el relleu (2,10,13,107,108).

Benattar *et al.* (13,107) va realitzar una anàlisi morfomètrica i colorimètrica dels limfòcits anormals de sang perifèrica en neoplàsies limfoides B i van crear un sistema de puntuació per a caracteritzar les diferents malalties mitjançant la combinació de set criteris morfològics rellevants: la forma del nucli, la forma i l'àrea de la cèl·lula sencera, la relació N/C, el rati vermell/blau del nucli, el rati verd/blau del citoplasma i la

proporció de cèl·lules amb nuclèol. Aquest estudi també inclou neoplàsies limfoides T (la SS, la leucèmia prolimfocítica T (LPT) i la leucèmia/limfoma de cèl·lules T de l'adult) però, donada la baixa quantitat de mostres (n=10), no van poder treure conclusions. Pel que fa a les neoplàsies limfoides B, aquest estudi ressalta que la introducció de mesures colorimètriques pot ser d'ajuda a l'hora de quantificar altres característiques citològiques rellevants com la densitat de la cromatina, però que és imprescindible realitzar-ho en condicions estandarditzades.

Angulo *et al.* (10,108) utilitza eines de morfologia matemàtica per a processar imatges de color de sang perifèrica, extreu mesures quantitatives de mida, forma, color i textura (mitjançant la mida de partícules) i analitza irregularitats específiques del nucli i del citoplasma. L'autor construeix un marc ontològic per a descriure de manera automàtica els limfòcits anormals de la LLC, HCL, LZME, LF, LCM i LPT.

Jahanmehr *et al.* (2) quantifica paràmetres citològics de limfòcits malignes de la LLC, el LCM i la leucèmia prolimfocítica B (LPB) mitjançant l'anàlisi computacional d'imatges i un conjunt de paràmetres són calculats per a la quantificació de l'àrea i el diàmetre cel·lular, l'àrea citoplasmàtica, la relació N/C i l'àrea i densitat nuclear. Els resultats demostren que les tres entitats, LLC, LCM i LPB, poden ser diferenciades mitjançant paràmetres quantitius.

Aquest capítol presenta una metodologia construïda a partir de 20 noves variables citològiques (geomètriques i de color i textura) amb les següents propietats:

- 1) tenen formulacions quantitatives
- 2) permeten interpretacions morfològiques qualitatives
- 3) són eficients per a discernir entre un nombre considerable de grups de cèl·lules limfoides diferents, sobretot si considerem parelles de cèl·lules limfoides anormals en què el seu reconeixement a partir de la morfologia de sang perifèrica pot ésser difícil.

4.2 Material i mètodes

A la Figura 4.1 es mostra l'esquema de les etapes de les que consta el present Capítol 4 i que són les següents: 1) obtenció dels pacients i les mostres, 2) adquisició de les imatges digitals, 3) processament digital d'imatges i 4) anàlisi estadística. A continuació, s'explica cada una d'elles.

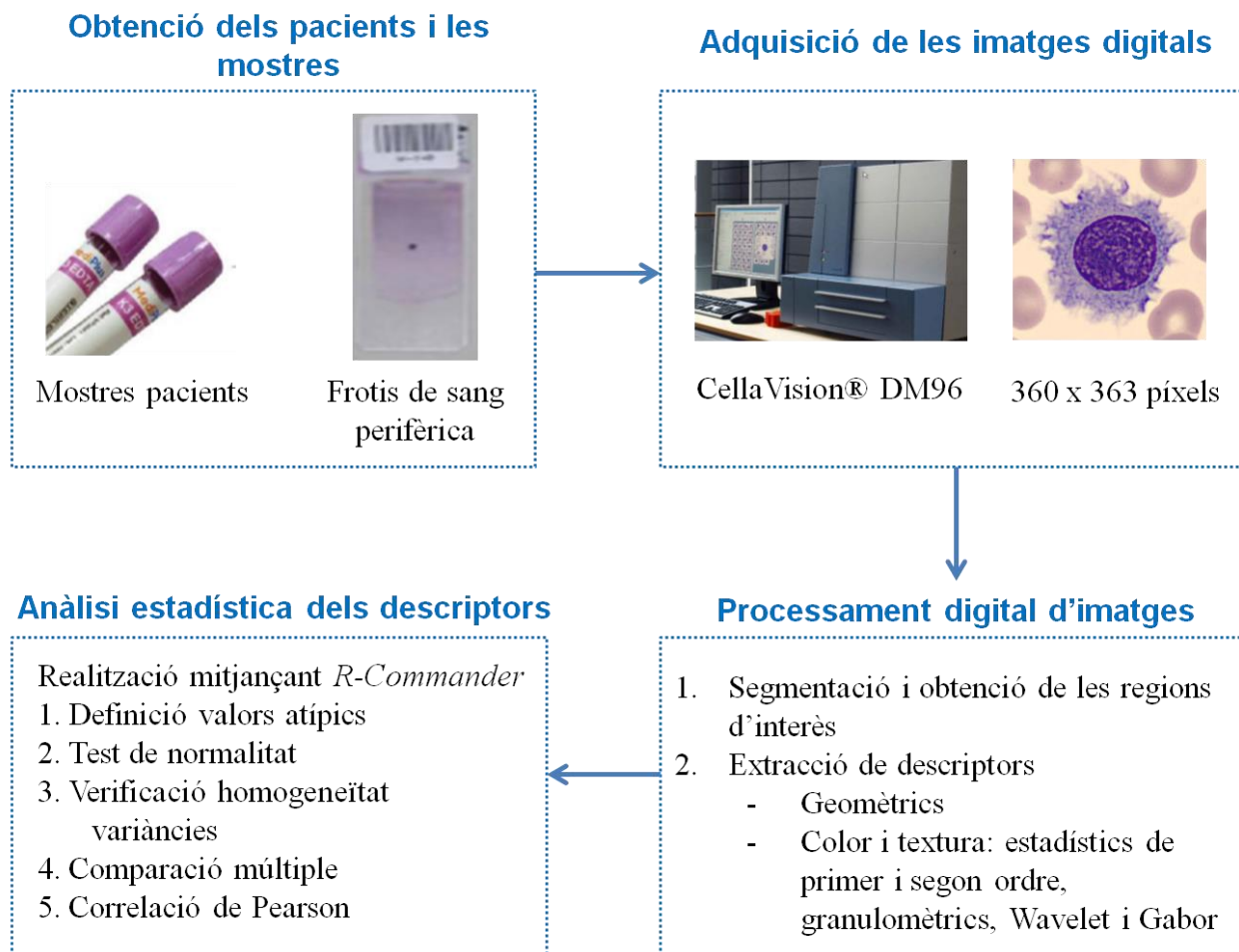


Figura 4.1: Esquema del Capítol 4.

4.2.1 Obtenció dels pacients i les mostres

En aquest estudi s'han inclòs limfòcits normals; limfòcits anormals de pacients amb LLC, LPB, HCL, LZME, LCM, LF, LPT, LLGG-T i SS així com cèl·lules blàstiques limfoides (BL) de pacients amb leucèmies limfoblàstiques agudes i limfòcits reactius (CLR) de pacients amb infeccions virals o d'altres etiologies (Fig. 4.2). Els diagnòstics malignes han estat confirmats mitjançant troballes clíniques i morfològiques així com l'immunofenotip de les cèl·lules limfoides analitzades per citometria de flux i altres proves complementàries seguint la classificació de l'OMS de les neoplàsies limfoides (38). Un total de 325 pacients han estat inclosos i a la Taula 4.1 se'n mostra la distribució.

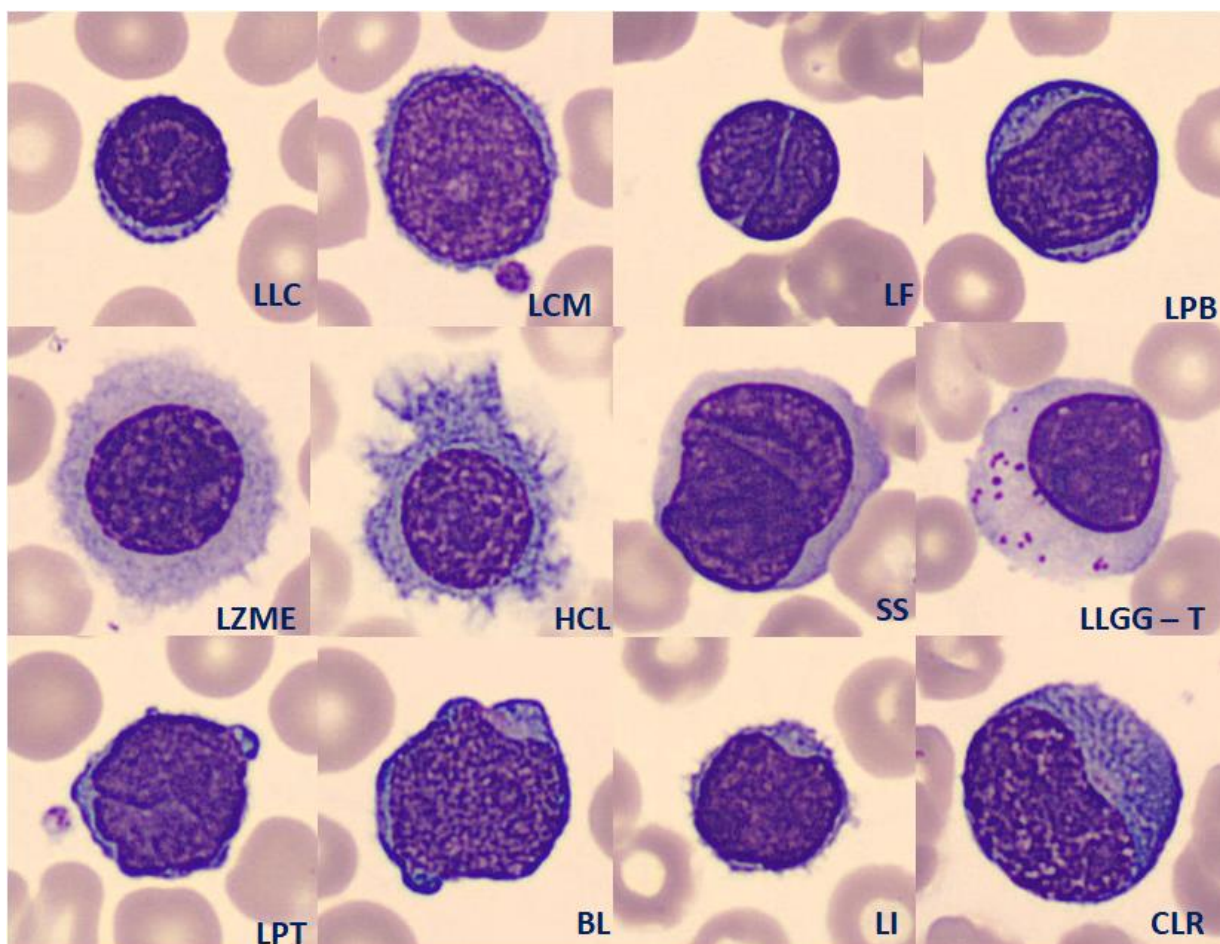


Figura 4.2: Exemples de cada un dels tipus de cèl·lules limfoides incloses a l'estudi. Imatges obtingudes amb el CellaVision DM96 a x1.000 augments.

Taula 4.1: Distribució dels diferents grups de cèl·lules limfoides, el nombre de pacients i imatges incloses a l'estudi.

Grup cel·lular limfoide	Pacients, n	Imatges, n
Limfòcits (LI)	49	1.085
Leucèmia limfàtica crònica (LLC)	56	2.468
Leucèmia prolimfocítica B (LPB)	3	292
Tricoleucèmia (HCL)	18	611
Limfoma de la zona marginal esplènica (LZME)	15	819
Leucèmia de cèl·lules del mantell (LCM)	21	1.438
Limfoma fol·licular (LF)	18	1.119
Leucèmia prolimfocítica T (LPT)	4	332
Leucèmia de limfòcits grans granulars (LLGG)	23	894
Síndrome de Sézary (SS)	11	756
Cèl·lules limfoides reactives (CLR)	82	1.199
Blasts limfoides (BL)	25	1.561
TOTAL	325	12.574

Pel que fa a l'immunofenotip de les cèl·lules incloses a estudi, les de la LLC expressaven dèbilment IgM/IgD de superfície, CD5, CD19, CD23, CD25 i CD20. CD10 i FMC7 van ser negatius. Els casos de LLC atípica i els centrocits que es troben freqüentment a la LLC no s'han inclòs en aquest estudi. Els prolimfòcits B expressaven fortament antígens cel·lulars de tipus B (CD19, CD20, CD22, CD79b), el FMC7 era negatiu i el CD5 i CD23 van ser positius en tots els pacients. Els tricoleucòcits van ser positius per CD11c, CD25, FMC7, CD103, CD123 i Annexina A1 i negatius per CD10 i CD5.

Les cèl·lules tumorals del LZME expressaven IgM de superfície, CD19, CD20 i FMC7, mentre que CD5, CD10 i CD43 van ser negatius. Les cèl·lules limfoides del LCM van ser positives per CD20, CD5, FMC7 i CD43, però negatives per CD10 i negatives o dèbilment positives per CD23. La translocació t(11;14)(q31;q32) entre els gens IGH i la ciclina D1 hi era present. Els limfòcits anormals del LF van ser positius per CD19, CD20, CD22, CD79a i CD10, mentre que el CD5 i CD43 van ser negatius i la translocació t(14;18)(q32;q21) i el reordenament del gen *BCL2* hi eren presents.

Els prolimfòcits T de la LPT van ser negatius per TdT i CD1a mentre que el CD2, CD3 i CD7 van ser positius. CD4+ i CD8- es va trobar en dos pacients i en els altres dos pacients, CD4 i CD8 van trobar-se coexpressats. Pel que fa als limfòcits grans granulars de la LLGG-T, aquests van expressar CD2, CD3 i CD8. En el 91 % dels casos es va observar una disminució o pèrdua d'expressió del CD5 i/o CD7 mentre que el CD57 es va trobar expressat per sobre del 70 %. Les cèl·lules de Sézary van ser positives per CD2, CD3, CD4 i CD5, mentre que característicament no van expressar CD7, CD8 ni CD26.

Els frotis de sang perifèrica es van obtenir dels pacients que formen part del flux de treball diari del Laboratori Core de l'Hospital Clínic de Barcelona. Es van realitzar a partir de sang venosa recollida en EDTA i utilitzant l'extensor-tenyidor automàtic SP1000i (Sysmex, Kobe, Japó) i seguint el procediment de tinció de May Grünwald-Giemsa durant les quatre primeres hores des de l'extracció sanguínia.

4.2.2 Adquisició de les imatges digitals

La qualitat dels frotis de sang perifèrica així com la morfologia cel·lular de les imatges van ser valorades per les facultatius especialistes (LP, AM) abans de realitzar l'estudi de les imatges. Cada imatge cel·lular tenia una resolució de 363 x 360 píxels i va ser obtinguda mitjançant el sistema CellaVision DM96 (CellaVision AB, Lund, Suècia). Un total de 12.574 imatges de cèl·lules limfoides van ser recollides i distribuïdes com es mostra a la Taula 4.1. La resta de processos desenvolupats sobre les imatges es van realitzar i implementar independentment del DM96, a una estació de treball Dell Precision 5810 amb processador Intel® Xeon® E5-1620 v3 4 x 3.50 GHz, 16 GB de RAM, una targeta gràfica NVIDIA Quadro K2200 i MATLAB® com a software científic.

4.2.3 Processament digital de les imatges

El primer pas del processament digital de les imatges va consistir en separar les regions d'interès (RI) de la cèl·lula limfoide d'altres objectes de la imatge, com els glòbuls vermells o les plaquetes i es va assolir mitjançant un procés automàtic de segmentació. Aquest procediment és crític dins el marc del processament digital d'imatges ja que una segmentació robusta és indispensable per a l'obtenció de descriptors fiables (75). En treballs previs del nostre grup de recerca (19–21), s'ha desenvolupat un procediment automàtic de segmentació basat en *color clustering* i morfologia matemàtica per tal de separar tres RI principals: el nucli, la cèl·lula sencera i la zona perifèrica al voltant de la cèl·lula. Una quarta RI s'obté pel citoplasma calculant la diferència entre la regió de la cèl·lula i del nucli. Alguns detalls respecte a la segmentació es descriuen al Capítol 3 i secció 3.3. Un cop obtingudes les RI, és hora d'extreure els descriptors geomètrics, de color i de textura.

4.2.4 Extracció de descriptors

L'objectiu principal de l'extracció de descriptors és obtenir mesures quantitatives dels objectes de la imatge. En el treball que es presenta en aquest capítol, s'han extret 2.676 descriptors: 27 geomètrics i 2.649 de color i textura. La Taula 4.2 mostra els diferents tipus de descriptors utilitzats en aquest treball. A continuació se'n presenta un resum.

4.2.4.1 *Descriptors geomètrics*

Aquest tipus de descriptors són els més intuïtius i estan estretament relacionats amb els patrons visuals que són familiars per als citòlegs. Des del punt de vista geomètric, hi ha dues RI principals: la cèl·lula sencera i el nucli. En aquest treball s'han calculat un total de 27 descriptors geomètrics.

Com mostra la Taula 4.2, per ambdues RI (nucli i cèl·lula) s'han calculat 12 descriptors geomètrics com el perímetre nuclear, l'àrea de la cèl·lula i la circularitat. Els tres restants descriptors impliquen dues RI per al seu càlcul com són la relació N/C, l'excentricitat del nucli i el *hairiness*. La relació N/C s'obté de dividir l'àrea nuclear de l'àrea del citoplasma i aquesta última s'obté restant l'àrea nuclear de la cel·lular. Pel que fa a l'excentricitat del nucli, aquesta s'ha calculat com la distància entre el centre de la cèl·lula i el centre del nucli (10). Per últim, el descriptor del perfil citoplasmàtic o *hairiness* s'ha calculat tal com el va descriure Alférez *et al.* (29) com a un nou mètode per a caracteritzar el perfil citoplasmàtic. Per a una interpretació més detallada dels descriptors geomètrics veure el Capítol 3 i secció 3.4.

4.2.4.2 *Descriptors de color i textura*

La idea d'utilitzar els descriptors de color i textura prové de la capacitat humana innata de reconèixer les diferències texturals utilitzant la visió i el tacte. La textura d'una imatge digital es defineix per la seva uniformitat, densitat, gruix, rugositat, regularitat, intensitat, direccionalitat, to del píxel i les seves relacions

espacials, entre d'altres, per identificar diferents patrons de textura (109). Diversos procediments es poden trobar a la literatura per analitzar la textura d'imatges digitals mèdiques (110).

Taula 4.2: Conjunt de 2.676 descriptors que s'han extret en la metodologia descrita en aquest capítol.

DESCRIPTORS GEOMÈTRICS (27)				
Cèl·lula	Nucli	Nucli i cèl·lula	Citoplasma	
	Àrea Diàmetre equivalent Perímetre Solidesa Extensió Circularitat Elongació Rodonesa <i>Convex hull</i> Convexitat Variància circular Variància de l'el·lipse	Excentricitat	<i>Hairiness</i>	
		Nucli i citoplasma		
		Relació N/C		
DESCRIPTORS DE COLOR I TEXTURA (2.649)				
Regió d'interès	Cèl·lula	Nucli	Citoplasma	Espais de color
Estadístics de primer ordre (6x3x19)		Mitjana Desviació estàndard Asimetria Curtosi Energia Entropia		RGB CMGN HSV Lab Luv XYZ
Estadístics de segon ordre (17x3x19)		Energia Entropia Contrast Homogeneïtat Correlació Variància <i>Sum average</i> <i>Sum variance</i> <i>Sum entropy</i> <i>Difference variance</i> <i>Difference entropy</i> <i>Information measure of correlation 1</i> <i>Information measure of correlation 2</i> Coeficient de correlació màxima Probabilitat màxima <i>Cluster shade</i> <i>Cluster prominence</i>		
Wavelet (12 x 3 x 19)		Mitjana d'H1, H2, V1, V2, D1 i D2 Desviació estàndard d'H1, H2, V1, V2, D1 i D2		
Granulometria (8x3x19)		Corba granulomètrica: mitjana, desviació estàndard, asimetria i curtosi Corba pseudo-granulomètrica: mitjana, desviació estàndard, asimetria i curtosi		
Gabor (22x3x3)		Mitjana a 8°, 12°, 15° Desviació estàndard a 8°, 12°, 15° Mitjana a 22°, 45°, 67°, 90°, 112°, 135°, 157° i 180° Desviació estàndard a 22°, 45°, 67°, 90°, 112°, 135°, 157° i 180°		

Des d'un punt de vista numèric, una imatge digital està formada per un rectangle de píxels. Cada píxel s'identifica per la seva localització específica i té un valor d'intensitat. Com a pràctica habitual, una imatge de color original es descompon en tres o quatre components de color, que defineixen un espai de color. Per exemple, el RGB és l'espai de color més utilitzat, que està format per tres components de color: vermell (R), verd (G) i blau (B). Altres espais de color utilitzats en processament digital d'imatges són el CMGN (cian, magenta, groc i negre) i l'HSV (tonalitat, saturació, valor), entre d'altres. Cada component de color es tradueix en una imatge individual en escala de grisos, la qual es compon exclusivament de tons grisos, que van des del negre (intensitat més dèbil) fins al blanc (intensitat més forta) (111).

L'histograma, tal i com hem explicat al Capítol 3 i secció 3.4.2.2, és una forma molt pràctica de descriure quantitativament una imatge i la base de moltes tècniques de processament com les desenvolupades en aquest treball (75). A mode d'exemple, la Figura 4.3 mostra (I) dos exemples d'imatges de cèl·lules de Sézary amb (II) les seves corresponents imatges en escala de grisos per al component de color magenta.

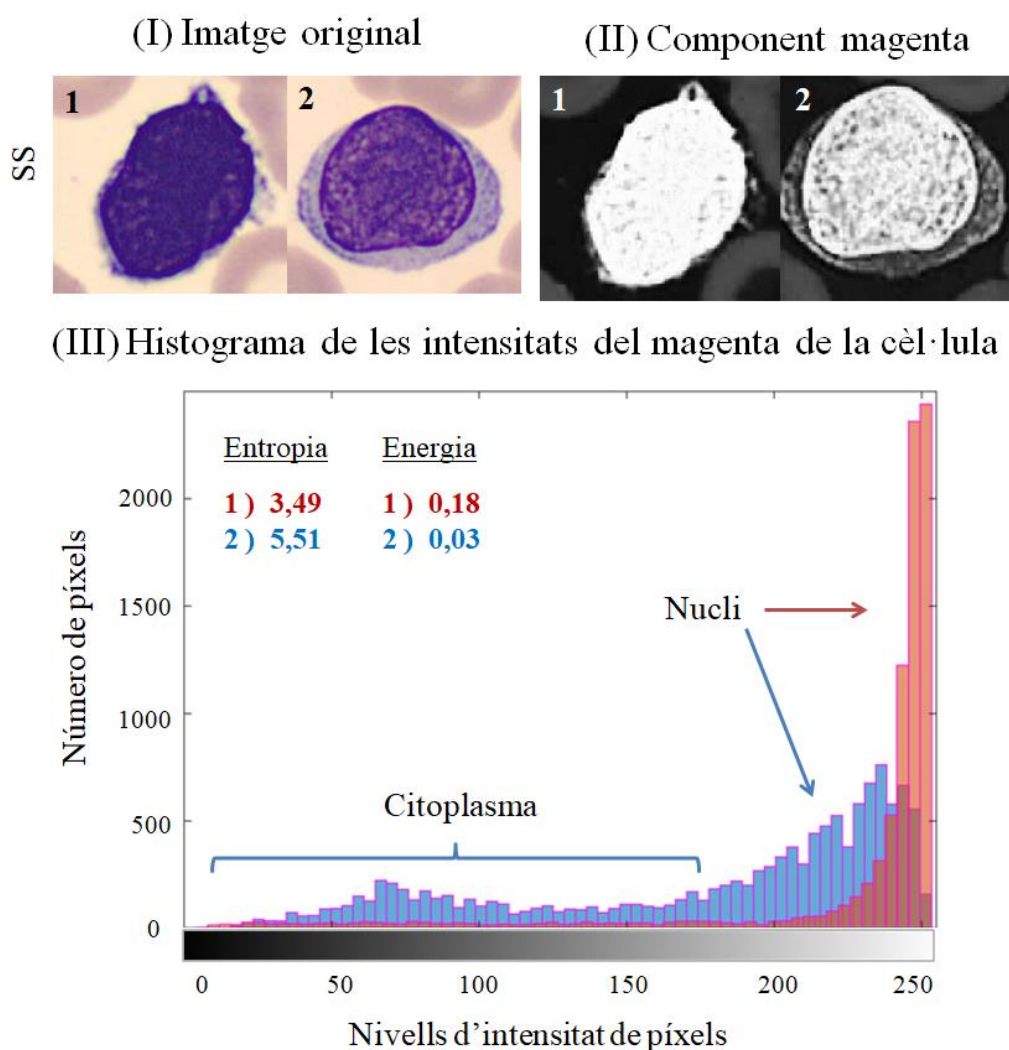


Figura 4.3 : Anàlisi d'imatges de l'entropia de l'histograma del component magenta de la cèl·lula. (I) Imatges originals de dues cèl·lules de la síndrome de Sézary (SS). (II) Component magenta en escala de grisos. (III) Histogrames de les intensitats de magenta de la cèl·lula.

L'histograma és un diagrama format per dos eixos on l'eix horitzontal representa la intensitat de grisos amb valors que van des de 0 (negre) fins a un valor màxim normalitzat (blanc), dividit en un nombre discret d'interval. L'eix vertical conté el número de píxels a la imatge amb les intensitats corresponents a cada interval.

En aquest estudi, per a identificar els patrons de textura, s'han utilitzat els següents quatre tipus de descriptors de color i textura (Taula 4.2): 1) estadístics (primer i segon ordre) (88,112), 2) *Wavelet* (91), 3) granulomètrics (29,113) i 4) Gabor (92,109,114).

Els descriptors estadístics de primer ordre, ja descrits al Capítol 3, són paràmetres calculats a partir de l'histograma d'una imatge digital en escala de grisos (115) i són els següents:

- (1) Mitjana: mitjana de tots els valors de la intensitat de la regió d'interès
- (2) Desviació estàndard (DE): dispersió de la intensitat dels píxels al voltant de la mitjana
- (3) Asimetria: simetria de l'histograma al voltant de la mitjana
- (4) Curtosi: quant de pla és l'histograma
- (5) Energia: uniformitat de l'histograma
- (6) Entropia: variabilitat de l'histograma de la imatge. Energia i entropia tenen una relació inversa (3).

Ambdós descriptors de textura s'exemplifiquen a la Figura 4.3.

Com es pot observar, l'histograma del component magenta corresponent a la imatge 2 (distribució blava) mostra un nombre significatiu de píxels en un rang ampli d'intensitat, el que indica l'existència d'una variabilitat significativa (entropia alta i energia baixa). En canvi, l'histograma de la imatge 1 (distribució vermella) presenta la majoria dels píxels en un interval petit d'elevada intensitat, que correspon a una distribució d'intensitat poc distribuïda (entropia baixa, energia elevada).

D'altra banda, els descriptors estadístics de segon ordre proporcionen més informació sobre la textura de la regió d'interès i es basen en la matriu de co-ocurrència del nivell de gris d'una imatge digital (88). Alguns d'aquests descriptors són l'homogeneïtat i el *cluster shade* (112).

La granulometria estima la distribució de la mida dels grànuls brillants i foscos de la imatge. Per a això s'utilitza morfologia matemàtica per a estimar la distribució dels grànuls indirectament, aplicant operacions morfològiques matemàtiques successives d'*opening* i *closing* (erosió i dilatació), respectivament (116). Unint les densitats de distribució en un gràfic, s'obté la corba granulomètrica, tal i com s'ha descrit al Capítol 3 i secció 3.4.3.3. D'aquesta corba s'han calculat la mitjana, la desviació estàndard, l'asimetria i la curtosi.

Pel que fa a les funcions Gabor, aquestes són filtres sensibles a l'orientació utilitzats per a la detecció de contorns (117). En resum, en aquest estudi s'han extret un total de 43 descriptors de color i textura diferents, els quals han estat calculats sobre sis espais de colors diferents, incloent 19 components de color (RGB, CMYK, XYZ, *Lab*, *Luv* i HSV) a excepció dels 22 descriptors Gabor, que només s'han calculat sobre el RGB. Encara que l'espai de color *Lab* és el més similar a la percepció humana, s'han utilitzat sis espais de color diferents perquè treballs anteriors han demostrat que això ofereix una gran quantitat d'informació sobre les cèl·lules limfoides (3,31).

4.2.5 Anàlisi estadística dels descriptors

Per tal de determinar els 20 descriptors més rellevants i menys redundants dels 2.676 descriptors obtinguts, s'ha aplicat la selecció de descriptors seguint la teoria de la informació mútua (118,119). Sobre els descriptors seleccionats, s'han realitzat les següents anàlisis estadístiques: definició dels valors atípics mitjançant el mètode de Tukey (120), comprovació de la normalitat amb el test de Kolgomorov-Smirnov, verificació de l'homogeneïtat de les variàncies mitjançant la prova de Fligner-Killeen, la prova Kruskal-Wallis per al càlcul dels *p* valors, la prova Kruskal-Wallis juntament amb el test de Dunn aplicant l'ajust de Bonferroni per a la comparació múltiple (121,122) i la correlació de Pearson.

D'altra banda, donada la importància del reconeixement de les cèl·lules limfoides reactives (benignes) per a diferenciar-les de les neoplàsiques (malignes), s'ha analitzat quins descriptors mostren diferències significatives entre les CLR i cada un dels subtipus de cèl·lula limfoide anormal.

Per a minimitzar la influència dels valors atípics, s'han pres els valors de les medianes (quartil 50) en comptes dels valors de la mitjana com a valors representatius per a la comparació múltiple de tots els descriptors. *P* valors inferiors a 0,05 s'han considerat estadísticament significatius. Totes les anàlisis estadístiques s'han realitzat mitjançant el software *R-Commander* (93).

4.3 Resultats

Els 20 descriptors més rellevants i menys redundants obtinguts per a la discriminació dels 12 tipus de cèl·lules limfoides incloses a l'estudi es mostren a la Taula 4.3. Tots ells han permès quantificar diferents característiques morfològiques i s'han obtingut diferències estadísticament significatives en els 20 descriptors ($p < 0,0001$). Cap dels 20 descriptors segueixen una distribució normal i tots ells mostren variàncies heterogènies.

Taula 4.3: Els 20 descriptors més rellevants seleccionats en base a maximitzar la rellevància i minimitzar la redundància. Per a cada descriptor s'inclou la posició, el component i l'espai de color en el cas dels descriptors de color i textura, la regió d'interès i la classe de descriptor.

Posició	Descriptor	Component de color	Espai de color	Regió d'interès	Classe
1	Relació nucli/citoplasma	-	-	cèl·lula i nucli	Geomètric
2	Perímetre	-	-	nucli	Geomètric
3	Diàmetre equivalent	-	-	cèl·lula	Geomètric
4	Curtosi de la corba pseudo-granulomètrica	Cian	CMGN	cèl·lula	Granulomètric
5	Asimetria de l'histograma	Blau	RGB	cèl·lula	Estadístic primer ordre
6	<i>Information measure of correlation 1</i>	Magenta	CMGN	cèl·lula	Estadístic segon ordre
7	Desviació estàndard de l'histograma	Blau	RGB	cèl·lula	Estadístic primer ordre
8	Entropia de l'histograma	Magenta	CMGN	cèl·lula	Estadístic primer ordre
9	Asimetria de l'histograma	Verd	RGB	cèl·lula	Estadístic primer ordre
10	Mitjana de la corba pseudo-granulomètrica	Negre	CMGN	citoplasma	Granulomètric
11	Mitjana de l'histograma	Blau	RGB	nucli	Estadístic primer ordre
12	<i>Cluster shade</i>	Saturació	HSV	cèl·lula	Estadístic segon ordre
13	Mitjana de l'histograma	Magenta	CMGN	citoplasma	Estadístic primer ordre
14	Homogeneïtat	Magenta	CMGN	cèl·lula	Estadístic segon ordre
15	Curtosi de l'histograma	Saturació	HSV	cèl·lula	Estadístic primer ordre
16	<i>Information measure of correlation 1</i>	Cian	CMGN	nucli	Estadístic segon ordre
17	<i>Hairiness</i>	-	-	citoplasma	Geomètric
18	Desviació estàndard de l'histograma	Saturació	HSV	cèl·lula	Estadístic primer ordre
19	Circularitat	-	-	nucli	Geomètric
20	Homogeneïtat	Negre	CMGN	citoplasma	Estadístic segon ordre

Pel que fa a les regions d'interès involucrades, 11 dels 20 descriptors es calculen sobre la cèl·lula sencera, quatre sobre el citoplasma i quatre més sobre el nucli. En canvi, la relació N/C s'obté a partir de dues regions. En aquest treball, aquest descriptor s'ha calculat a partir del nucli i del citoplasma (Taula 4.3). La Taula 4.4 mostra la significança estadística obtinguda al comparar els valors quantitius dels 20 descriptors geomètrics entre sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil.

Taula 4.4: p valors obtinguts mitjançant el test de Kruskal-Wallis-Dunn i l'ajust de Bonferroni. Comparació múltiple dels valors quantitatius dels 20 descriptors entre sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil. Un p valor inferior a 0,05 s'ha considerat significatiu.

Posició	Descriptor	LI vs LLC	HCL vs LZME	LCM vs LF	LPB vs LPT	LPT vs SS	CLR vs BL
1	Relació nucli/citoplasma	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	0,001	< 2 ⁻¹⁶
2	Perímetre nuclear	NS	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	NS	< 0,00001
3	Diàmetre equivalent	NS	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	NS	< 0,00001
4	Curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula	0,02	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶
5	Asimetria de l'histograma del blau de la cèl·lula	< 0,00001	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶
6	<i>Information measure of correlation 1</i> del magenta de la cèl·lula	NS	< 0,00001	< 2 ⁻¹⁶	NS	< 2e-16	< 2 ⁻¹⁶
7	Desviació estàndard de l'histograma del blau de la cèl·lula	< 0,00001	< 0,00001	< 2 ⁻¹⁶	< 0,00001	< 2e-16	< 2 ⁻¹⁶
8	Entropia de l'histograma del magenta de la cèl·lula	< 0,00001	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2e-16	< 2 ⁻¹⁶
9	Asimetria de l'histograma del verd de la cèl·lula	NS	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶
10	Mitjana de la corba pseudo-granulomètrica del negre del citoplasma	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶
11	Mitjana de l'histograma del blau del nucli	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶
12	<i>Cluster shade</i> de la saturació de la cèl·lula	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶
13	Mitjana de l'histograma del magenta del citoplasma	< 2 ⁻¹⁶	0,0001	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶
14	Homogeneïtat del magenta de la cèl·lula	< 0,00001	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶
15	Curtosi de l'histograma de la saturació de la cèl·lula	NS	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
16	<i>Information measure of correlation 1</i> del cian del nucli	< 2 ⁻¹⁶	< 0,00001	0,001	< 0,00001	< 0,00001	< 2 ⁻¹⁶
17	<i>Hairiness</i>	< 2 ⁻¹⁶	0,0001	< 2 ⁻¹⁶	< 0,00001	< 0,00001	NS
18	Desviació estàndard de l'histograma de la saturació de la cèl·lula	< 0,00001	0,008	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
19	Circularitat nuclear	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	NS
20	Homogeneïtat del negre del citoplasma	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	0,005	< 0,00001	< 2 ⁻¹⁶

BL, cèl·lula blàstica limfoide; CLR, cèl·lula limfoide reactiva; HCL, tricoleucèmia; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LI, limfòcit normal; LLC, leucèmia limfàtica crònica; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; NS, no significació estadística; SS, síndrome de Sézary.

El resultat de la Taula 4.4 mostren que la majoria dels 20 descriptors han estat significatius: un total de 19 comparant els limfòcits anormals del LF i el LCM, 18 per LB i CLR, 17 per a prolimfòcits T respecte cèl·lules de Sézary i 16 per a limfòcits normals i limfòcits anormals de la LLC, entre els dos subconjunts de limfòcits vellosos (HCL i LZME) i entre prolimfòcits B i T.

La Taula 4.5 inclou els cinc descriptors de color i textura que mostren significació estadística quan es fan comparacions entre les CLR (benignes) amb cadascun dels nou subconjunts de limfòcits anormals (malignes).

Taula 4.5: p valors dels descriptors de color i textura que mostren significació estadística quan es comparen els seus valors de les cèl·lules limfoides reactives amb cadascun dels nous subconjunts de limfòcits anormals.

	DE del blau de la cèl·lula (7è)	Entropia del magenta de la cèl·lula (8è)	Mitjana del blau del nucli (11è)	IMC1 del cian del nucli (16è)	DE de la saturació de la cèl·lula (18è)
LLC	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
HCL	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	0,0001	< 0,00001
LZME	< 0,00001	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LCM	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	0,001	< 0,00001
LF	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶
LPB	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LPT	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	0,005	< 2 ⁻¹⁶
LLGG-T	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
SS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 0,00001

CLR, cèl·lula limfoide reactiva; DE, desviació estàndard; HCL, tricoleucèmia; IMC1, *Information measure of correlation I*; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LLC, leucèmia limfàtica crònica; LLGG-T, leucèmia de limfòcits grans granulars T; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

A continuació es presenten els resultats corresponents als cinc descriptors geomètrics amb més rellevància per a la discriminació entre els grups d'estudi a l'apartat 4.3.1. Seguidament, l'apartat 4.3.2 presenta els 15 descriptors de color i textura.

4.3.1 Descriptors geomètrics

Els tres descriptors més rellevants per a la discriminació dels 12 grups cel·lulars són geomètrics: (1) *relació N/C*, (2) *perímetre nuclear* i (3) *diàmetre equivalent cel·lular*. Per tant, la *relació N/C* ha resultat ser el millor descriptor per a la diferenciació dels 12 tipus de cèl·lules limfoides i els seus valors concorden amb la morfologia dels diferents grups de cèl·lules. Dos descriptors geomètrics addicionals es troben dins els primers 20 descriptors: el descriptor del perfil citoplasmàtic o *hairiness* (posició 17) i la *circularitat nuclear* (posició 19).

La Figura 4.4 mostra les imatges de cèl·lules individuals corresponents als valors més baixos i més alts de la *relació N/C* per a cada grup cel·lular, així com els corresponents diagrames de caixes.

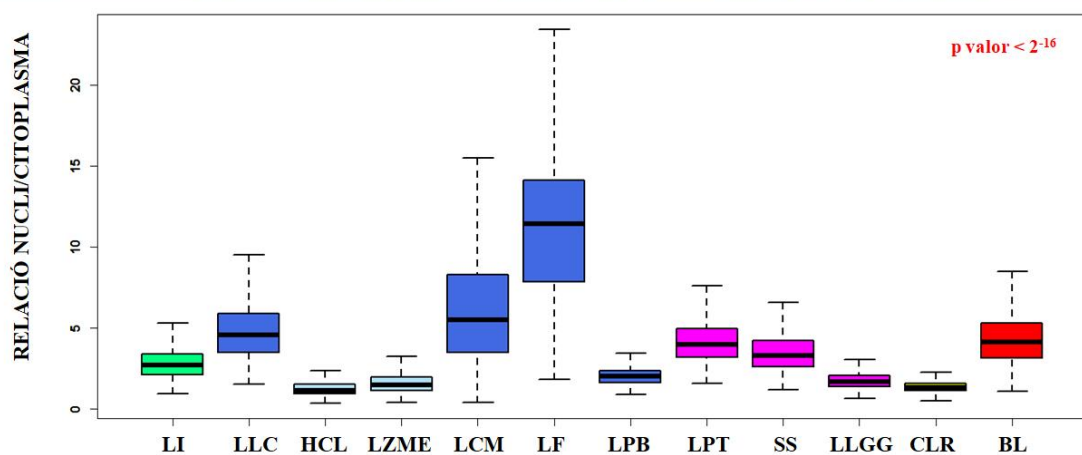
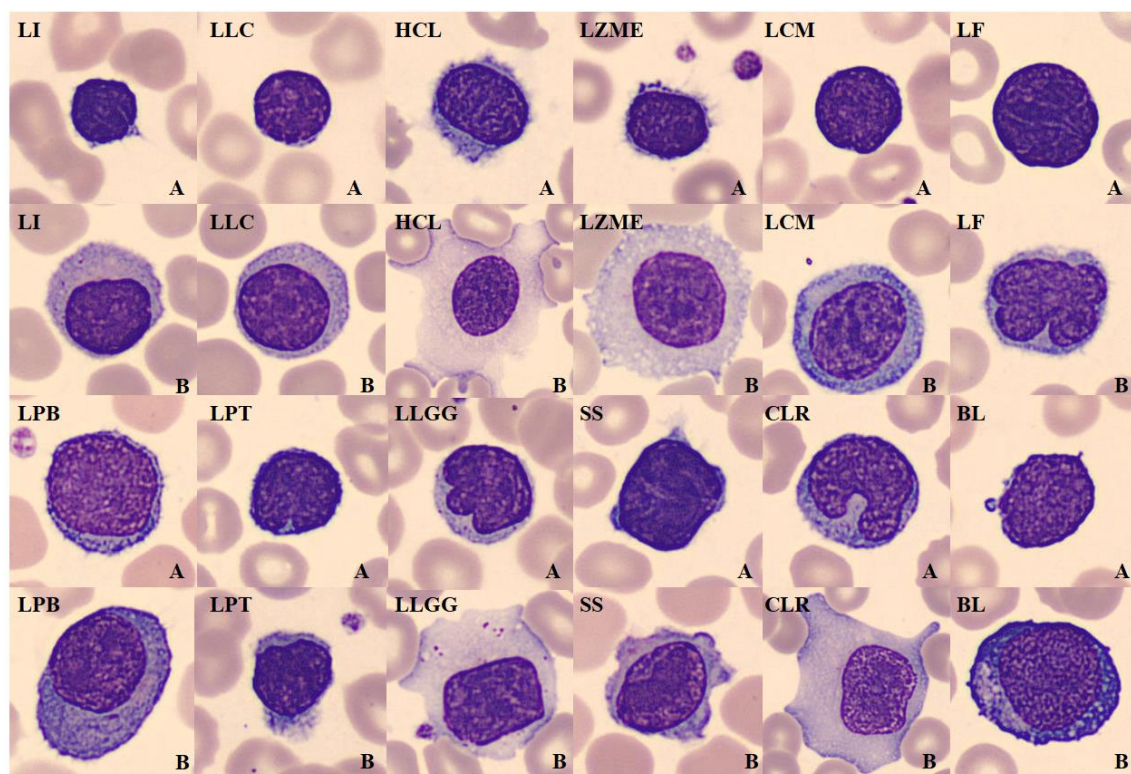


Figura 4.4: Exemples d'imatges de cèl·lules individuals que corresponen al valor més alt (A) i més baix (B) de la *relació nucli/citoplasma* i els corresponents diagrames de caixes.

La Taula 4.6 presenta els valors quantitatius dels cinc descriptors geomètrics més rellevants (els tres primers, el 17è i el 19è) per a diferenciar entre els 12 grups de cèl·lules limfoides.

Taula 4.6: Valors quantitatius dels descriptors geomètrics més rellevants. Els valors que es mostren són les medianes (rang interquartílic).

Posició	1	2	3	17	19
Descriptor	Relació nucli/citoplasma	Perímetre nuclear (píxels)	Diàmetre equivalent (píxels)	Hairiness (píxels)	Circularitat nuclear
LI	2,7 (2,1 - 3,4)	278 (264 - 293)	100 (92 - 107)	170 (102 - 268)	0,977 (0,958 - 0,991)
LLC	4,5 (3,5 - 5,9)	281 (265 - 298)	98 (92 - 105)	46 (23 - 89)	0,994 (0,984 - 1,002)
HCL	1,2 (0,9 - 1,5)	320 (305 - 335)	135 (125 - 150)	647 (345 - 1056)	0,993 (0,980 - 1,002)
LZME	1,5 (1,1 - 2,0)	323 (308 - 339)	130 (120 - 143)	518 (222 - 979)	0,982 (0,963 - 0,992)
LCM	5,5 (3,5 - 8,3)	338 (309 - 372)	116 (103 - 130)	81 (33 - 160)	0,954 (0,913 - 0,980)
LF	11,4 (7,8 - 14,1)	274 (253 - 298)	90 (84 - 97)	29 (10 - 71)	0,980 (0,941 - 1,002)
LPB	2,0 (1,7 - 2,4)	344 (328 - 368)	130 (122 - 141)	210 (93 - 338)	0,948 (0,919 - 0,970)
LPT	4,0 (3,2 - 5,0)	307 (290 - 327)	103 (97 - 108)	79 (49 - 133)	0,909 (0,868 - 0,938)
SS	3,3 (2,6 - 4,2)	308 (284 - 343)	109 (100 - 120)	138 (82 - 234)	0,963 (0,932 - 0,984)
LLGG	1,7 (1,4 - 2,1)	305 (292 - 321)	118 (112 - 129)	109 (46 - 189)	0,958 (0,931 - 0,979)
CLR	1,4 (1,2 - 1,6)	368 (348 - 400)	147 (138 - 157)	44 (21 - 83)	0,908 (0,852 - 0,944)
BL	4,1 (3,2 - 5,3)	357 (333 - 384)	119 (111 - 128)	51 (23 - 100)	0,913 (0,863 - 0,951)

BL, cèl·lula blàstica limfoide; CLR, cèl·lula limfoide reactiva; HCL, tricoleucèmia; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LI, limfòcit normal; LLC, leucèmia limfàtica crònica; LLGG-T, leucèmia de limfòcits grans granulars T; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

Pel que fa a la *relació N/C* (1a posició), les dues cèl·lules limfoides velloses (HCL i LZME) i les CLR són els subconjunts amb les medianes més baixes (1,2, 1,5 i 1,4, respectivament), mentre que el LF és el grup amb la mediana més elevada (11,4) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, la *relació N/C* ha presentat significança estadística a totes elles ($p < 0,0001$, Taula 4.4).

El segon descriptor més rellevant és el *perímetre nuclear* (2a posició), que es mesura en píxels, el LCM i la SS són els subconjunts cel·lulars que mostren més variabilitat en els seus valors: el rang interquartílic és de 309 – 372 i 284 – 343 píxels, respectivament (Fig. 4.5). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, el *perímetre nuclear* ha estat significatiu per a discriminar entre LCM *versus* LF, LPB *versus* LPT i CLR *versus* BL ($p < 0,0001$, Taula 4.4).

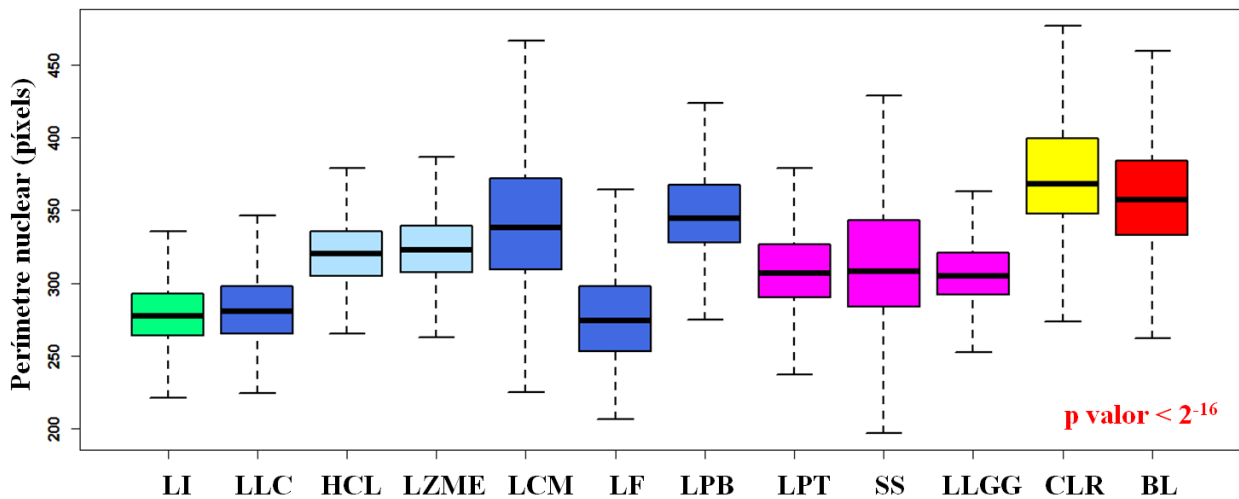


Figura 4.5: Diagrames de caixes obtinguts per al descriptor del *perímetre nuclear* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

En relació al *diàmetre equivalent* (3a posició), aquest descriptor ressalta les diferències de mida entre alguns subconjunts de limfòcits anormals. El grup de les CLR presenta les medianes més altes, 147 píxels, mentre que el LF les més baixes, 90 píxels ($p < 0,0001$, Fig. 4.6). La LPB presenta valors més alts en comparació amb la LPT, 130 i 103 píxels, respectivament ($p < 0,0001$, Fig. 4.6). Els resultats obtinguts per aquest descriptor geomètric en els diferents subconjunts cel·lulars limfoides concorden amb la seva morfologia (Fig. 4.6). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, el *diàmetre equivalent* ha presentat significança estadística en les mateixes tres parelles que el *perímetre nuclear* ($p < 0,0001$, Taula 4.4).

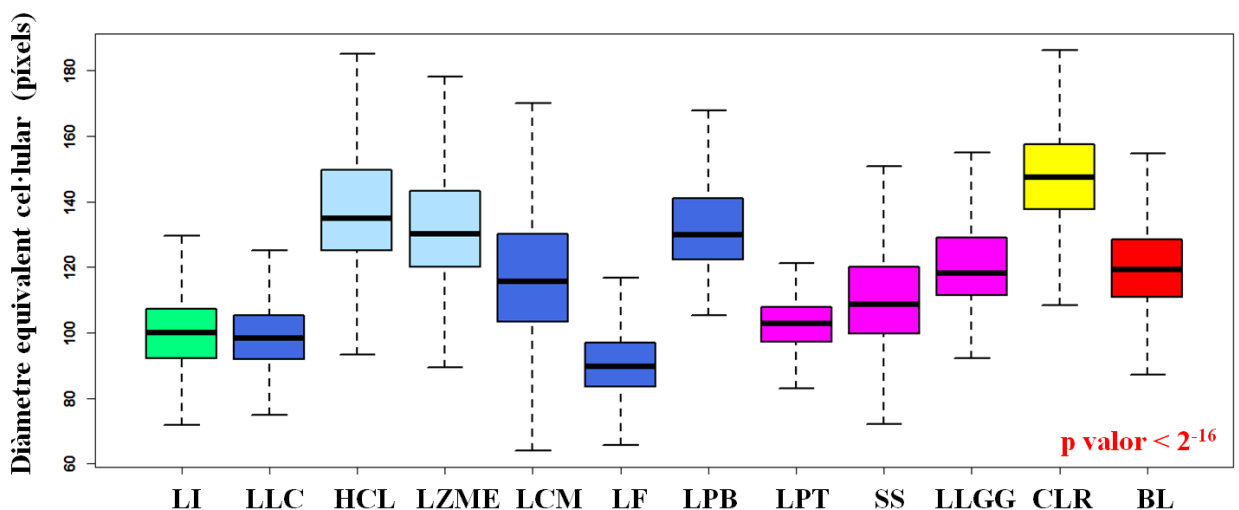


Figura 4.6: Diagrames de caixes obtinguts per al descriptor del *diàmetre equivalent cel·lular* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Pel que fa al *hairiness* (17a posició), la quantificació del perfil de la regió citoplasmàtica ha permès discriminar els limfòcits vellosos d'altres limfòcits anormals. L'HCL mostra medianes de 647 i el LZME de 518, mentre que els valors de la resta de grups es troben entre 29 i 210 píxels ($p < 0,0001$, Fig. 4.7). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, el *hairiness* ha estat significatiu per a discriminar entre totes elles a excepció del binomi CLR-BL ($p < 0,0001$, Taula 4.4).

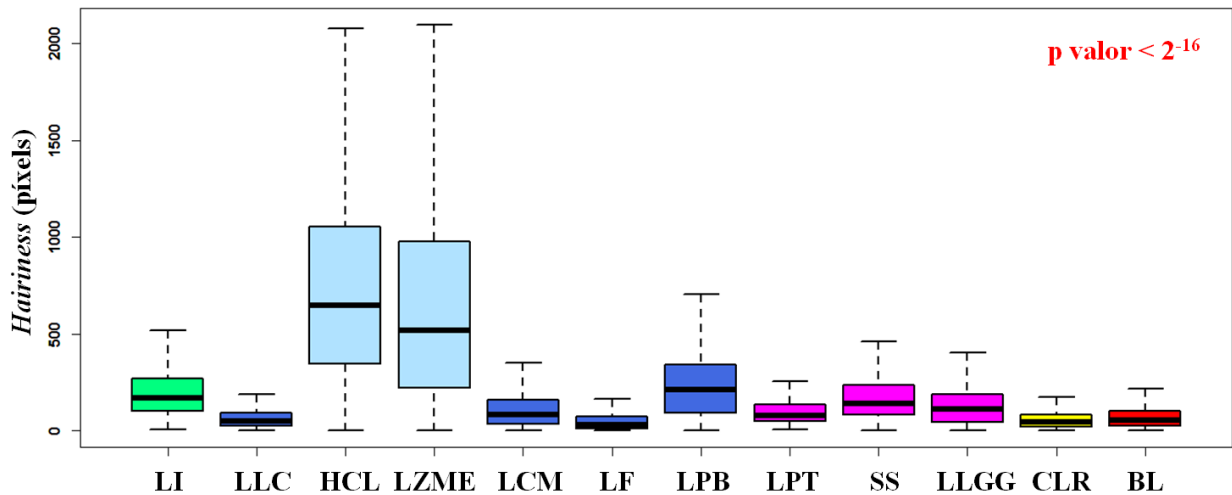


Figura 4.7: Diagrames de caixes obtinguts per al descriptor del *hairiness* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Per últim, la *circularitat nuclear* (19a posició), el valor del qual va des de zero (baixa circularitat nuclear) fins a un (nucli rodó), els limfòcits anormals de la LLC i HCL mostren els valors més alts (0,994 i 0,993, respectivament), mentre que les CLR i la LPT presenten els valors més baixos (0,908 i 0,909, respectivament, $p < 0,0001$, Fig. 4.8). En relació a la comparació múltiple, la *circularitat nuclear* ha presentat significança estadística en les mateixes tres parelles que el *hairiness* ($p < 0,0001$, Taula 4.4).

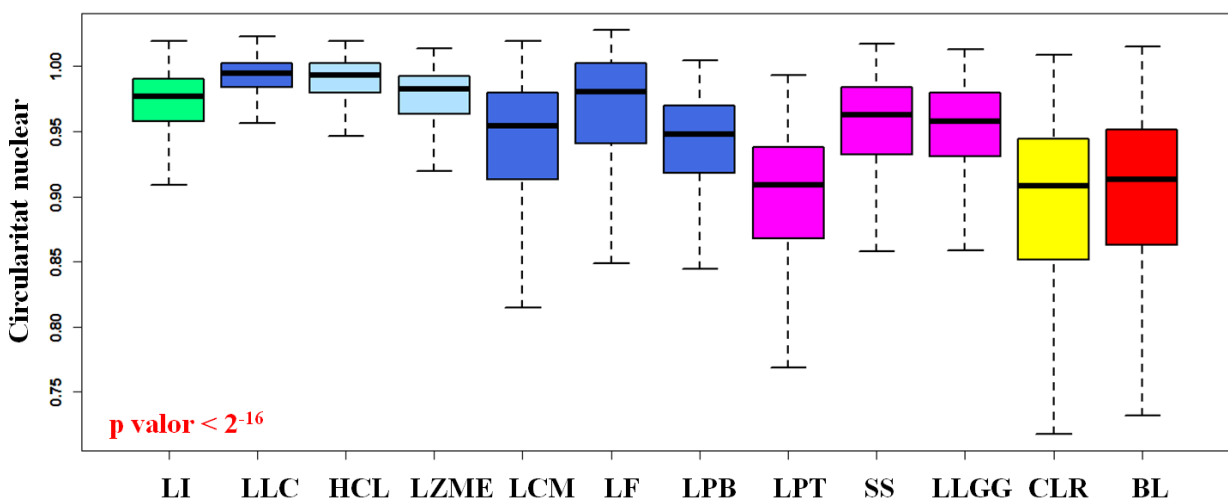


Figura 4.8: Diagrames de caixes obtinguts per al descriptor de la *circularitat nuclear* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

4.3.2 Descriptors de color i textura

Els 15 descriptors restants que van mostrar rellevància per a la discriminació entre els subconjunts de cèl·lules limfoides van ser de color i textura. A la Taula 4.2 es pot observar que només tres espais de color (CMGN, RGB i HSV) i només sis dels 19 components de color inicials es van veure involucrats en els primers 20 descriptors: el blau, el verd, el magenta, el negre, el cian i la saturació. A més a més, tenint en compte els quatre tipus de descriptors de color i textura utilitzats, 13 són estadístics (set de primer ordre i sis de segon ordre) i dos granulomètrics.

Pel que fa a la *curtosi de la corba pseudo-granulomètrica del component cian de la cèl·lula* (4t descriptor més rellevant), la Figura 4.9 té com a objectiu mostrar algunes nocions més de la granulometria amb la finalitat d'entendre la seva relació amb les característiques discriminants morfològiques. La resta de descriptors basats en corbes granulomètriques tenen interpretacions similars.

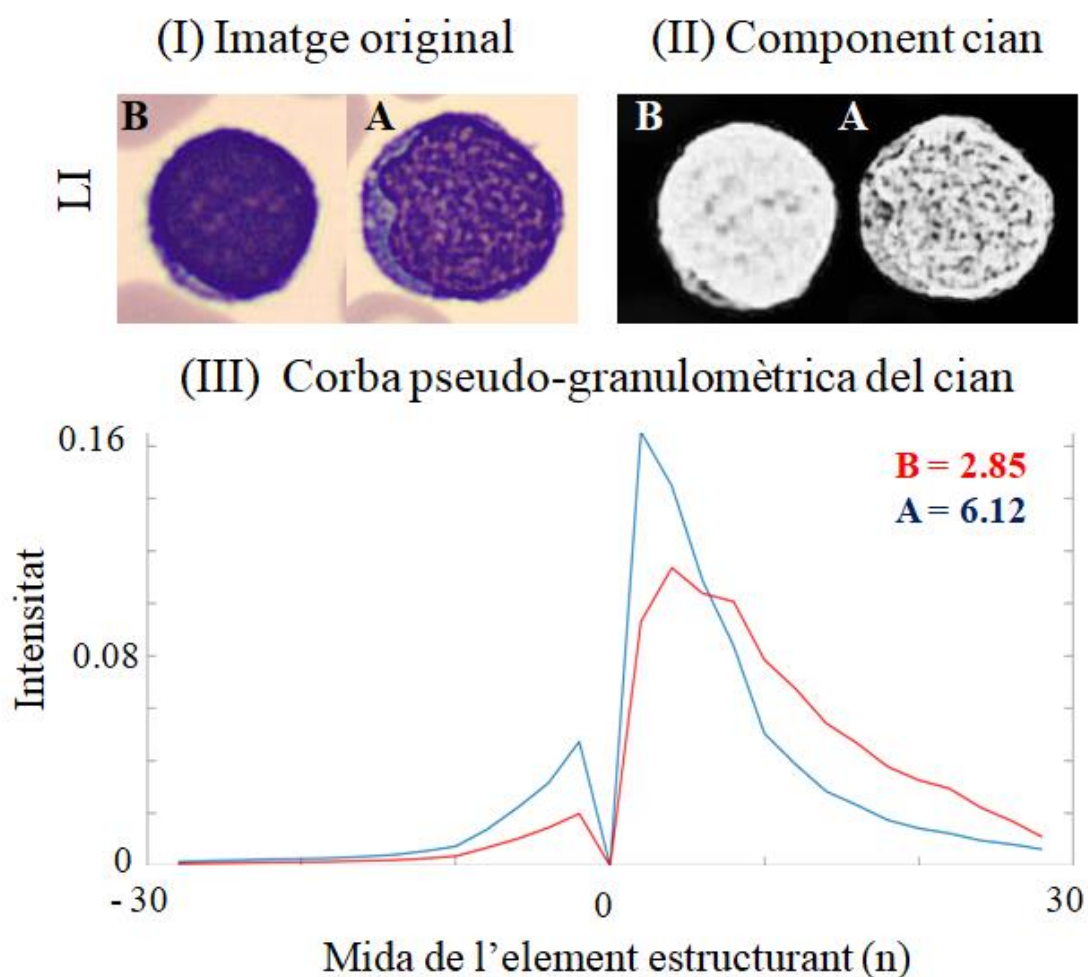


Figura 4.9: Anàlisi d'imatges de la *curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula* (4a posició). (I) Imatges originals de limfòcits normals. (II) Components de color en escala de grisos. (III) Corbes pseudo-granulomètriques. B, baix; A, alt.

Com es pot observar, la Figura 4.9 conté (I) dues imatges originals de limfòcits normals amb el valor més baix ($B = 2,85$) i més alt ($A = 6,12$) de curtosi, (II) els seus components cian que s'obtenen transformant les imatges originals en escala de grisos i (III) les corbes pseudo-granulomètriques corresponents dels components cian. La part esquerra de la corba (mida de partícula negativa) representa la densitat de mida dels grànuls foscos, mentre que la part dreta de la corba representa la densitat de mida dels grànuls brillants. S'ha de tenir en compte que quan hi ha més proporció de grànuls brillants a la imatge de la cèl·lula limfoide (imatge B), es veu una major contribució del component cian a la imatge, resultant en una major uniformitat i una corba més ampla (menor curtosi, corba vermella).

La Figura 4.10 mostra els diagrames de caixes de la curtosi de la corba pseudo-granulomètrica del component cian de la regió de la cèl·lula per als 12 grups de cèl·lules limfoides: les medianes més elevades es troben a la LPB (5,9) i els BL (5,6), mentre que les més baixes es troben a l'HCL (3,3) i la LLGG (3,6) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística en totes les parelles ($p < 0,0001$, Taula 4.4).

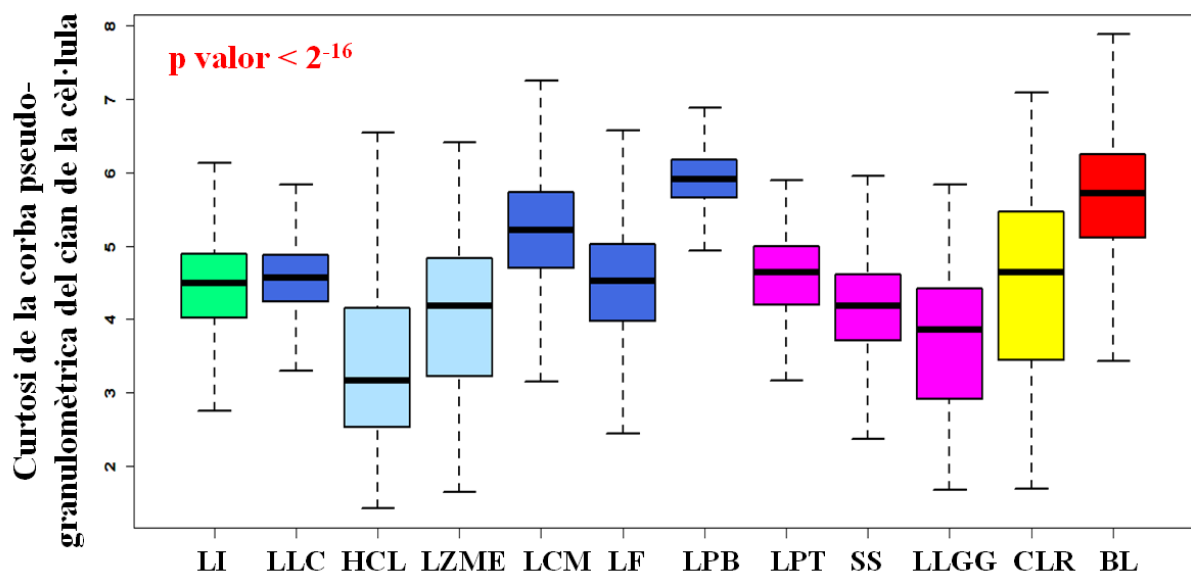


Figura 4.10: Diagrames de caixes obtinguts per al descriptor de la curtosi de la corba pseudo-granulomètrica del component cian de la cèl·lula per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El cinquè millor descriptor va resultar ser l'asimetria de l'histograma del component blau de la cèl·lula. Es tracta d'un descriptor de color, concretament, un estadístic de primer ordre. Com hem vist anteriorment, aquest tipus de descriptors es basen en l'histograma i a la Figura 4.11 es mostra un exemple per tal de comprendre millor la seva relació amb les característiques discriminants morfològiques. La resta de descriptors basats en l'histograma d'un component de color tenen interpretacions similars.

A la Figura 4.11, podem veure (I) dues imatges de CLR amb el valor més baix ($B = -0,29$) i el més alt ($A = 0,64$) d'asimetria, (II) els corresponents components blaus de la cèl·lula i (III) els histogrames de les intensitats de blau. Ambdós histogrames tenen una distribució bimodal, però tenint en compte la distribució de la intensitat de blau, s'observa que la distribució B té una asimetria negativa, mentre que és positiva a la distribució A. La imatge B mostra una major quantitat de píxels brillants (localitzats al citoplasma), de manera que la distribució de l'histograma es desplaça cap a la dreta. Per contra, una asimetria positiva s'observa a la imatge A i és degut a que la distribució de l'histograma (blau) es desplaça cap a l'esquerra perquè la majoria de píxels són foscos (pertanyents al nucli).

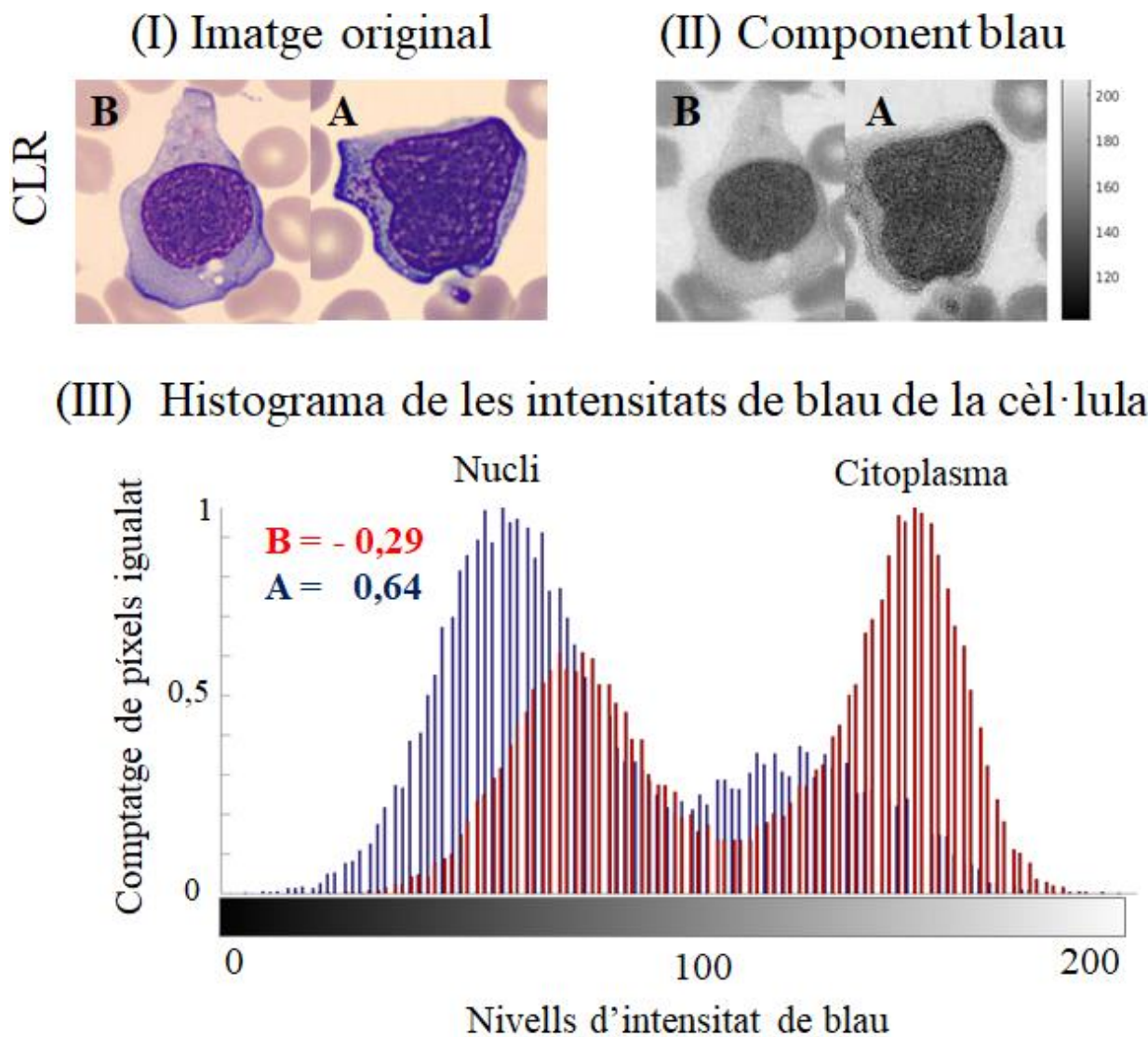


Figura 4.11: Anàlisi d'imatges de l'asimetria de l'histograma del component blau de la cèl·lula (5a posició). (I) Imatges originals. (II) Components de color blau en escala de grisos. (III) Histogrames. B, baix; A, alt.

La Figura 4.12 conté els diagrames de caixes corresponents a aquest descriptor. Les medianes més elevades s'observen als limfòcits anormals de la LLC (0,97) i SS (0,92), mentre que les més baixes es troben a l'HCL (0,18) i les CLR (0,23) ($p < 0,0001$). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, l'asimetria de l'histograma del component blau de la cèl·lula ha estat significativa per a discriminar entre les següents parelles: LI versus LLC, HCL versus LZME, LPB versus LPT i CLR versus BL ($p < 0,0001$, Taula 4.4).

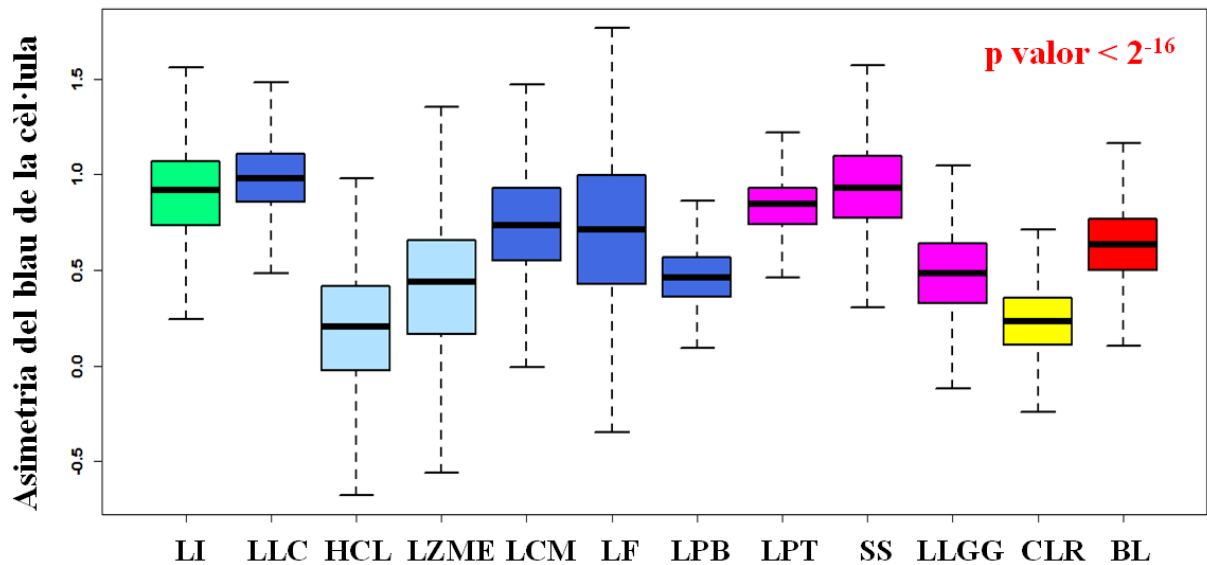


Figura 4.12: Diagrames de caixes obtinguts per al descriptor de l'asimetria de l'histograma del component blau de la cèl·lula per als diferents grups de cèl·lules limfoides incloses a l'estudi.

En relació al sisè descriptor, la *information measure of correlation 1 (IMC1)* del component magenta de la regió de la cèl·lula, els BL i el LF presenten les medianes més elevades (ambdós -0,29), mentre que els tricoleucòcits la més baixa (-0,42) ($p < 0,0001$, Fig. 4.13). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, la *IMC1* del component magenta de la cèl·lula ha estat significativa per a discriminar entre totes les parelles a excepció de LI versus LLC i LPB versus LPT ($p < 0,0001$, Taula 4.4).

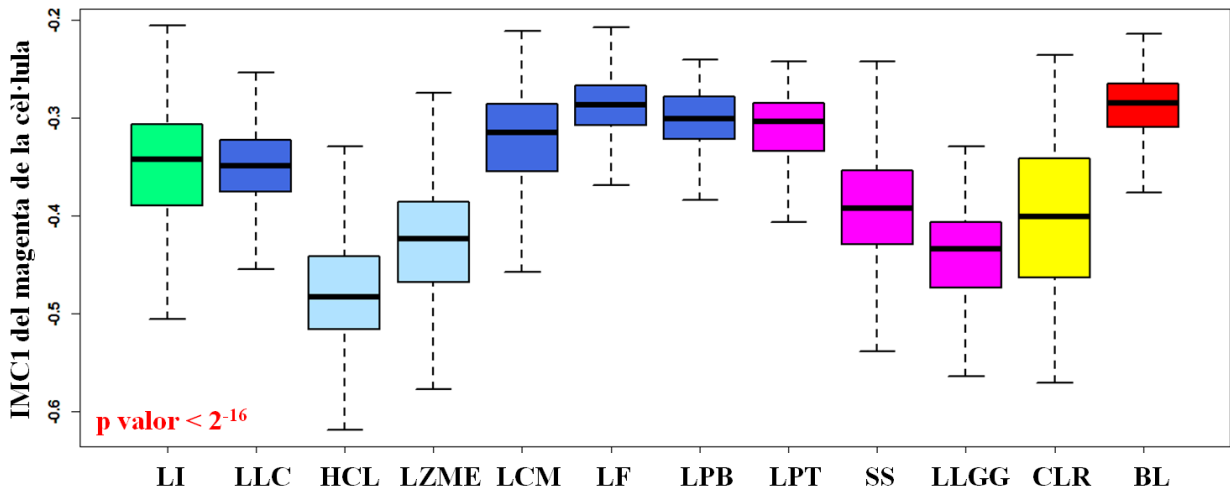


Figura 4.13: Diagrames de caixes obtinguts per al descriptor de la *information measure of correlation 1 del component magenta de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

La DE de l'histograma del component blau de la regió de la cèl·lula és el setè descriptor més rellevant i es calcula a partir del mateix histograma que el cinquè descriptor, però en aquest cas el paràmetre estadístic és la DE de la distribució. La Figura 4.14 mostra (I) dues imatges de prolifèrants B amb el valor més baix ($B = 14,4$) i més alt ($A = 18,2$) per aquest descriptor, (II) els corresponents components blaus de la cèl·lula i (III) els seus histogrames. La DE de la imatge A és major que la de la imatge B ja que la distribució en la imatge A és bimodal, mentre que és unimodal en la imatge B, el que significa menys dispersió (menor DE).

La Figura 4.15 conté els diagrames de caixes corresponents a aquest descriptor. Les medianes més baixes d'aquest descriptor s'observen al LF (10,7) i als BL (12,6), mentre que els limfòcits vellosos de l'HCL (21,0) i el LZME (18,7) mostren les més altes ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística a totes elles ($p < 0,0001$, Taula 4.4).

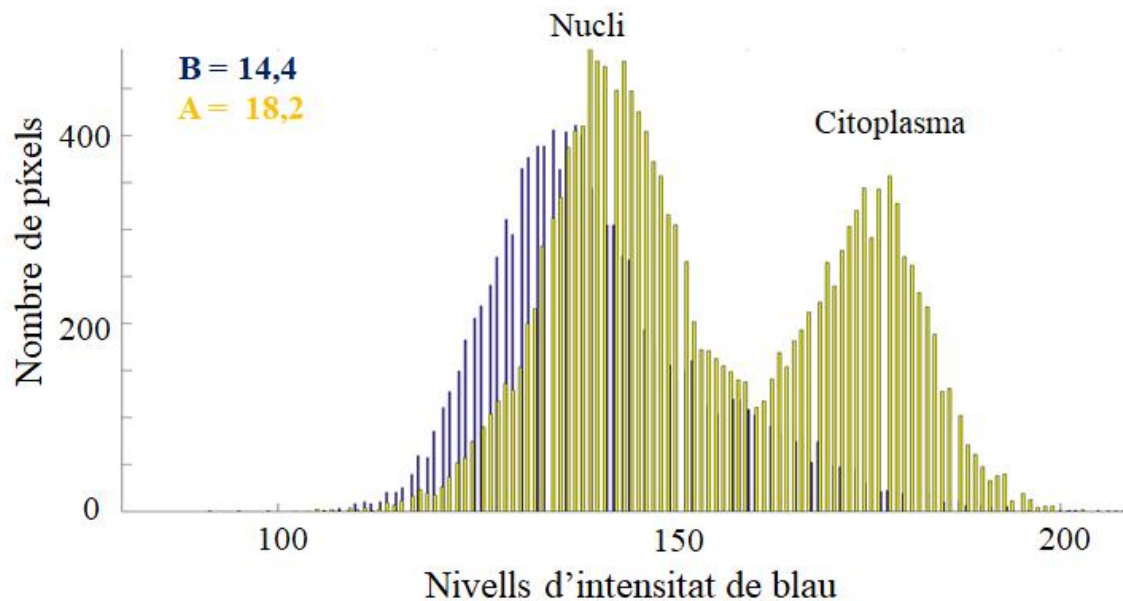
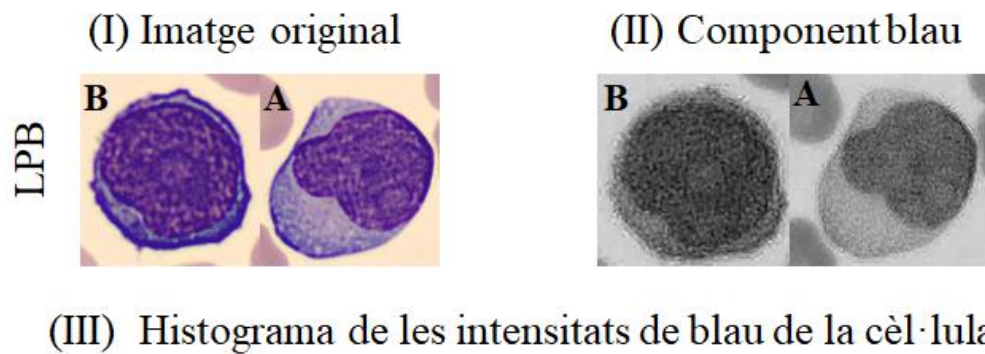


Figura 4.14: Anàlisi d'imatges de la desviació estàndard de l'histograma del component blau de la cèl·lula (7a posició). (I) Imatges originals de dos prolifmòcits B. (II) Components de color blau en escala de grisos. (III) Histogrames de les intensitats de blau. B, baix; A, alt.

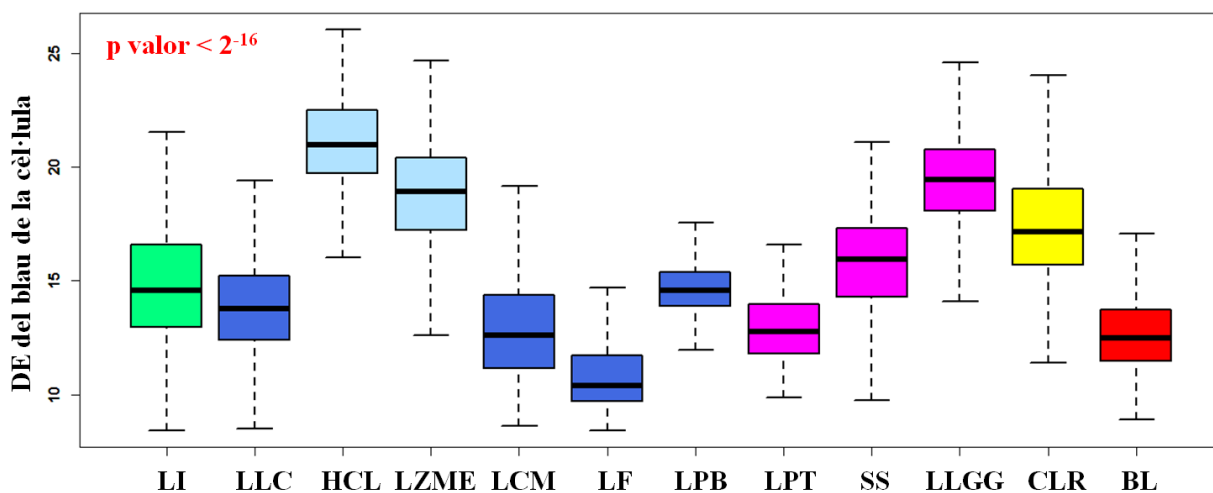


Figura 4.15: Diagrames de caixes obtinguts per al descriptor de la desviació estàndard del component blau de la regió de la cèl·lula per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El cinquè i el setè descriptor s'associen a la intensitat blava de la regió cel·lular, però val la pena observar una relació amb la *relació N/C*. Per a les cèl·lules limfoides amb una baixa relació N/C, l'histograma té una menor asimetria ja que hi ha un equilibri entre els píxels del nucli i del citoplasma. En canvi, la *DE de l'histograma del component blau de la cèl·lula* té una interpretació inversa en comparació amb l'asimetria. De fet, els limfòcits anormals de la LLC, LPT, LF i els BL, són cèl·lules que acostumen a tenir una relació N/C elevada i presenten valors alts d'asimetria (0,97, 0,84, 0,71 i 0,63, respectivament) i valors baixos de DE (13,8, 12,9, 10,7 i 12,6, respectivament) ($p < 0,0001$).

Per contra, cèl·lules amb una baixa relació N/C com ara les de l'HCL, LZME i CLR presenten valors baixos d'asimetria (0,18, 0,41 i 0,23, respectivament) i alts de DE (21,0, 18,7 i 17,4, respectivament, $p < 0,0001$).

Tal i com es mostra a la Taula 4.7, s'observa correlació entre el descriptor de la *relació N/C* i l'*asimetria de l'histograma del component blau de la regió de la cèl·lula* (5è) en la majoria dels grups cel·lulars limfoides mentre que una correlació inversa la trobem a la *DE de l'histograma del component blau de la cèl·lula* (7è). En altres paraules, quan una cèl·lula limfoide conté poc citoplasma, a més de presentar una relació N/C elevada, tindrà valors alts d'asimetria perquè les intensitats de blau seran més elevades a la regió del nucli en comparació amb el citoplasma mentre que la DE presentarà valors baixos degut a la petita variació de les intensitats de blau de la cèl·lula (la majoria de píxels correspondran al nucli).

Taula 4.7: Correlacions de la relació nucli/citoplasma amb quatre descriptors (5è, 7è, 8è i 9è). A la taula es mostren els coeficients r i el rang dels p valors, els qual s'han obtingut mitjançant la correlació de Pearson. Un p valor inferior a 0,05 s'ha considerat estadísticament significatiu.

	Asimetria de l'histograma del blau de la cèl·lula (5è)	DE de l'histograma del blau de la cèl·lula (7è)	Entropia del magenta de la cèl·lula (8è)	Asimetria de l'histograma del verd de la cèl·lula (9è)	Rang dels p valors
LI	0,777	-0,741	-0,678	0,696	$< 2^{-16}$
LLC	0,271	-0,812	-0,583	0,492	$< 2^{-16}$
HCL	0,937	-0,532	-0,087	0,881	$0,03887 - < 2^{-16}$
LZME	0,921	-0,718	-0,598	0,852	$< 2^{-16}$
LCM	-0,136	-0,823	-0,645	0,302	$3,7^{-7} - < 2^{-16}$
LF	-0,603	-0,834	-0,312	-0,345	$< 2^{-16}$
LPB	0,769	-0,451	-0,423	0,513	$1^{-13} - < 2^{-16}$
LPT	0,075*	-0,721	-0,414	0,214	$2,1^{-4} < 2^{-16}$
LLGG	0,945	-0,834	-0,555	0,909	$< 2^{-16}$
SS	0,754	-0,767	-0,520	0,634	$< 2^{-16}$
CLR	0,889	-0,470	-0,184	0,663	$5,3^{-10} - < 2^{-16}$
BL	0,216	-0,648	-0,455	0,269	$4,4^{-15} - < 2^{-16}$

* p valor no significatiu

El vuitè descriptor més rellevant correspon a *l'entropia de l'histograma del component magenta de la regió de la cèl·lula*. L'entropia és una mesura d'incertesa o d'aleatorietat. És a dir, com més alta sigui l'entropia, més informació s'inclourà. A la Figura 4.3, introduïda anteriorment, veiem que conté (I) dues imatges corresponents a cèl·lules de Sézary amb valors d'entropia baixos (3,49) i alts (5,51) d'entropia, (II) els corresponents components magenta en escala de grisos de la regió de la cèl·lula i (III) els seus histogrames. Podem observar que quan l'histograma mostra una distribució heterogènia, tal com es veu per la imatge 2 (histograma blau), l'entropia és més alta. Per contra, a la imatge A, hi ha una distribució més àmplia de les intensitats de píxels magenta (més homogènia), que dona lloc a una menor entropia (histograma vermell).

La Figura 4.16 conté els diagrames de caixes corresponents a aquest descriptor i es pot veure que les CLR, els limfòcits anormals de l'HCL, LZME, LLGG i LPB són els grups de cèl·lules limfoides que presenten valors més alts d'entropia (2,7, 2,6, 2,5, 2,5 i 2,4, respectivament). Per contra, els limfòcits anormals del LF i els prolimfòcits T mostren els valors més baixos d'aquest descriptor (1,4 i 1,5, respectivament, $p < 0,0001$). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, *l'entropia de l'histograma del component magenta de la regió de la cèl·lula* ha estat significativa per a discriminar entre totes les parelles a excepció dels limfòcits vellosos ($p < 0,0001$, Taula 4.4).

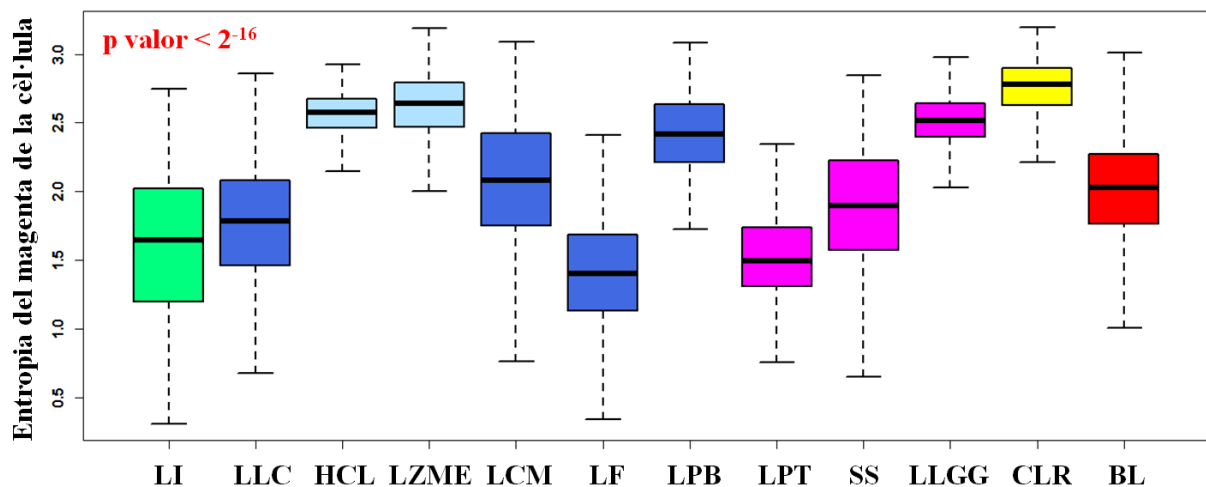


Figura 4.16: Diagrames de caixes obtinguts per al descriptor de *l'entropia de l'histograma de component magenta de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Quan es correlaciona aquest descriptor amb la *relació N/C*, s'observa que els valors més baixos d'entropia corresponen a valors més elevats de *relació N/C* en alguns grups de cèl·lules limfoides (Taula 4.5). Això vol dir que les cèl·lules limfoides que tenen una major proporció de citoplasma i en conseqüència una menor relació N/C, com les CLR o els limfòcits vellosos de l'HCL o el LZME, presentaran valors elevats d'entropia perquè tindran una distribució dels nivells d'intensitat del component de color molt agrupada en una franja de píxels i no hi haurà heterogeneïtat en els valors.

El novè descriptor obtingut és *l'asimetria de l'histograma del component verd de la cèl·lula*, el qual és un descriptor de primer ordre i està relacionat amb la intensitat de verd de tota la cèl·lula. La Figura 4.17 mostra (I) dues imatges de cèl·lules del mantell amb el valor més baix ($B = 0,13$) i alt ($A = 1,98$), (II) els corresponents components verds en escala de grisos i (III) els histogrames d'aquests components verds de la regió cel·lular. Aquest descriptor es calcula com el cinquè (*asimetria del component blau de la cèl·lula*), però, en aquest cas, s'utilitza el color verd com a component de color.

Comparant les dues distribucions dels histogrames del component verd de cada imatge de la Figura 4.17, s'observa que la distribució A (blava) es desplaça cap a l'esquerra, fent-se més asimètrica (més esbiaixada). En canvi, les cèl·lules limfoides que tenen una baixa proporció N/C (imatge B) presenten valors d'asimetria més baixos a causa de la quantitat similar de píxels verds brillants i foscos que hi ha a les dues regions cel·lulars (nucli i citoplasma). Tal i com es mostra a la Taula 4.5, es troba una bona correlació entre la relació N/C i l'asimetria de l'histograma del component verd de la regió de la cèl·lula de cèl·lules limfoides amb citoplasma ampli ($r = 0,852 - 0,909$; $p < 0,00001$).

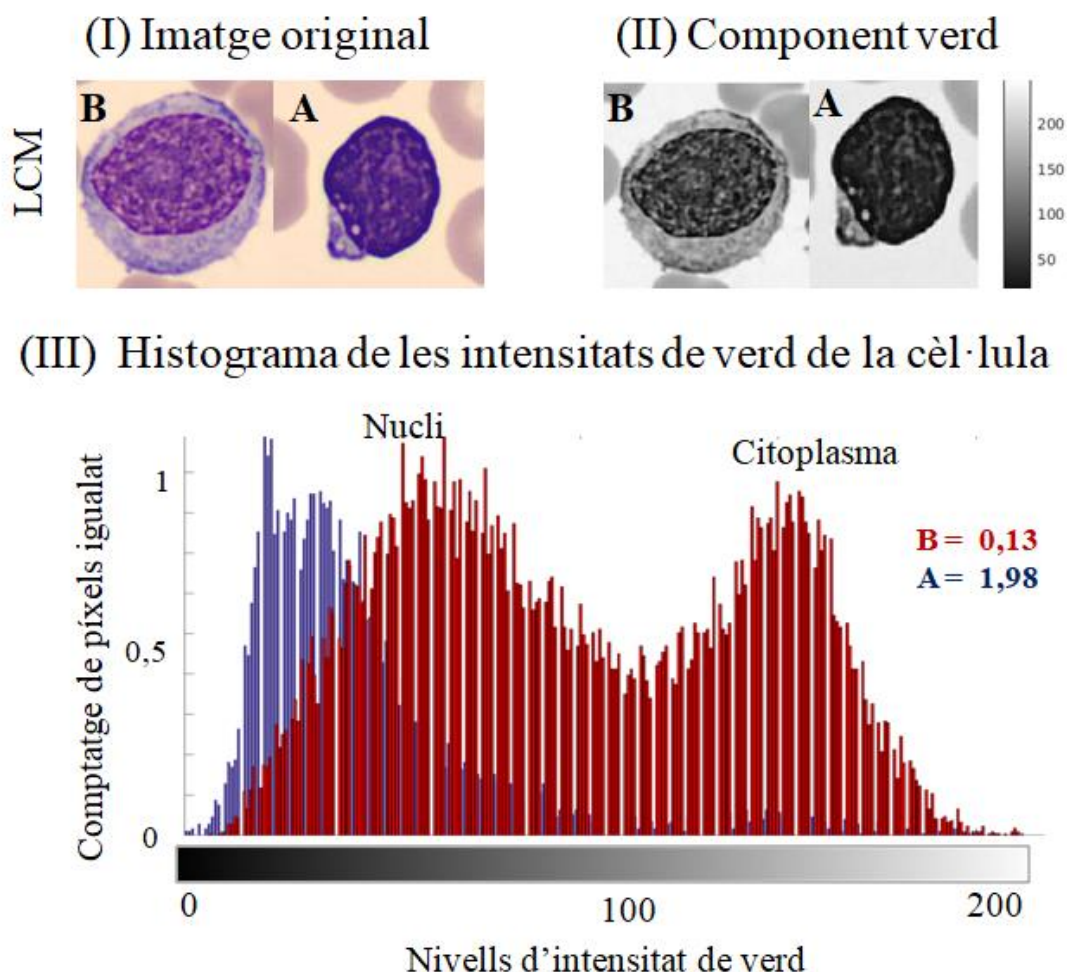


Figura 4.17: Anàlisi d'imatges de *l'asimetria de l'histograma del component verd de la cèl·lula* (9a posició). (I) Imatges originals de dues cèl·lules limfoides anormals d'un limfoma de cèl·lules del mantell. (II) Components de color verd en escala de grisos. (III) Histogrames de les intensitats de verd. B, baix; A, alt.

La Figura 4.18 mostra els corresponents diagrames de caixes *l'asimetria de l'histograma del component verd de la cèl·lula*. Els tricoleucòcits i les CLR són els subconjunts de cèl·lules limfoides que mostren les medianes més baixes (0,30 i 0,41, respectivament) mentre que les més altes s'observen als prolimfòcits T (1,54) ($p < 0,0001$).

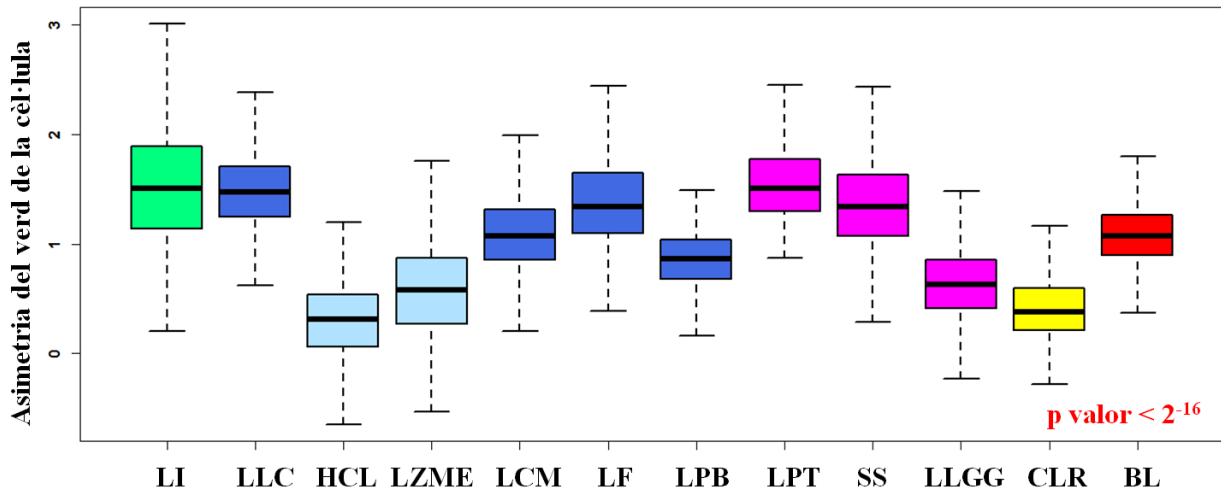


Figura 4.18: Diagrames de caixes obtinguts per al descriptor de *l'asimetria de l'histograma del component verd de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, *l'asimetria de l'histograma del component verd de la cèl·lula* ha presentat significança estadística en totes les parelles a excepció del binomi LI i LLC ($p < 0,0001$, Taula 4.4).

El descriptor més rellevant obtingut del citoplasma és el desè descriptor: *la mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma*. Com més contribueix el component negre a la regió del citoplasma, major valor s'obté d'aquest descriptor. A la Figura 3.22 hem vist un exemple d'aquest descriptor amb una cèl·lula blàstica limfoide. Els prolimfòcits B presenten les medianes més altes (0,106), mentre que els limfòcits anormals del LF les més baixes (0,030) ($p < 0,0001$, Fig. 4.19). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, *la mitjana de la corba pseudo-granulomètrica del component negre del citoplasma* ha estat significativa per a discriminar entre totes les parelles a excepció de l'HCL *versus* LZME i LPB *versus* LPT ($p < 0,0001$, Taula 4.4).

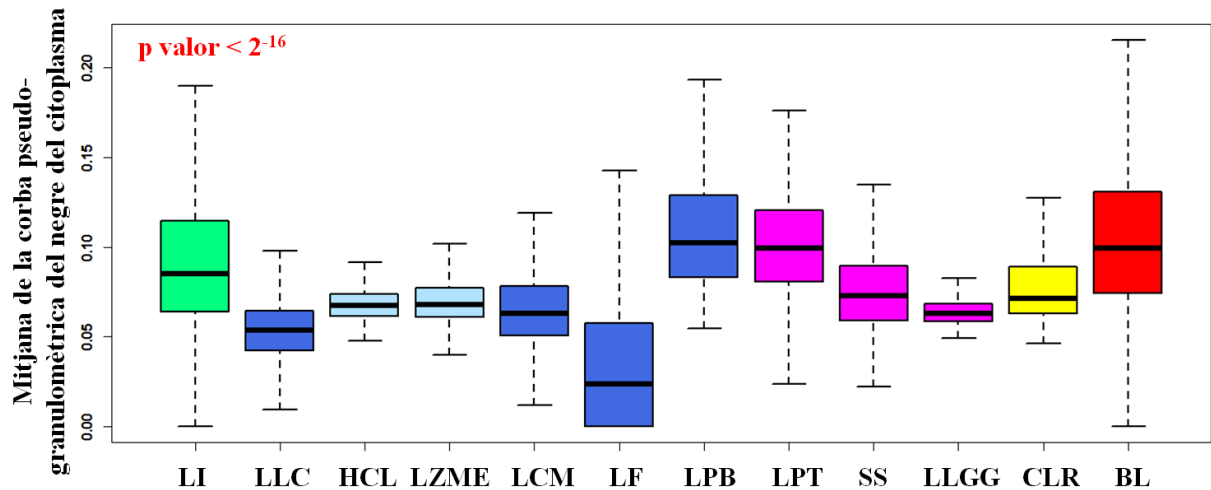


Figura 4.19: Diagrames de caixes obtinguts per al descriptor de *la mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

La mitjana de l'histograma del component blau del nucli és el descriptor més important obtingut del nucli i ocupa l'onzena posició. La Figura 4.20 conté els corresponents diagrames de caixes. La mediana més alta es troba a les CLR (142), mentre que la més baixa als limfòcits normals (128) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística a totes elles ($p < 0,0001$, Taula 4.4).

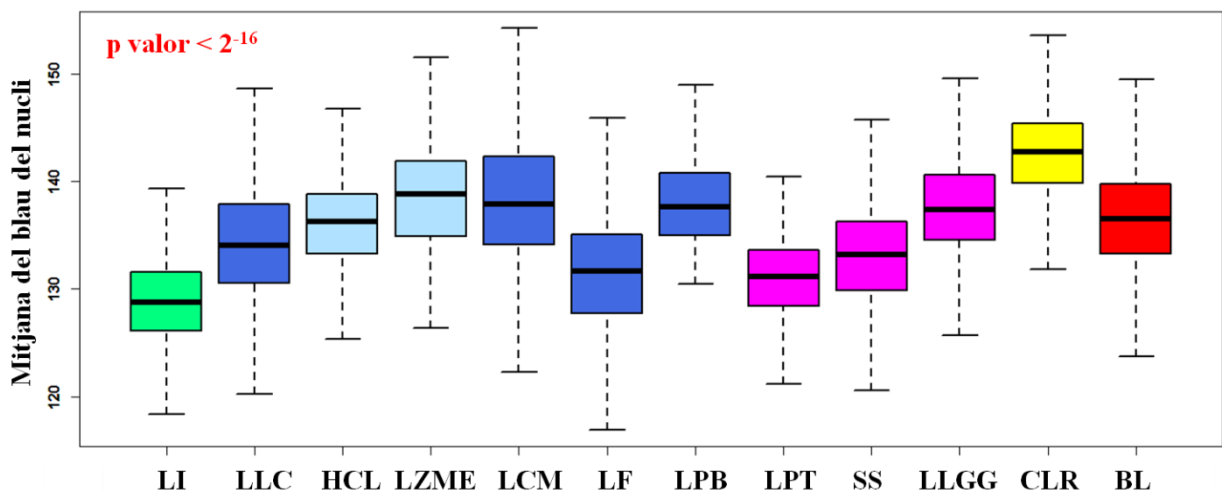


Figura 4.20: Diagrames de caixes obtinguts per al descriptor de *la mitjana de l'histograma del component blau de la regió del nucli* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El dotzè descriptor més rellevant és el *cluster shade del component de la saturació de la regió de la cèl·lula*, el qual és un descriptor estadístic de segon ordre basat en la GLCM. La Figura 4.21 conté els diagrames de caixes corresponents a aquest descriptor.

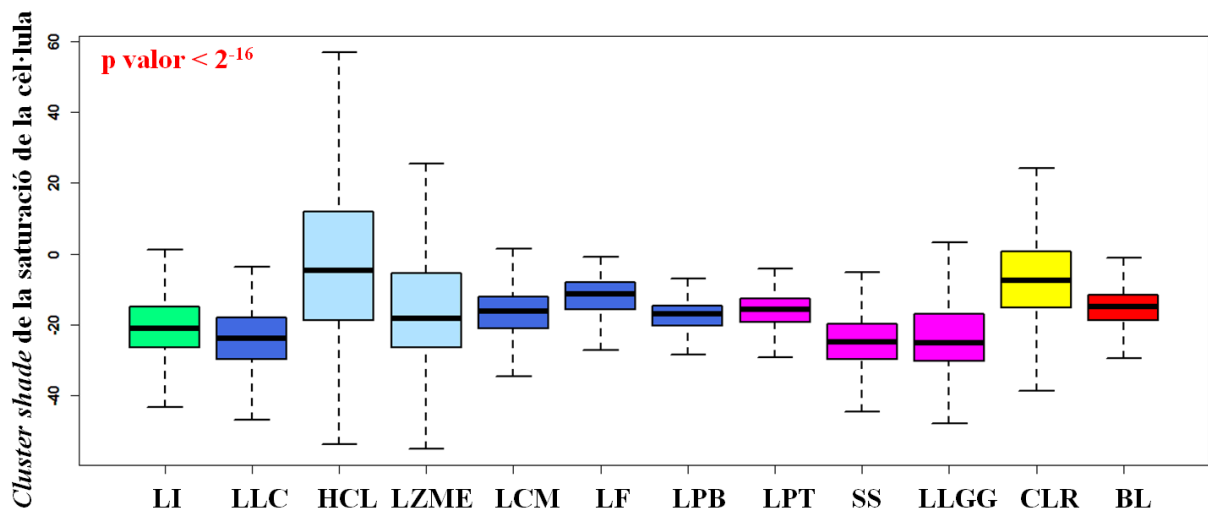


Figura 4.21: Diagrames de caixes obtinguts per al descriptor del *cluster shade del component de la saturació de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Es pot veure que els limfòcits vellosos de l'HCL i les CLR són els grups de cèl·lules limfoides que presenten valors més alts d'aquest descriptor (-4,6 i -7,6, respectivament). Per contra, els limfòcits grans granulars i les cèl·lules de Sézary mostren els valors més baixos (-25,3 i -24,9, respectivament, $p < 0,0001$). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules limfoides anormals analitzades, el *cluster shade del component de la saturació de la cèl·lula* ha estat significatiu per a discriminar entre totes les parelles a excepció de la LPB *versus* LPT ($p < 0,0001$, Taula 4.4).

El segon descriptor més rellevant obtingut de la regió del citoplasma ha estat *la mitjana de l'histograma del component magenta del citoplasma* (posició 13). Les medianes més baixes d'aquest descriptor s'observen al LF (90,7) i als tricoleucòcits (98,7), mentre que els prolimfòcits T i B les més altes (147,7 i 157,7, respectivament) ($p < 0,0001$, Fig. 4.22). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, *la mitjana de l'histograma del component magenta del citoplasma* ha presentat significança estadística en les mateixes cinc parelles que el *cluster shade del component de la saturació de la regió de la cèl·lula* ($p < 0,0001$, Taula 4.4).

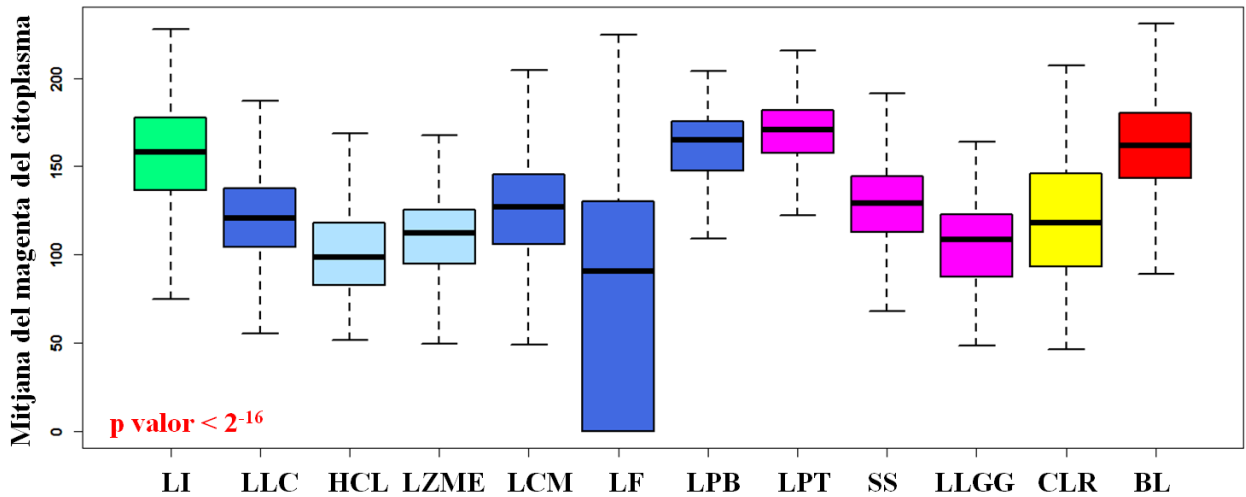


Figura 4.22: Diagrames de caixes obtinguts per al descriptor de *la mitjana de l'histograma del component magenta de la regió del citoplasma* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El catorzè descriptor més rellevant, a l'igual que el tretzè, també s'obté a partir del component magenta però en aquest cas és *la homogeneïtat de la matriu de co-ocurrència de la regió de la cèl·lula*. La Figura 4.23 mostra els diagrames de caixes corresponents a aquest descriptor i es pot veure que els limfòcits anormals del LF i els prolimfòcits T són els grups de cèl·lules limfoides que presenten valors més alts d'aquest descriptor (0,878 i 0,874, respectivament).

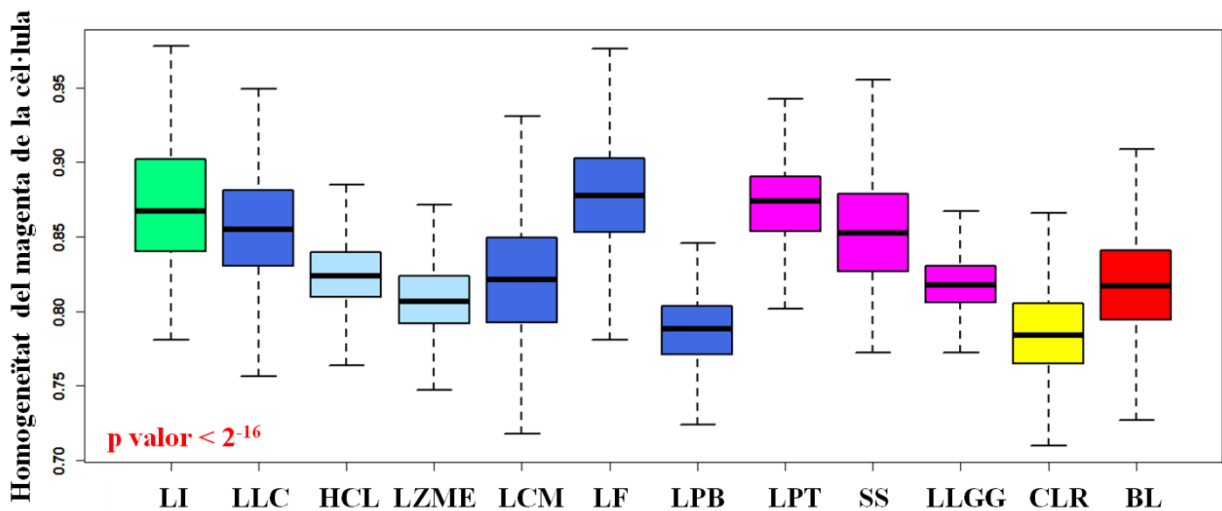


Figura 4.23: Diagrames de caixes obtinguts per al descriptor de *la homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Per contra, els prolimfòcits B i les CLR presenten els valors més baixos (0,789 i 0,784, respectivament, $p < 0,0001$). Pel que fa a la significació estadística obtinguda en les sis parelles de subconjunts de cèl·lules

limfoides anormals analitzades, *la homogeneïtat de la matriu de co-ocurrència de la cèl·lula* ha estat significativa en totes elles ($p < 0,0001$, Taula 4.4).

El quinzè descriptor més rellevant per la diferenciació de les cèl·lules limfoides incloses en el present treball és *la curtosi de l'histograma de la saturació de la cèl·lula*. La Figura 4.24 conté els corresponents diagrames de caixes. Els valors més elevats es troben en les cèl·lules amb una relació N/C elevada com és el cas del LF (1,06), la LPT (0,96) o els LI (0,94) mentre que les medianes més baixes es troben en els tipus de cèl·lules limfoides que tenen una relació N/C baixa com l'HCL (-1,41), el LZME (-1,14), la LLGG (-1,15) i les CLR (-1,13) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística en totes les parelles a excepció del binomi LI i LLC ($p < 0,0001$, Taula 4.4).

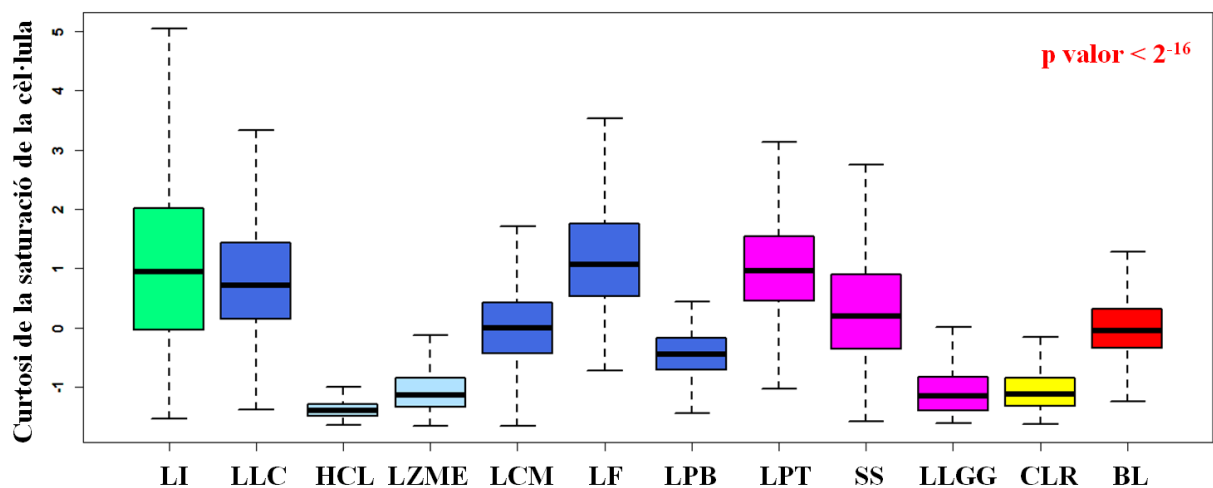


Figura 4.24: Diagrames de caixes obtinguts per al descriptor de *la curtosi de l'histograma del component de la saturació de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El segon descriptor més rellevant obtingut del nucli ocupa la setzena posició i és *la Information measure of correlation 1 del component cian de la regió del nucli*. La Figura 4.25 conté els corresponents diagrames de caixes.

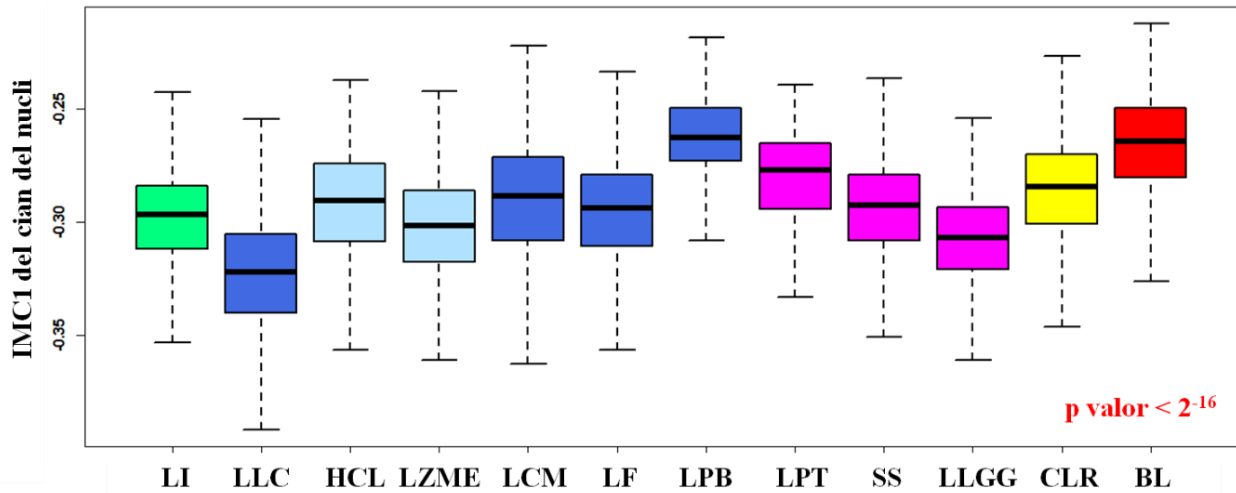


Figura 4.25: Diagrames de caixes obtinguts per al descriptor de *la Information measure of correlation 1 del component cian de la regió del nucli* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

Les medianes més altes es troben a les cèl·lules blàstiques limfoides (-0,26) i als prolimfòcits B (-0,26), mentre que la més baixa als limfòcits de la LLC (-0,32) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística a totes elles ($p < 0,0001$, Taula 4.4).

El divuitè descriptor més rellevant correspon a *la desviació estàndard de l'histograma de la saturació de la regió de la cèl·lula*. La Figura 4.26 conté els corresponents diagrames de caixes. Els valors més elevats es troben a l'HCL (64) i a la LLGG (62) mentre que les medianes més baixes es troben al LF (36) i a la LPT (40) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, aquest descriptor ha presentat significança estadística en totes les parelles d'interès ($p < 0,0001$, Taula 4.4).

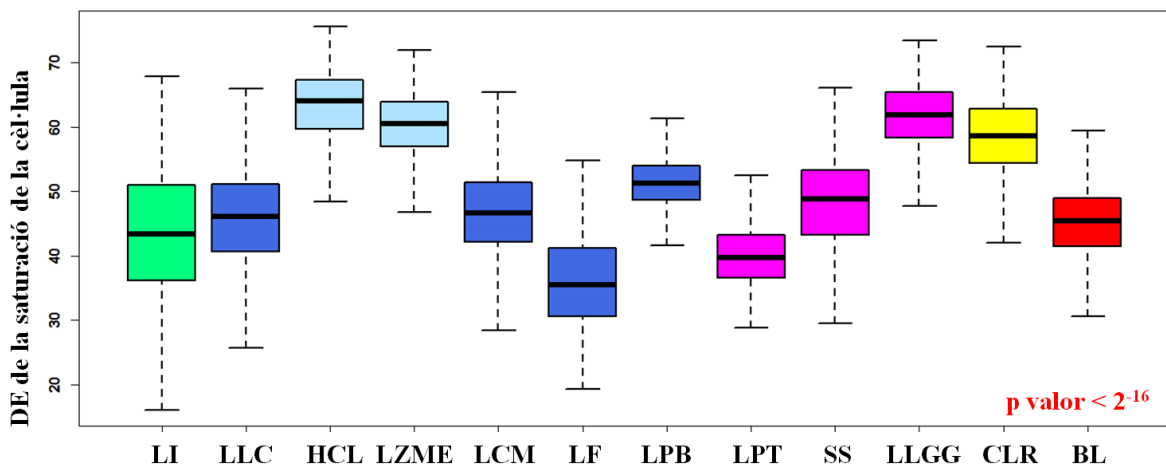


Figura 4.26: Diagrames de caixes obtinguts per al descriptor de *la desviació estàndard del component de la saturació de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi.

El vintè descriptor més rellevant és *l'homogeneïtat del component negre de la regió del citoplasma*, la qual és un descriptor de primer ordre i està relacionat amb la intensitat de negre del citoplasma. La Figura 4.27 conté els corresponents diagrames de caixes. Els valors més elevats es troben a les cèl·lules limfoides reactives (0,75) i les anormals del LF (0,73) mentre que la mediana més baixa es troba a la LLC (0,63) ($p < 0,0001$). En relació a les sis parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil, aquest descriptor ha presentat significança estadística en totes les parelles ($p < 0,0001$, Taula 4.4).

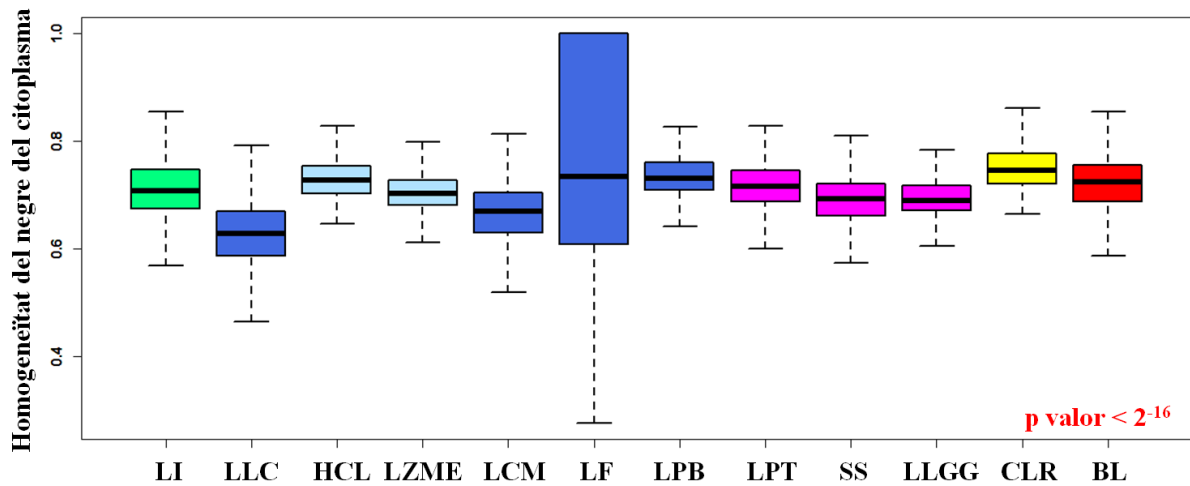


Figura 4.27: Diagrames de caixes obtinguts per al descriptor de *la homogeneïtat del component negre de la regió del citoplasma* als diferents grups de cèl·lules limfoides incloses a l'estudi.

4.4 Discussió

La diferenciació morfològica entre les cèl·lules limfoides anormals continua sent un problema sense resoldre que requereix experiència i habilitat, ja que són les cèl·lules més difícils de reconèixer si s'utilitzen només característiques morfològiques (4).

L'objectiu d'aquesta recerca era utilitzar l'anàlisi digital d'imatges per a contribuir a resoldre aquest problema amb un conjunt de variables citològiques quantitatives que mostressin les següents propietats:

- (i) tenir formulacions quantitatives,
- (ii) tenir interpretació qualitativa en termes de característiques morfològiques
- (iii) ser eficient en la discriminació entre un nombre significatiu de diferents subconjunts de cèl·lules limfoides.

El resultat ha estat la troballa de 20 variables quantitatives (descriptors) que han demostrat la capacitat de diferenciar entre 12 grups de cèl·lules limfoides: normals, reactives, anormals (incloent sis neoplàsies limfoides B i tres de tipus T) i blàstiques. Aquests descriptors es classifiquen en dos grups: cinc són geomètrics i 15 es basen en propietats de color i textura.

Les mesures geomètriques quantitatives són fàcils d'interpretar en termes de característiques morfològiques visuals. En canvi, els descriptors de color i textura tenen una base física però requereixen una base matemàtica més complexa per a una expressió quantitativa i el seu posterior càlcul. Això fa que la interpretació morfològica sigui més complicada en comparació amb els descriptors geomètrics, però aquest treball ha elaborat interpretacions citològiques raonables per ajudar a comprendre les seves capacitats discriminatòries entre els tipus cel·lulars estudiats. Aquesta discussió se centra a destacar qüestions rellevants relatives als descriptors obtinguts en aquest treball.

Pel que fa als descriptors utilitzats en treballs anteriors per a la diferenciació cel·lular limfoide, Benattar i Flandrin (13) i Jahanmehr *et al.* (2) calculen descriptors geomètrics per discriminar entre grups de cèl·lules limfoides quantificant variables citològiques. D'altra banda, Sabino *et al.* (14) utilitza descriptors de textura com les probabilitats de co-ocurrència com a mesura de textura d'una imatge en escala de grisos amb la finalitat de diferenciar entre cinc tipus de leucòcits normals i només un grup de cèl·lules limfoides anormals: la leucèmia limfàtica crònica.

Angulo *et al.* (10) aplica els descriptors geomètrics i de color i textura mitjançant els espais de color *Lab* i *Luv*. En aquest treball s'ha considerat interessant utilitzar sis espais de color per treure profit de l'abundant informació que pot aportar el fet d'utilitzar més espais de color per tal de discriminar entre la gran varietat de tipus de cèl·lules limfoides estudiades. L'ús de sis espais de color dóna un gran nombre de descriptors potencials. Encara que molts d'ells poden ser redundants o no són rellevants per al reconeixement cel·lular, això no és conegut *a priori*.

El mètode de selecció dels descriptors utilitzat ha reduït el nombre de descriptors de color i textura a 15, incloent tres espais de color (CMGN, RGB i HSV) i sis components diferents (cian, magenta, negre, verd, blau i saturació). Els resultats han demostrat que els cinc descriptors geomètrics més rellevants concorden amb les característiques morfològiques de tots els grups cel·lulars limfoides estudiats. En particular, la relació N/C ha estat el descriptor més rellevant, per tant, és el que conté la major informació sobre el tipus de cèl·lula en relació als 12 grups de cèl·lules limfoides.

En el present treball, s'ha demostrat una alta variabilitat en la relació N/C en el LCM a causa del pleomorfisme característic d'aquestes cèl·lules limfoides. Com es pot observar a la Figura 4.4, les dues cèl·lules limfoides velloses (HCL i LZME) i les CLR mostren els valors més baixos, mentre que les cèl·lules limfoides anormals del LF presenten els més alts. Treballs anteriors del nostre grup de recerca també han trobat que la relació N/C és el descriptor més rellevant en la diferenciació entre subconjunts de limfòcits anormals (31) i també entre cèl·lules blàstiques limfoides en comparació amb les cèl·lules limfoides reactives (123).

Altres estudis (2,10,13) han calculat la relació N/C, però sense especificar la seva rellevància. A més, els valors del perímetre nuclear presenten una gran variabilitat en el LCM i la SS. Això és degut a l'heterogeneïtat morfològica en el primer i a la inclusió de les cèl·lules de Sézary petites (variant Lutzner) en la segona.

Com es va descriure per primera vegada en un treball anterior del grup de recerca, el descriptor *hairiness* (17è) es va proposar per a quantificar el perfil del citoplasma en les cèl·lules limfoides velloses de l'HCL. El present treball ha demostrat que aquest descriptor també es pot utilitzar pel LZME. Per tant, aquest descriptor, en efecte, presenta una elevada especificitat per als limfòcits vellosos.

Si analitzem els descriptors de color i textura obtinguts per a tota la cèl·lula, podem observar relacions interessants amb el descriptor de la relació N/C. Fins al coneixement dels autors, no hi ha estudis previs que quantifiquin la relació N/C utilitzant descriptors de color i textura. Les cèl·lules amb un gran citoplasma i una baixa relació N/C (com la de la imatge B de la Figura 4.11) tenen píxels més brillants i mostren un histograma amb menor densitat a causa de la seva simetria. Per aquest motiu, les cèl·lules limfoides, en què la proporció del seu citoplasma és molt baixa i, per tant, la proporció N/C és alta, presenten valors de correlació més baixos per al cinquè i novè descriptor, els quals contenen informació sobre la proporció de píxels brillants i foscos de tota la cèl·lula (veure Taula 4.7).

A més a més, s'ha observat que la desviació estàndard de l'histograma del component blau de la cèl·lula (setè) presenta una correlació inversa amb la relació N/C de tots els grups cel·lulars limfoides. En el cas de l'entropia de l'histograma del component magenta de la regió de la cèl·lula s'ha trobat que valors baixos d'entropia corresponen a valors més elevats de relació N/C però només en alguns grups cel·lulars limfoides. Fins al nostre coneixement, no hi ha treballs previs que hagin utilitzat descriptors basats en l'entropia per al reconeixement de cèl·lules limfoides anormals.

En resum, en aquest treball s'han analitzat els 20 descriptors quantitius més rellevants per a la discriminació entre els 12 grups de cèl·lules limfoides estudiades. És important remarcar que la majoria d'ells han mostrat diferències significatives entre parelles de limfòcits anormals, el reconeixement per morfologia de les quals en frotis de sang perifèrica pot ser difícil, com ara els limfòcits normals *versus* LLC, HCL *versus* LZME, LF *versus* LCM, BL *versus* CLR, LPB *versus* LPT i LPT *versus* SS. A més a més, cinc descriptors quantitius de color i textura han mostrat ser significatius per discriminar les CLR presents en infeccions dels limfòcits anormals observats quan hi ha una neoplàsia limfoide.

L'aplicabilitat d'aquesta investigació rau en el potencial de reduir la subjectivitat associada a la morfologia de sang perifèrica. L'estandardització de l'ús de descriptors quantitius descrits aquí podria ser interessant per donar suport al diagnòstic morfològic diferencial i, molt especialment, quan l'immunofenotip de les cèl·lules limfoides per citometria de flux no és conclouent.

A més a més, a l'actualitat s'estan desenvolupant nous autoanализadors hematològics basats en l'anàlisi digital d'imatges de cèl·lules individuals (52), els quals es podrien beneficiar de l'ús de descriptors quantitius rellevants, com els descrits en aquest treball, per tal de discriminar entre cèl·lules limfoides reactives, anormals i blàstiques. És qüestió de temps que la integració de la citologia combinada amb l'anàlisi digital d'imatges conjuntament amb la citometria de flux es produeixi a la pràctica diària dels laboratoris clínics.

4.5 Conclusions

L'objectiu de la recerca plantejat en aquest Capítol és l'obtenció dels 20 descriptors més rellevants un cop aplicada la teoria de la informació mútua sobre un gran nombre de descriptors (27 geomètrics i 2.649 de color i textura) per a la diferenciació de 12 tipus diferents de cèl·lules limfoides. D'altra banda, s'ha ampliat el nombre de neoplàsies limfoides estudiades a la literatura, ja que la recerca inclou les cèl·lules tipus B anormals del LZME i cèl·lules de tipus T anormals circulant a sang perifèrica a les entitats SS, LLGG-T i LPT.

S'han identificat els 20 descriptors geomètrics, de color i de textura més eficients per a la diferenciació entre nou grups de cèl·lules limfoides circulant a sang perifèrica de pacients amb diferents tipus de neoplàsies limfoides B i T, blasts limfoides, cèl·lules limfoides reactives i limfòcits normals.

Els descriptors obtinguts tenen formulacions quantitatives explícites i admeten interpretacions qualitatives en termes de característiques morfològiques visuals. Els més rellevants per al reconeixement de les diferents cèl·lules limfoides incloses a l'estudi són geomètrics i, en ordre de importància, són els següents: 1) la *relació nucli/citoplasma*, 2) el *perímetre del nucli* i 3) el *diàmetre de la cèl·lula*.

Dels 20 descriptors obtinguts, cinc són geomètrics i 15 de color i textura, dels quals 13 són estadístics (set de primer ordre i sis de segon ordre) i dos són granulomètrics. Només tres espais de color (CMGN, RGB i HSV) i sis components de color (el blau, el verd, el magenta, el negre, el cian i la saturació) dels 19 utilitzats es troben involucrats en la formulació dels descriptors de color i de textura.

Un total de 19 descriptors identificats són significatius per a discriminar els limfòcits anormals del limfoma fol·licular amb expressió perifèrica, respecte als del limfoma del mantell.

Un total de 18 descriptors resulten significatius per a diferenciar els limfoblasts de les cèl·lules limfoides reactives.

Un total de 17 descriptors mostren diferències significatives entre els prolimfòcits T i les cèl·lules de Sézary. Aquesta diferenciació és particularment rellevant donat que ambdues són cèl·lules anormals de neoplàsies limfoides T que comparteixen similituds morfològiques.

Un total dels 16 descriptors són significativament discriminants per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica respecte als limfòcits normals.

Un total de 16 descriptors permeten també diferenciar els limfòcits amb prolongacions citoplasmàtiques de la tricoleucèmia respecte als del limfoma de la zona marginal esplènic leucemitzat i entre prolimfòcits de tipus B i T.

Els cinc descriptors de color i textura següents mostren significació estadística quan es fan comparacions entre les cèl·lules limfoides reactives (benignes) amb cadascun dels diferents tipus de limfòcits anormals

(malignes): 1) *Desviació estàndard de l'histograma del component blau de la cèl·lula*, 2) *Entropia del component magenta de la cèl·lula*, 3) *Mitjana de l'histograma del component blau del nucli*, 4) *Information measure of correlation 1 del component cian del nucli* i 5) *Desviació estàndard del component de la saturació de la cèl·lula*.

Al Capítol 6 utilitzem una nova font per a l'adquisició de les imatges (microscopi BX43) i s'analitzen si els 20 descriptors presentats al present capítol són reproduïbles per a la diferenciació dels diferents grups cel·lulars limfoides.

Seguint amb el propòsit de la *Tesi Doctoral* de la caracterització morfològica dels diferents grups cel·lulars limfoides, el proper Capítol 5 utilitza la mateixa font d'adquisició d'imatges (el sistema automatitzat DM96) però el punt de partida no és el processament digital d'imatges, sinó que partint d'algunes anormalitats cel·lulars establertes *a priori* pels citòlegs, les quals estan associades a cèl·lules i malalties específiques i de particular interès clínic, l'objectiu serà identificar descriptors quantitius específics. Aquests descriptors haurien de descriure objectivament característiques morfològiques que els citòlegs reconeixen de forma visual i expressen habitualment amb conceptes subjectius. El resultat és l'obtenció de nous descriptors que poden contribuir a la identificació objectiva dels limfòcits de la leucèmia limfàtica crònica, la síndrome de Sézary, els limfòcits grans granulars i els limfòcits vellosos no només a l'etapa inicial de diagnòstic, sinó també després de rebre els tractaments, per detectar limfòcits anormals residuals.

Capítol 5

5 Descriptors citològics quantitius per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, les cèl·lules de Sézary, els limfòcits vellosos i els limfòcits grans granulars mitjançant l'anàlisi d'imatges utilitzant el sistema DM96

Basat en: Puigví L, Merino A, Alférez S, *et al.* *Quantitative cytologic descriptors to differentiate CLL, Sézary, granular and villous lymphocytes through image analysis. Am J Clin Pathol.* 2019;5; 152:74-85. doi: 10.1093/ajcp/aqz025

Resum

El present Capítol 5, tot i que utilitza la mateixa font d'adquisició d'imatges (sistema DM96) i procediment d'extracció dels descriptors respecte el Capítol 4, el punt de partida és diferent. En comptes de seguir un procediment purament algorítmic, ara s'aborda el problema d'obtenir descriptors quantitius específics en un entorn on el punt de partida són les anomalies cel·lulars específiques donades *a priori*, que estan associades a cèl·lules i malalties específiques. Aquests descriptors han de descriure característiques morfològiques objectives que els patòlegs generalment reconeixen a partir de la inspecció visual i expressen sobretot amb conceptes subjectius. En aquest sentit, l'objectiu d'aquest Capítol 5 és buscar descriptors per a identificar als limfòcits anormals a la leucèmia limfàtica crònica, a la tricoleucèmia, al limfoma de la zona marginal esplènica, a la síndrome de Sézary i a la leucèmia de limfòcits grans granulars quan aquestes malalties tenen expressió a sang perifèrica. En primer lloc, l'anàlisi d'imatges s'ha aplicat a 768 imatges de 15 tipus diferents de cèl·lules limfoides i monòcits obtingudes de 144 pacients per tal d'obtenir quatre descriptors discriminants de les següents característiques citològiques: la cromatina madura i condensada i la cerebriforme, la granulació citoplasmàtica i les vellositats citoplasmàtiques. En una segona etapa, per a cada descriptor s'han obtingut escales numèriques utilitzant 627 imatges de 79 pacients nous. Per últim, s'ha realitzat una avaluació de l'eficiència diagnòstica dels quatre descriptors específics per a la classificació d'un frotis en una única característica cel·lular, en la qual s'han inclòs frotis de sang perifèrica de 209 nous pacients. Pel que fa als resultats obtinguts, la *correlació del cian del nucli* identifica la cromatina madura i condensada de la leucèmia limfàtica crònica i la *desviació estàndard de la corba granulomètrica del cian del nucli* és específica per a la cromatina cerebriforme. Per últim, *l'asimetria de l'histograma del component de color u de la regió del citoplasma* ha mostrat una gran especificitat per a la detecció de la granulació citoplasmàtica. El descriptor *hairiness* ha mostrat especificitat per les vellositats citoplasmàtiques. Pel que fa als frotis utilitzats en l'avaluació de l'exactitud dels quatre descriptors específics, el 96 % van ser classificats correctament. En resum, els descriptors presentats en aquest Capítol 5 poden contribuir a la identificació morfològica de les cèl·lules limfoides anormals considerades en aquest estudi.

5.1 Introducció

Algunes característiques citològiques proporcionen una descripció específica de cèl·lules diana en diferents malalties hematològiques (56). Tot i que la morfologia sanguínia és important en el diagnòstic de malalties hematològiques i no hematològiques, les característiques citològiques distintives són subjectives i es veuen influenciades per la interpretació del citòleg (2).

L'anàlisi d'imatges és una àrea de recerca molt dinàmica que ha vist una gran quantitat de publicacions en l'última dècada (124) i les imatges digitals s'han convertit en un component integral de la pràctica i l'educació en la patologia contemporània (125).

Les característiques geomètriques s'interpreten fàcilment en termes quantitius per descriure i comparar variables com el perímetre, l'àrea o les prolongacions velloses. A més, l'experiència adquirida a partir de treballs (2,10,123,126) mostra que l'anàlisi d'imatges pot proporcionar descriptors quantitius basats en eines matemàtiques ben establertes, per a altres patrons geomètrics menys visibles i altres atributs físics com la textura i el color. Tot i que en alguns casos pot no tenir una interpretació visual directa intuïtiva, l'interès pràctic és que aquests descriptors es puguin calcular mitjançant un algoritme dedicat directament des de la sang i tinguin possibilitat de distingir entre cèl·lules malignes caracteritzades per presentar diferències morfològiques subtils, com per exemple, patrons anormals de cromatina, grànuls o prolongacions velloses (127,128).

Al treball presentat al Capítol anterior (126), s'han presentat un conjunt de 20 descriptors quantitius (geomètrics, de color i textura) que eren rellevants per a la diferenciació entre cèl·lules limfoides normals, reactives, anormals i blàstiques de sang perifèrica. Com en la majoria de la literatura actual, l'aprenentatge automàtic és el punt de partida de l'estudi, de manera que els descriptors s'obtenen mitjançant un procediment purament algorítmic i després s'analitzen les seves habilitats discriminants. D'altra banda, és interessant abordar el problema d'obtenir descriptors quantitius específics en un entorn on el punt de partida són les anomalies cel·lulars específiques donades *a priori*, que estan associades a cèl·lules i malalties específiques. Aquests descriptors han de descriure característiques morfològiques objectives que els patòlegs reconeixen generalment a partir de la inspecció visual i les expressen sobretot de forma subjectiva. Per tant, el present estudi utilitza la mateixa metodologia que el Capítol 4 en quant a l'adquisició automàtica de les imatges mitjançant el sistema DM96 com l'extracció dels descriptors, però el punt de partida són, com ja s'ha comentat, determinades alteracions morfològiques específiques de neoplàsies limfoides concretes.

En aquest sentit, l'objectiu d'aquest Capítol 5 és proposar i analitzar nous descriptors per a proporcionar escales numèriques objectives per tal de quantificar aquelles característiques morfològiques que el citòleg utilitza de manera subjectiva per tal de reconèixer les cèl·lules limfoïdes anormals al microscopi. Per tant, aquest treball representa un avenç en comparació amb els treballs anteriors en el camp (3,10,113,123,126) i proporciona objectivitat en l'avaluació morfològica de cèl·lules limfoïdes anormals de sang perifèrica (SP).

El present Capítol 5 se centra en la vessant morfològica, de manera que el fons tècnic i la formulació dels descriptors són bastant reduïts. Al Capítol 3 i secció 3.4 es pot trobar més informació sobre els descriptors de color i textura així com a (126,128). La present recerca considera tres aspectes:

- 1) Descriu com s'utilitza l'anàlisi d'imatges per a determinar els millors descriptors per a quantificar quatre característiques citològiques (la cromatina madura i condensada, la cromatina cerebriforme, la granulació citoplasmàtica i les vellositats citoplasmàtiques).
- 2) Presenta un estudi estadístic que proporciona escales numèriques per a cada descriptor per tal de relacionar-lo amb la corresponent característica qualitativa i cèl·lula limfoïde anormal específica utilitzant 15 tipus de cèl·lules limfoïdes diferents
- 3) Presenta un estudi d'avaluació on un algoritme que implementa el càlcul dels descriptors, s'utilitza per a frotis individuals en una prova cega per verificar que són específics i sensibles per a identificar una característica citològica específica.

Fins al nostre coneixement, només dues publicacions (2,13) han quantificat característiques citològiques, tot i que no han analitzat tants grups diferents com el treball que es presenta en aquest Capítol. A més, aquest estudi quantifica per primera vegada les següents anomalies cel·lulars, dues per patró de cromatina nuclear i dues per a les característiques del citoplasma:

- 1) Cromatina madura i condensada per limfòcits anormals de la leucèmia limfàtica crònica
- 2) Cromatina cerebriforme per a cèl·lules Sézary a la síndrome de Sézary
- 3) Grànuls citoplasmàtics azuròfils per a limfòcits de la leucèmia de limfòcits grans granulars T
- 4) Prolongacions velloses a per limfòcits vellosos de la tricoleucèmia i el limfoma esplènic de zona marginal

5.2 Material i mètodes

Aquest estudi s'ha realitzat en dues etapes que es descriuen a continuació.

5.2.1 Obtenció dels descriptors citològics quantitius i les corresponents escales numèriques

La Figura 5.1 detalla els diferents processos inclosos en la primera etapa, la qual conté cinc subetapes. La primera consisteix en l'adquisició de les imatges, seguidament la seva anàlisi amb la segmentació i l'extracció dels descriptors així com la selecció de les característiques citològiques objecte d'estudi. Posteriorment es seleccionen els descriptors mitjançant anàlisi estadística i per últim, s'obtenen escales numèriques per a 12 cèl·lules limfoides anormals diferents i limfoblasts així com per als limfòcits normals i reactius.

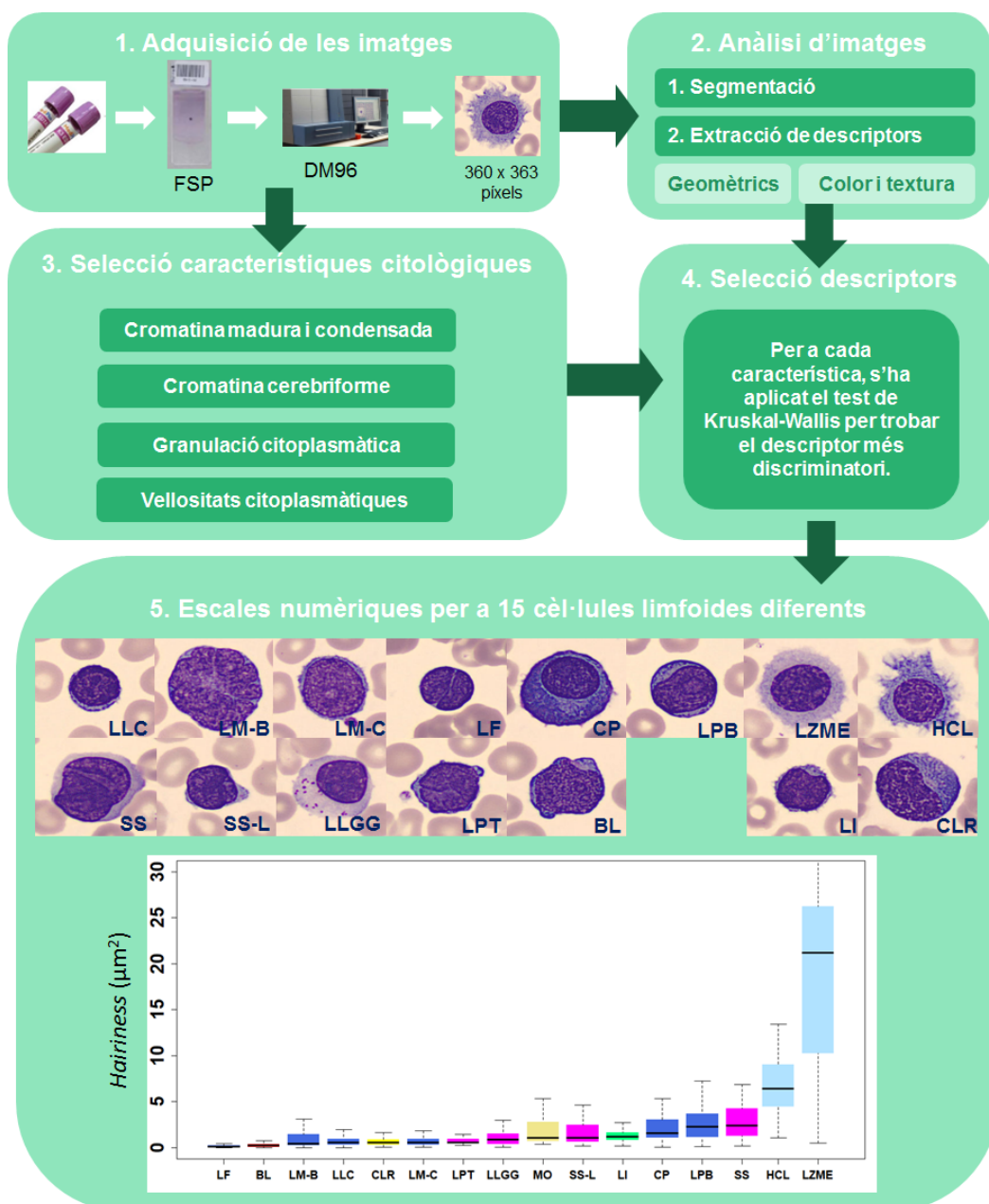


Figura 5.1: Esquema de la primera etapa de l'estudi.

A continuació es detallen les subetapes anteriorment mencionades:

1) Preparació de les mostres de sang i adquisició de les imatges

En aquesta primera etapa s'ha utilitzat un primer conjunt de 768 imatges corresponents a 144 pacients (veure Taula 5.1) amb les següents neoplàsies limfoides: leucèmia limfàtica crònica (LLC), leucèmia prolimfocítica B (LPB), tricoleucèmia (HCL), limfoma de la zona marginal esplènica (LZME), LCM amb morfologia blàstica (LM-B) o clàssica (LM-C), limfoma fol·licular (LF), leucèmia prolimfocítica de cèl·lules T (LPT), leucèmia de limfòcits grans granulars tipus T (LLGG) i la síndrome de Sézary, incloent la variant clàssica (SS) i la variant Lutzner de cèl·lules petites (SS-L).

La raó per dividir algunes malalties en les variants corresponents és per assegurar-se que els descriptors analitzats cobreixen ambdues variants morfològiques. En el cas de la síndrome de Sézary, es pretén mostrar que el descriptor quantitatiu de la cromatina cerebriforme no només és capaç d'identificar la variant clàssica sinó també la de mida petita, la qual és més difícil de reconèixer.

A més també s'han considerat cèl·lules limfoides reactives (CLR) de pacients amb infeccions virals (Epstein-Barr, Cytomegalovirus, Dengue, hepatitis, entre d'altres), cèl·lules plasmàtiques de pacients amb mieloma múltiple, cèl·lules blàstiques limfoides de pacients amb leucèmia limfoblàstica aguda així com limfòcits i monòcits d'individus sans.

Taula 5.1: Tipus de cèl·lula, nombre (n) de pacients i imatges utilitzades en el primer conjunt d'imatges.

Tipus de cèl·lula (primer conjunt)	Pacients, n	Imatges, n
Monòcit	16	40
Limfòcit	7	46
Limfòcit anormal de la leucèmia limfàtica crònica	15	124
Prolimfòcit B	2	20
Limfòcit vellós de la tricoleucèmia	18	56
Limfòcit vellós del limfoma de la zona marginal esplènica	13	58
Cèl·lula del limfoma del mantell morfologia blàstica	1	20
Cèl·lula del limfoma del mantell morfologia clàssica	1	15
Limfòcit anormal del limfoma fol·licular	7	46
Cèl·lula plasmàtica	11	30
Prolimfòcit T	1	20
Limfòcit gran granular	8	99
Cèl·lula de Sézary variant clàssica	4	62
Cèl·lula de Sézary variant Lutzner o petita	2	36
Cèl·lula limfoide reactiva	35	69
Blast limfoide	3	27
Total	144	768

De tots els pacients s'ha obtingut el consentiment informat. El *gold standard* per a la classificació cel·lular ha estat el diagnòstic basat en dades clíniques, morfològiques, immunològiques i altres proves complementàries seguint la classificació de l'OMS de 2016 de neoplàsies limfoides (38). En la majoria dels pacients amb neoplàsies limfoides tipus B, el diagnòstic es va realitzar no només tenint en compte la sang perifèrica, sinó també a partir de la biòpsia del gangli limfàtic. A més a més, es va utilitzar la biòpsia cutània per fer el diagnòstic en els pacients amb la síndrome de Sézary.

Els diferents grups de limfòcits inclosos en aquesta etapa han estat seleccionats perquè representen la majoria dels tipus de cèl·lules limfoides que es poden observar a SP. Les mostres de sang s'han obtingut de la rutina diària del Laboratori Core de l'Hospital Clínic de Barcelona. La sang venosa s'ha recollit mitjançant el K3-EDTA com a anticoagulant i els frotis de SP s'han preparat durant les quatre hores posteriors a la seva recollida de manera estandarditzada amb l'extensor-tenyidor i la tinció May Grünwald-Giemsa (SP1000i, Sysmex, Kobe, Japó). Les imatges cel·lulars s'han obtingut amb un sistema d'adquisició d'imatges estandarditzat (sistema DM96, CellaVision, AB, Lund, Suècia) amb una resolució de 360 x 363 píxels i un augment de x 1.000. Abans de procedir al processament digital d'imatges, les facultatives especialistes (LP, AM) han avaluat la qualitat de la morfologia cel·lular.

2) Anàlisi d'imatges

Aquest procés consisteix en dos passos utilitzant algorismes computacionals desenvolupats prèviament pels autors i implementats en el software científic MATLAB® (30,31,59,129): (1) segmentació automàtica per separar quatre regions d'interès: nucli, cèl·lula, citoplasma i zona perifèrica al voltant de la cèl·lula; i (2) extracció automàtica de 2.678 descriptors (29 geomètrics i 2.649 de color i textura). Per descriure la textura de les imatges de SP, s'han considerat quatre tipus de descriptors diferents: estadístics (primer i segon ordre), granulomètrics, *Wavelet* i Gabor. Tal i com s'ha explicat al Capítol 3 secció 3.4.2.2, els estadístics de primer ordre estimen les propietats dels píxels individuals mentre que els de segon ordre descriuen propietats de parells de píxels relatius entre si.

3) Selecció característiques citològiques

La Figura 5.2 il·lustra les quatre característiques citològiques qualitatives objecte d'estudi i les dues categories considerades mútuament exclusives (p. ex., granular *versus* no granular), així com el nombre d'imatges utilitzades.

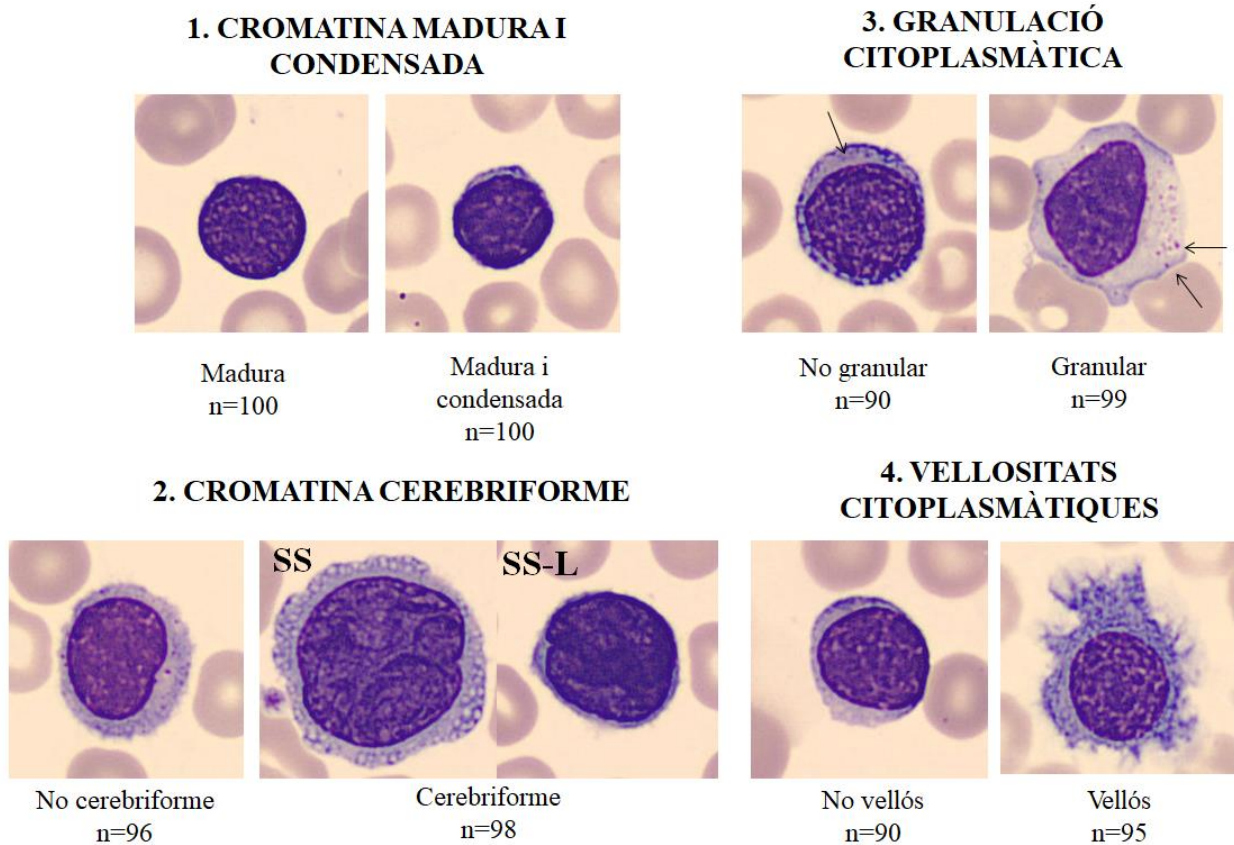


Figura 5.2: Exemples d'imatges de les diferents categories qualitatives considerades per a les quatre característiques citològiques. Per a la cromatina cerebriforme s'han inclòs dues imatges, una per a il·lustrar la cèl·lula de Sézary clàssica (SS) i l'altra de la variant petita (SS-L). Les fletxes indiquen l'absència i presència de grànuls al citoplasma. Les imatges han estat adquirides pel DM96. n, nombre d'imatges

4) Selecció del millor descriptor quantitatiu discriminant

Aquesta subetapa consisteix a determinar el millor descriptor quantitatiu per a distingir cadascuna de les quatre característiques qualitatives entre la categoria corresponent. Atès que en el treball presentat al Capítol 4 (126) tots els descriptors van seguir una distribució no normal i van mostrar variàncies heterogènies, s'ha realitzat el test de Kruskal-Wallis sobre tots ells per tal de seleccionar el descriptor quantitatiu més discriminatori entre categories (per exemple, granular *versus* no granular) per a cadascuna de les característiques citològiques a estudi. Una $p < 0,05$ s'ha considerat estadísticament significativa. R Commander (60), una interfície gràfica d'usuari per al codi R (93), s'ha utilitzat com a programari estadístic.

5) Establiment d'escala numèriques objectives

Un cop obtingut el millor descriptor discriminant, s'han establert escales numèriques objectives de cada característica i categoria mitjançant el càlcul de la mediana i el rang interquartílic (RIQ), que és la diferència entre els quartils superior i inferior. A continuació, s'ha realitzat una prova de validació, la qual és necessària per a la validació dels descriptors i per aquest motiu, s'ha utilitzat un segon conjunt independent format per 627 imatges de cèl·lules de 79 nous pacients (veure el segon conjunt, Taula 5.2). Totes les imatges han estat obtingudes, adquirides i processades a través de les mateixes etapes descrites anteriorment. En total s'han inclòs els mateixos 16 tipus cel·lulars limfoides utilitzats a la primera subetapa. Per a cada descriptor, s'han utilitzat les 627 imatges per obtenir els valors quantitius.

Com s'ha presentat al Capítol 3 secció 3.4.1.5, el descriptor *hairiness* pot prendre múltiples unitats (píxels, micròmetres al quadrat). En el present estudi, es mostra en micròmetres al quadrat i s'ha utilitzat la fórmula 3.4 prèviament descrita al Capítol 3 on 1 píxel equival a $0,009865 \mu\text{m}^2$.

Taula 5.2: Tipus de cèl·lula, nombre (n) de pacients i imatges utilitzades en el segon conjunt d'imatges.

Tipus de cèl·lula (segon conjunt)	Pacients, n	Imatges, n
Monòcit	10	40
Limfòcit	7	40
Limfòcit anormal de la leucèmia limfàtica crònica	4	40
Limfòcit vellós de la tricoleucèmia	4	40
Limfòcit vellós del limfoma de la zona marginal esplènica	2	40
Cèl·lula del limfoma del mantell morfologia blàstica	3	34
Cèl·lula del limfoma del mantell morfologia clàssica	2	40
Limfòcit anormal del limfoma fol·licular	4	42
Cèl·lula plasmàtica	6	36
Prolimfòcit B	1	39
Prolimfòcit T	2	42
Cèl·lula de Sézary variant clàssica	3	30
Cèl·lula de Sézary variant Lutzner o petita	3	50
Limfòcit gran granular	12	39
Cèl·lula limfoide reactiva	14	36
Blast limfoide	2	39
Total	79	627

S'han realitzat comparacions múltiples mitjançant la prova de Dunn aplicant l'ajust de Bonferroni amb el paquet "PMCMR" (121) entre tots els tipus de cèl·lules per verificar que els descriptors seleccionats correlacionen bé amb les característiques morfològiques i són eficients discriminant entre els tipus de cèl·lules corresponents.

5.2.2 Avaluació dels descriptors quantitius per a la classificació de frotis

La segona etapa de l'estudi ha consistit en l'avaluació de l'eficiència diagnòstica dels descriptors que són específics d'un subtipus cel·lular limfoide (veure Figura 5.3). Mitjançant els resultats de l'etapa anterior, s'han associat quatre característiques qualitatives i els seus descriptors quantitius corresponents a determinades neoplàsies limfoides de la manera següent:

- 1) Cromatina madura i condensada per la LLC
- 2) Cromatina cerebriforme per a les cèl·lules de Sézary
- 3) Granulació azuròfila citoplasmàtica per als limfòcits grans granulars
- 4) Vellositats citoplasmàtiques per l'HCL i el LZME

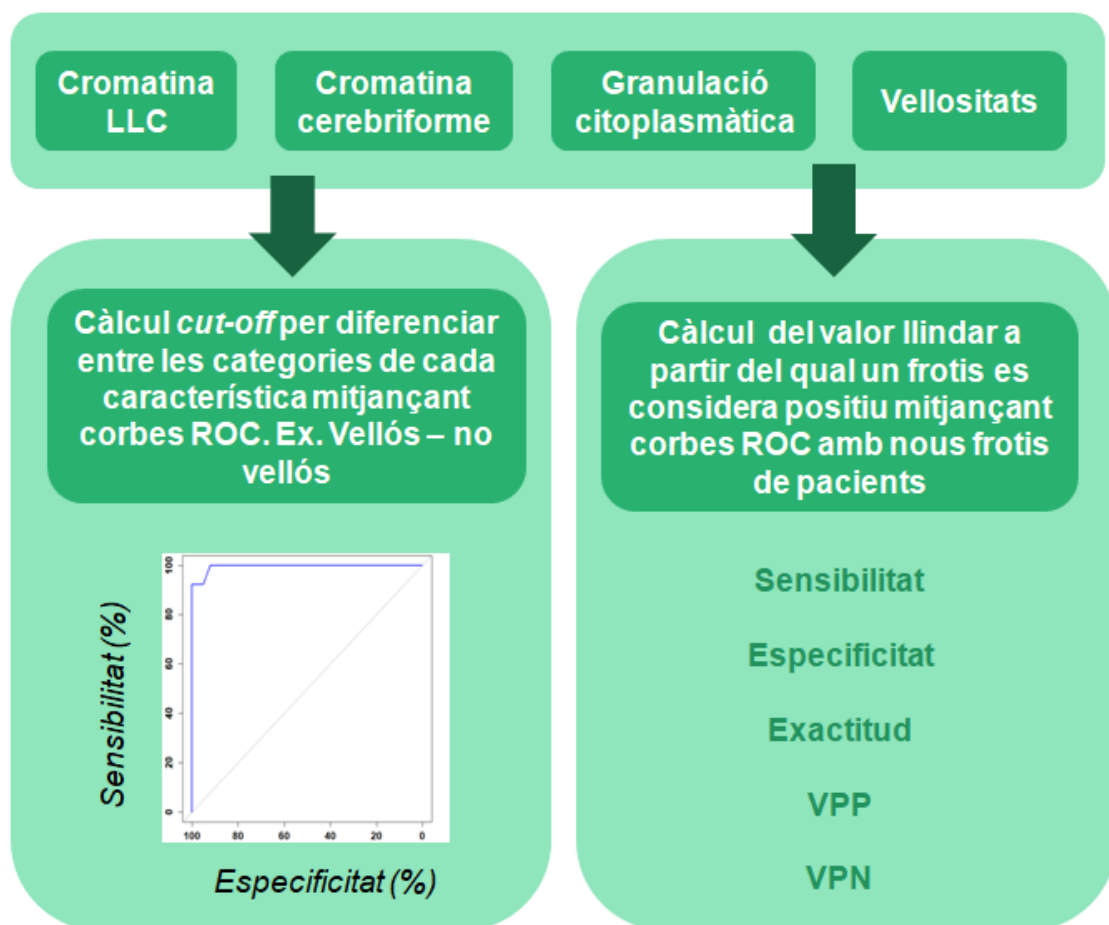


Figura 5.3: Esquema de la segona etapa de l'estudi.

La primera qüestió ha estat calcular un *cut-off* o valor de tall per a cada descriptor. Això vol dir que per a una determinada imatge cel·lular, després de ser segmentada, si el valor del descriptor obtingut és superior o inferior al valor de tall, podem concloure que la cèl·lula de la imatge mostra o no la característica citològica corresponent.

Els *cut-offs* o valors de tall s'han obtingut mitjançant anàlisi ROC sobre el primer conjunt d'imatges (veure Taula 5.1 i Figura 5.2) utilitzant el paquet R 'pROC' (99), el qual permet obtenir el punt de tall òptim per a cada descriptor.

A continuació, els valors de tall obtinguts han estat utilitzats en una prova d'avaluació utilitzant un tercer conjunt d'imatges, on el frotis de sang perifèrica s'ha pres com a unitat d'anàlisi (130). S'han inclòs un total de 209 frotis que contenen de 10 a 98 imatges cel·lulars cadascun, incloent un total de 8.223 imatges (veure Taules 5.3 i 5.4).

Taula 5.3: Diagnòstic dels pacients utilitzats a l'avaluació de l'eficiència diagnòstica dels descriptors (segona etapa de l'estudi) d'acord amb la classificació de l'OMS 2016, immunofenotip de les cèl·lules limfoides circulants a sang perifèrica i nombre de pacients (P) inclosos.

Diagnòstic de leucèmia/limfoma	Immunofenotip	P
Leucèmia limfàtica crònica	CD5+, CD19+, CD23+, CD25+, CD20+ i CD10-	95
Tricoleucèmia	CD19+, CD20+, CD11c+, CD25+, CD103+, CD123+. CD10- i CD5-	2
Limfoma de la zona marginal esplènic	CD19+, CD20+, CD5-, CD10- i CD43-	11
Limfoma fol·licular	CD19+, CD20+, CD22+, CD10+, CD5- i CD43-	14
Leucèmia de limfòcits grans granulars	CD2+, CD3+, CD8+, CD7+ i CD57+	2
	CD2+, CD3+, CD8+, CD7- i CD57+	9
Síndrome de Sézary	CD2+, CD3+, CD4+ i CD5+. CD7-, CD8- i CD26-	4
Limfoma/leucèmia limfoblàstica aguda tipus B	CD19+, CD22+, CD79a+, TdT+, CD10+, CD20- i CD200+	12
	CD19+, CD22+, CD79a+, TdT+, CD34+, CD10+, CD13+ i CD33 feble	3
	CD19+, CD22+, CD79a+, TdT+, CD34+, CD10- i CD20-	2
Limfoma/leucèmia limfoblàstica aguda tipus T	TdT+, CD34+, CD1a+, CD3+ citoplasmàtic, CD4+, CD5+, CD7+ i CD8+	3
Infeccions virals o d'altres etiologies		18
Individus sans		35

Taula 5.4: Nombre total d'imatges corresponents a cada diagnòstic i descriptor del tercer conjunt d'imatges.

Descriptor cromatina madura i condensada	
	Imatges
Imatges LLC	2.277
Imatges no LLC	1.557
Descriptor cromatina cerebriforme	
	Imatges
Imatges SS	125
Imatges no SS	840
Descriptor vellositats citoplasmàtiques	
	Imatges
Imatges HCL/LZME	387
Imatges no HCL/LZME	1.572
Descriptor granulació citoplasmàtica	
	Imatges
Imatges LLGG	292
Imatges no LLGG	1.173

HCL, tricoleucèmia; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulars; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

Per a cada característica (p. ex., el descriptor de vellositats citoplasmàtiques), s'han seleccionat dos grups d'imatges de frotis obtinguts de pacients amb diagnòstics confirmats:

- (1) *Positiu*: quan les imatges cel·lulars exhibeixen la morfologia específica de la malaltia hematològica (p. ex. limfòcits vellosos)
- (2) *Negatiu*: quan les imatges cel·lulars no mostren la morfologia específica característica (p. ex. limfòcits no vellosos).

A continuació, per a cada un dels quatre descriptors, s'ha realitzat una prova de classificació binària cega per a la totalitat del frotis. Totes les imatges del frotis s'han processat de la següent manera:

1. La imatge és capturada i segmentada en les regions d'interès.
2. Els descriptor quantitatiu és calculat i la imatge és considerada positiva si el valor corresponent es troba per sobre del punt de tall, i negativa si és el contrari.
3. La totalitat del frotis s'ha classificat com a positiu quan el rati d'imatges cel·lulars individuals classificades com a positives és superior a un valor límit preestablert (per exemple, classificat com a frotis d'HCL o LZME perquè el rati de cèl·lules positives és superior al llindar preestablert). Aquest llindar s'ha obtingut mitjançant una segona anàlisi ROC.

4. Càlcul dels següents paràmetres:

- Sensibilitat = $\frac{\text{Veritables positius} + \text{Falsos negatius}}{\text{Veritables positius}}$

- Especificitat = $\frac{\text{Veritables negatius}}{\text{Veritables negatius} + \text{Falsos positius}}$

- Valor predictiu positiu (VPP) = $\frac{\text{Veritables positius}}{\text{Veritables positius} + \text{Falsos positius}}$

- Valor predictiu negatiu (VPN) = $\frac{\text{Veritables negatius}}{\text{Veritables negatius} + \text{Falsos negatius}}$

- Exactitud (percentatge dels frotis que han estat identificats correctament) =

$$\frac{\text{Veritables positius} + \text{Veritables negatius}}{\text{Veritables positius} + \text{Veritables negatius} + \text{Falsos positius} + \text{Falsos negatius}}$$

5.3 Resultats

Aquest apartat es presenta d'acord amb les dues etapes detallades anteriorment.

5.3.1 Determinació dels descriptors citològics quantitius i les corresponents escales numèriques

Els resultats d'aquesta primera part es resumeixen a les Taules 5.5 i 5.6 i a les Figures 5.4 – 5.10. Els dos descriptors quantitius relacionats amb la distribució de la cromatina de la LLC i les cèl·lules Sézary han estat la "*correlació del cian del nucli*" (descriptor de textura) i la "*desviació estàndard (DE) de la corba granulomètrica del cian del nucli*" (descriptor de textura), respectivament. A més, els descriptors quantitius relacionats amb grànuls citoplasmàtics azuròfils i les vellositats citoplasmàtiques que s'han trobat són "*l'asimetria de l'histograma del component u del citoplasma*" (descriptor de color) i el "*hairiness*" (descriptor geomètric), respectivament.

A la Taula 5.5 es detallen les quatre característiques citològiques qualitatives inicials amb les seves categories seleccionades *a priori* i la corresponent característica quantitativa discriminatòria més adequada. També proporciona els valors (mitjana i RIQ) de tots els descriptors obtinguts per a cada categoria, el *p* valor i el *cut-off* obtingut mitjançant anàlisi ROC.

Taula 5.5: Regió d'interès, característica citològica qualitativa, categoria, descriptor quantitiu, nombre d'imatges, valors llindars mostren la mediana i el rang interquartílic (RIQ) i els p valors obtinguts de la comparació per parelles utilitzant el test de Kruskal-Wallis post-hoc i Dunn aplicant l'ajust de Bonferroni. Una $p < 0.05$ s'ha considerat estadísticament significativa. DE, desviació estàndard.

Regió d'interès	Característica citològica qualitativa	Categoria	Descriptor quantitiu	Nombre d'imatges (conjunt 1)	Mediana (RIQ)	<i>p</i> valors
Nucli	Cromatina LLC	No LLC	<i>Correlació del Cian del nucli</i> ($\times 10^{-2}$)	100	75,8 (4,3)	$p < 0.0001$
		LLC		100	79,7 (2,6)	
	Cromatina cerebriforme	No cerebriforme	<i>DE de la corba granulomètrica del cian del nucli</i> ($\times 10^{-3}$)	96	30,0 (4,7)	$p < 0.0001$
		Cebriforme		98	20,9 (5,3)	
Cito-plasma	Granulació citoplasmàtica	No granular	<i>Asimetria de l'histograma del component u del citoplasma</i>	90	0,26 (0,40)	$p < 0.0001$
		Granular		99	2,21 (1,59)	
Regió externa	Vellositats citoplasmàtiques	No vellós	<i>Hairiness</i> (μm^2)	90	1,0 (1,1)	$p < 0.0001$
		Vellós		95	5,9 (8,2)	

La Taula 5.6 proporciona els valors dels descriptors obtinguts per a cada neoplàsia limfoide utilitzant les imatges incloses en el conjunt 2.

Taula 5.6: Valors quantitatius dels quatre descriptors que han permès la quantificació de les corresponents característiques citològiques qualitatives per als diferents grups cel·lulars del conjunt 2 d'imatges així com la regió d'interès a partir de la qual s'ha calculat cada descriptor. Els valors mostren la mediana (rang interquartílic).

Característica qualitativa	Cromatina madura i condensada	Cromatina cerebriforme	Granulació citoplasmàtica	Velositats citoplasmàtiques
Regió d'interès	Nucli		Citoplasma	Regió externa de la cèl·lula
Grup cel·lular (conjunt 2)	Correlació del cian del nucli ($\times 10^{-2}$)	DE de la corba granulomètrica del cian del nucli ($\times 10^{-3}$)	Asimetria de l'histograma del component u del citoplasma	Hairiness (μm^2)
Monòcit	75,6 (1,8)	36,9 (4,5)	0,41 (0,11)	1 (2)
Limfòcit	76,4 (3,4)	28,6 (3,6)	0,24 (0,16)	1,2 (0,9)
Limfòcit anormal de la leucèmia limfàtica crònica	79,9 (1,1)	29,5 (3,7)	0,18 (0,33)	0,5 (0,7)
Limfòcit vellós de la tricoleucèmia	77,7 (2,7)	30,1 (4,7)	0,08 (0,35)	6,4 (4,6)
Limfòcit vellós del limfoma de la zona marginal esplènica	77,5 (2,6)	35,2 (4,4)	0,20 (0,11)	21,2 (15,7)
Cèl·lula del limfoma del mantell morfologia blàstica	77,0 (2,6)	36,7 (6,6)	0,19 (0,34)	0,4 (1,2)
Cèl·lula del limfoma del mantell morfologia clàssica	77,1 (2,0)	34,9 (7,2)	0,23 (0,41)	0,6 (0,6)
Limfòcit anormal del limfoma fol·licular	75,7 (2,4)	28,4 (6,0)	0,02 (0,40)	0,1 (0,2)
Cèl·lula plasmàtica	73,5 (5,2)	29,8 (4,3)	0,29 (0,31)	1,6 (2,0)
Prolimfòcit B	74,9 (2,7)	34,0 (5,5)	0,11 (0,29)	2,3 (2,5)
Prolimfòcit T	74,2 (2,1)	24,6 (4,0)	0,14 (0,32)	0,6 (0,5)
Cèl·lula de Sézary variant Lutzner o petita	77,0 (3,3)	24,3 (4,1)	0,05 (0,34)	1,1 (1,7)
Cèl·lula de Sézary variant clàssica	75,5 (3,3)	21,7 (4,9)	0,10 (0,34)	2,4 (3,0)
Limfòcit gran granular	78,2 (1,5)	27,1 (4,3)	2,00 (2,16)	0,9 (1,1)
Cèl·lula limfoide reactiva	77,5 (2,1)	33,4 (3,6)	0,07 (0,12)	0,5 (0,6)
Blast limfoide	76,8 (1,4)	36,7 (4,5)	0,29 (0,50)	0,2 (0,3)

La textura està relacionada amb patrons espacials de color o intensitats, que es poden detectar visualment. A l'anàlisi digital d'imatges, la textura es defineix quantitativament per la uniformitat, densitat, to de píxel i les seves relacions espacials, entre d'altres. Aquesta anàlisi es realitza normalment mitjançant dues eines principals: 1) matriu de co-ocurrència de nivell de grisos (GLCM) i 2) granulometria.

La GLCM es defineix com la probabilitat que els parells de píxels veïns tinguin intensitats similars (88). El present estudi mostra que la correlació de la GLCM del component cian del nucli és un descriptor capaç de discriminar entre la distribució de la cromatina madura i condensada de les cèl·lules de la LLC i la distribució de cromatina no madura i condensada (no LLC) d'altres cèl·lules limfoides diferents de la LLC ($p < 0,0001$, Taula 5.5).

En concret, s'han utilitzat GLCM calculades en quatre direccions ($\theta = 0^{\circ}, 45^{\circ}, 90^{\circ}$ i 135°) i considerant el píxel adjacent ($d = 1$) amb vuit nivells d'intensitat. Per tal de que aquesta funció sigui invariant respecte a la rotació, hem fet la mitjana de les GLCM obtingudes a partir de les quatre direccions. Diversos paràmetres

estadístics es poden calcular a partir del GLCM, com la correlació o l'energia, tal i com hem vist al Capítol 3 i secció 3.4.3.2.

En el nostre cas, el descriptor és la correlació del component cian de la imatge original. La correlació és una mesura de dependència lineal de nivell gris entre els píxels en les posicions especificades respecte als seus veïns. Les cèl·lules de la LLC presenten un elevat nivell correlació dels seus píxels veïns, ja que els píxels d'alta intensitat i baixa estan molt a prop (corresponent als grumolls característics). A continuació mostrem un exemple de com es calcula aquest descriptor. La Figura 5.4 presenta (I) dues imatges de limfòcits anormals corresponents a les següents neoplàsies limfoides:

- limfoma fol·licular (LF, imatge B): correlació = 0,7539
- leucèmia limfàtica crònica (LLC, imatge A): correlació = 0,8266

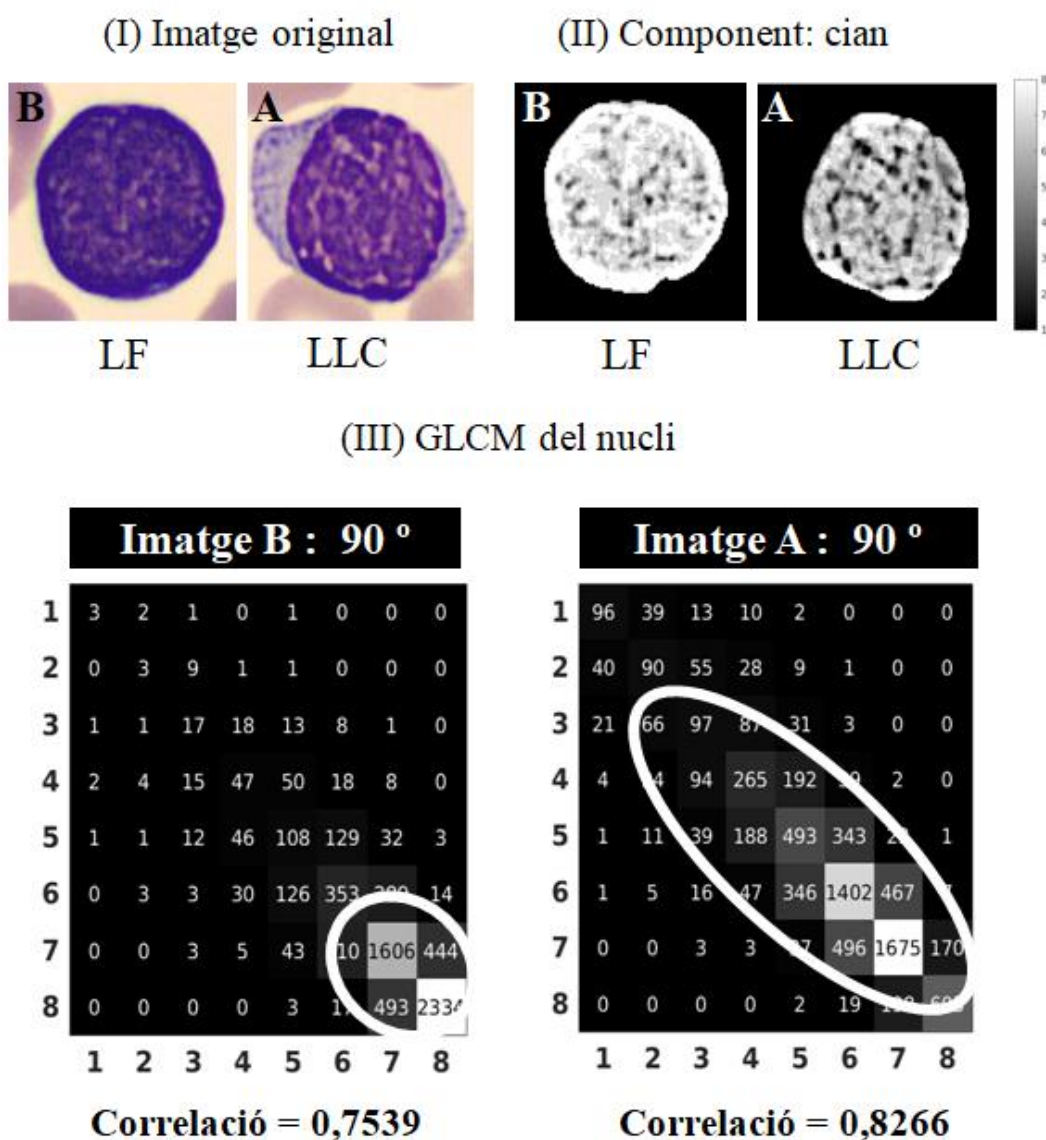


Figura 5.4: Anàlisi d'imatges de la correlació del component cian de la regió del nucli.

A continuació s'obtenen les corresponents imatges en escala de grisos per al component cian (II) i les corresponents GLCM calculades per a la regió del nucli per a una $d = 1$ (el píxel adjacent) i $\theta = 90^\circ$ (direcció oest) (III). La GLCM té una mida de 8×8 perquè els nivells d'intensitat a la imatge van des d'1 (negre) fins a 8 (blanc). A la GLCM de la imatge B, l'element de la posició (8,8) indica que el nivell 8 és adjacent al nivell 8 en direcció oest 2.334 vegades.

En el cas particular del limfòcit anormal de la LLC, la correlació dels píxels veïns del nucli, ja sigui en els nivells d'intensitat baixa i alta, és alta, ja que els píxels de poca intensitat es troben molt a prop (per exemple, 1-1, 2-2, 3-3) així com els d'alta intensitat (per exemple, 6-6, 7-7, 8-8). Els píxels de baixa intensitat es representen com les línies blanques observades al nucli i els d'alta intensitat corresponen als grumolls típics o "clumps", observats al nucli com píxels negres. Per contra, s'observa una baixa correlació quan un nucli conté més combinacions d'intensitats altes i baixes (per exemple, 2-7, 3-8, 1-6), ja que hi ha una gran variació en els nivells d'intensitat entre els píxels veïns (imatge B).

Les Figures 5.5, 5.7, 5.8 i 5.10 mostren els diagrames de caixes dels quatre descriptors. A simple vista, els diagrames de caixes destaquen com cada descriptor mostra un valor diferenciat per a la corresponent característica qualitativa i la cèl·lula diana.

A la Figura 5.5 es mostren els diagrames de caixes de la *correlació del component cian del nucli* per a tots els grups de 16 cèl·lules del conjunt 2 d'imatges i els valors quantitius corresponents es detallen a la Taula 5.6. La *correlació del component cian del nucli* a la LLC mostra diferències significatives respecte a tots els grups de cèl·lules en estudi ($p < 0,005$, Fig. 5.5).

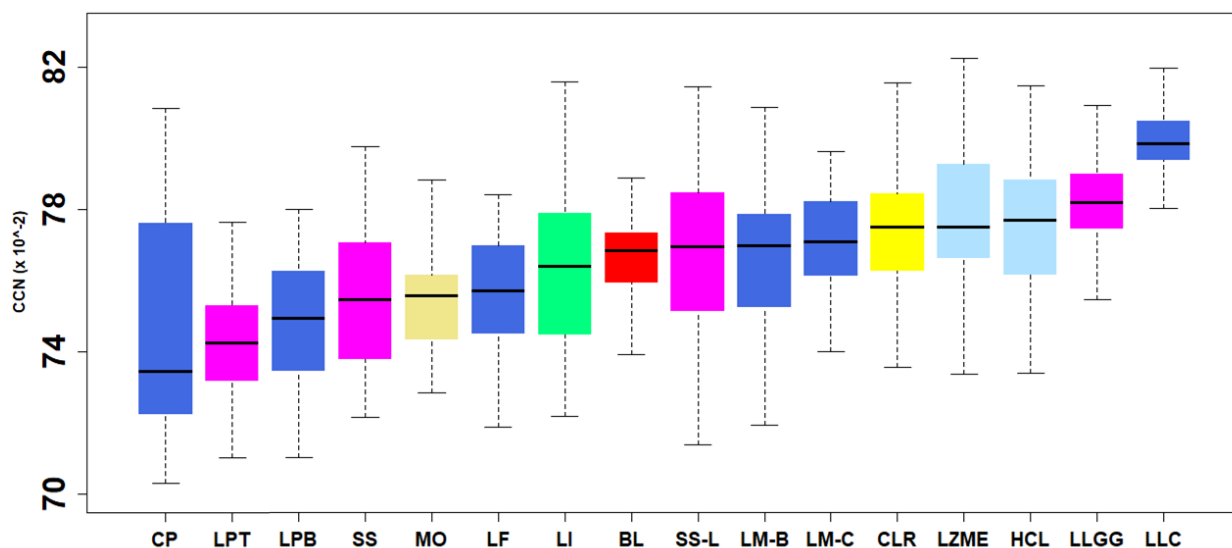


Figura 5.5: Valors de la *correlació del component cian del nucli (CCN)* calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.

D'altra banda, tal i com s'ha explicat al Capítol 3 i secció 3.4.3.3, la granulometria mesura la distribució de la grandària de partícules en una imatge mitjançant operacions de morfologia matemàtica, que es visualitzen mitjançant les anomenades corbes granulomètriques i pseudo-granulomètriques. Aquestes corbes representen la proporció de partícules negres i brillants a la imatge, i d'elles és possible calcular paràmetres, com la mitjana i DE, entre d'altres.

La Figura 5.6 il·lustra aspectes tècnics bàsics sobre com interpretar aquest descriptor granulomètric. En concret, mostra (I) dues imatges originals:

- Imatge 1: cèl·lula de Sézary amb un valor baix de *DE de la corba granulomètrica del component cian del nucli* (0,017)
- Imatge 2: cèl·lula limfoide reactiva amb un valor alt de *DE de la corba granulomètrica del component cian del nucli* (0,791)

A continuació s'obtenen els components cian d'ambdues imatges (II) i per últim, les corbes granulomètriques corresponents a la regió del nucli (III).

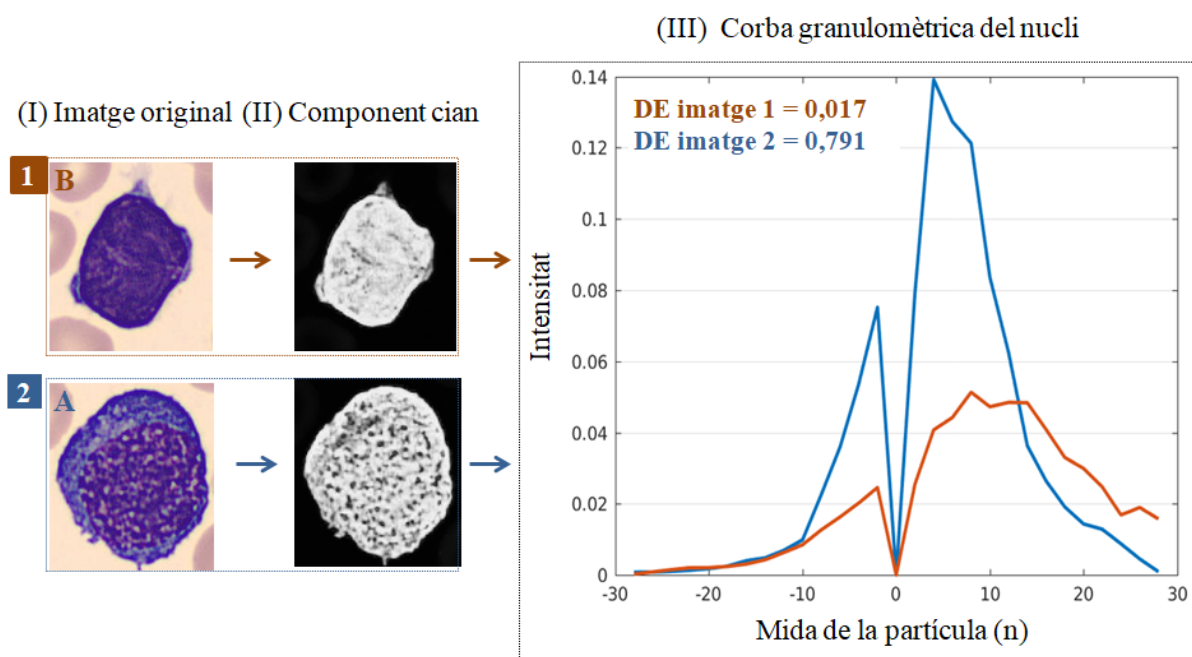


Figura 5.6: Anàlisi d'imatges de la desviació estàndard de la corba granulomètrica del component cian de la regió del nucli.

Com ja s'ha indicat al Capítol 3 i secció 3.4.3.3, la part esquerra de la corba granulomètrica (mida negativa) representa la mida dels grànuls foscos, mentre que la part dreta representa la dels grànuls brillants. Es pot veure que els nuclis cerebriformes tenen valors més baixos d'aquest descriptor en comparació amb els nuclis no cerebriformes. La corba granulomètrica d'una cèl·lula Sézary (vegeu la corba marró) no mostra un sol pic perquè les mides de les partícules que encaixen són de grandària variada. Per tant, hi ha molta més

variació en totes les mides de les partícules, la qual cosa fa que la corba sigui més ampla i més plana (menor desviació estàndard). Els plecs estan relacionats amb les mides de partícules més grans.

En canvi, en el nucli sense un patró de cromatina cerebriforme, tal com en un limfòcit reactiu (corba blava), la desviació estàndard és més alta, ja que el nucli mostra canvis més grans en la distribució de mida de les partícules, centrats en un rang de grandària més petita, on un sol pic és present. Una explicació més detallada sobre granulometria es pot trobar a (126).

En el cas de la present recerca, la *DE de la corba granulomètrica del component cian del nucli* és un descriptor de textura que considera el grau de dispersió de la distribució de les mides de les partícules. Aquest descriptor presenta diferències significatives quan es compara el nucli intensament irregular amb múltiples indentacions profundes (nucli cerebriforme) de les cèl·lules de Sézary (valors baixos del descriptor) respecte a les cèl·lules no cerebriformes ($p < 0,0001$, Taula 5.5). No només les cèl·lules Sézary clàssiques, sinó també la variant petita de les cèl·lules de Sézary, mostren valors significatius amb valors baixos d'aquest descriptor granulomètric en comparació amb tots els grups cel·lulars restants ($p < 0,05$, Taula 5.6 i Fig. 5.7), a excepció dels prolimfòcits T. És conegut pels citòlegs que aquestes cèl·lules mostren un nucli irregular amb plecs.

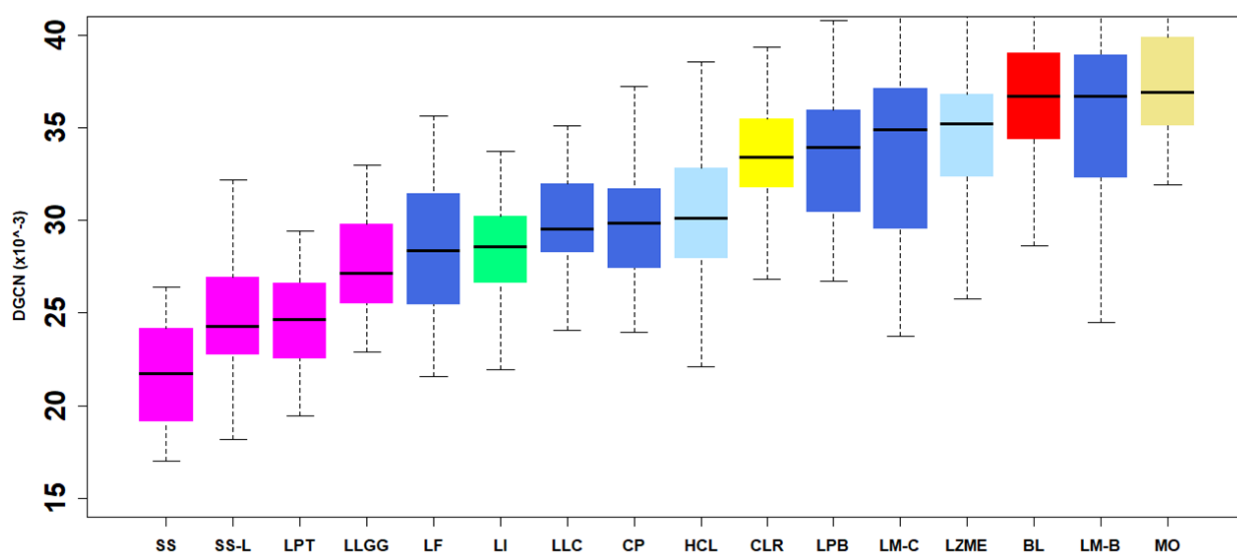


Figura 5.7: Valors de la desviació estàndard de la corba granulomètrica del component cian del nucli (DGCN) calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.

Pel que fa a les característiques citoplasmàtiques, hem trobat un descriptor de color per a la identificació dels grànuls citoplasmàtics azuròfils, generalment observats als limfòcits grans granulars. El color és una propietat física molt comú en la caracterització visual de les cèl·lules sanguínies.

Per tal d'explorar la rica quantitat d'informació present a les cèl·lules limfoides malignes, hem utilitzat diversos espais de color. Com hem detallat al Capítol 3 secció 3.4.2.1, alguns exemples d'espais de color són el RGB (vermell, verd i blau que corresponen als colors bàsics) o Luv (brillantor i cromaticitat), entre d'altres. Per obtenir descriptors quantitius de les imatges en escala de grisos, és molt útil representar l'histograma, un gràfic de dos eixos en la qual l'horitzontal es divideix en intervals que representen valors d'intensitat d'1 a 255. A partir de cada histograma, es calculen sis paràmetres estadístics, sent l'asimetria un d'ells.

En el present estudi s'ha demostrat que el descriptor de color "*asimetria de l'histograma del component u del citoplasma*" és útil per a la quantificació dels grànuls citoplasmàtics azuròfils, vistos com píxels violeta ($p < 0,0001$, Taula 5.5). Aquest descriptor es basa en el càlcul de l'asimetria de l'histograma del component u (espai de color Luv) obtingut a partir de la imatge en escala de grisos. A la Taula 5.6 i a la Figura 5.8, es pot veure que la leucèmia de limfòcits grans granulars mostra els valors més alts (2.00) quan es compara amb els grups cel·lulars restants ($p < 0.001$).

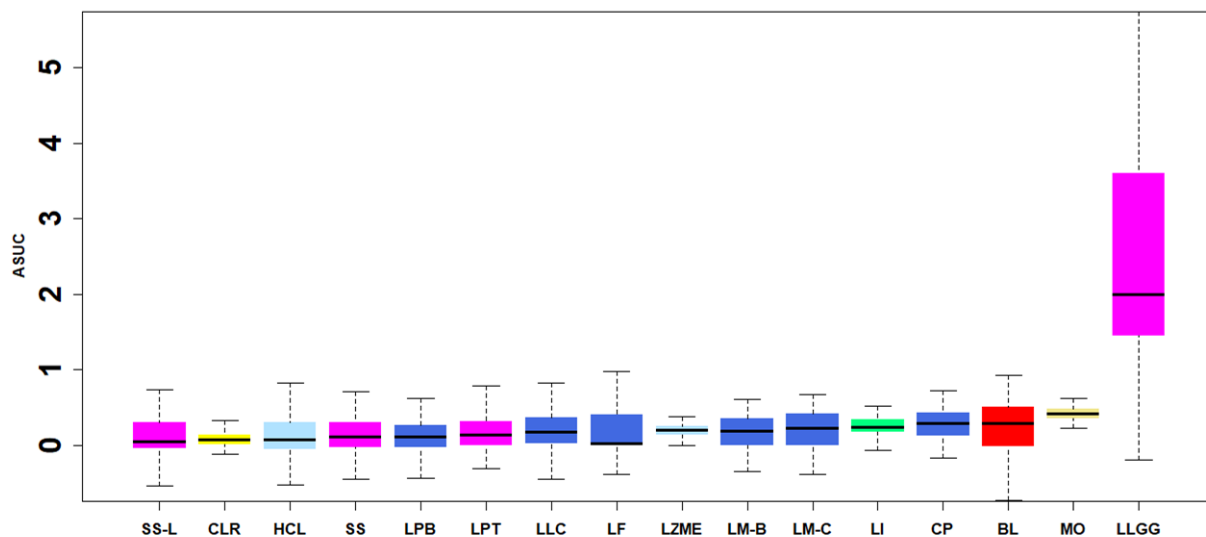


Figura 5.8: Valors de l'*asimetria de l'histograma del component u del citoplasma* (ASUC) calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.

La Figura 5.9 il·lustra aspectes tècnics bàsics sobre com interpretar aquest descriptor. En primer lloc, es mostren dues imatges originals d'una cèl·lula limfoide reactiva (CLR) i un limfòcit gran granular (LGG) amb els corresponents resultats de la segmentació de les regions d'interès (I). A continuació, s'obtenen els components de color u d'ambdues imatges (II) i per últim, els histogrames dels nivells d'intensitat del color u de la regió del citoplasma (III). El component de color u és una coordenada de la cromaticitat, mesura el color morat i a la imatge es correspon als grànuls.

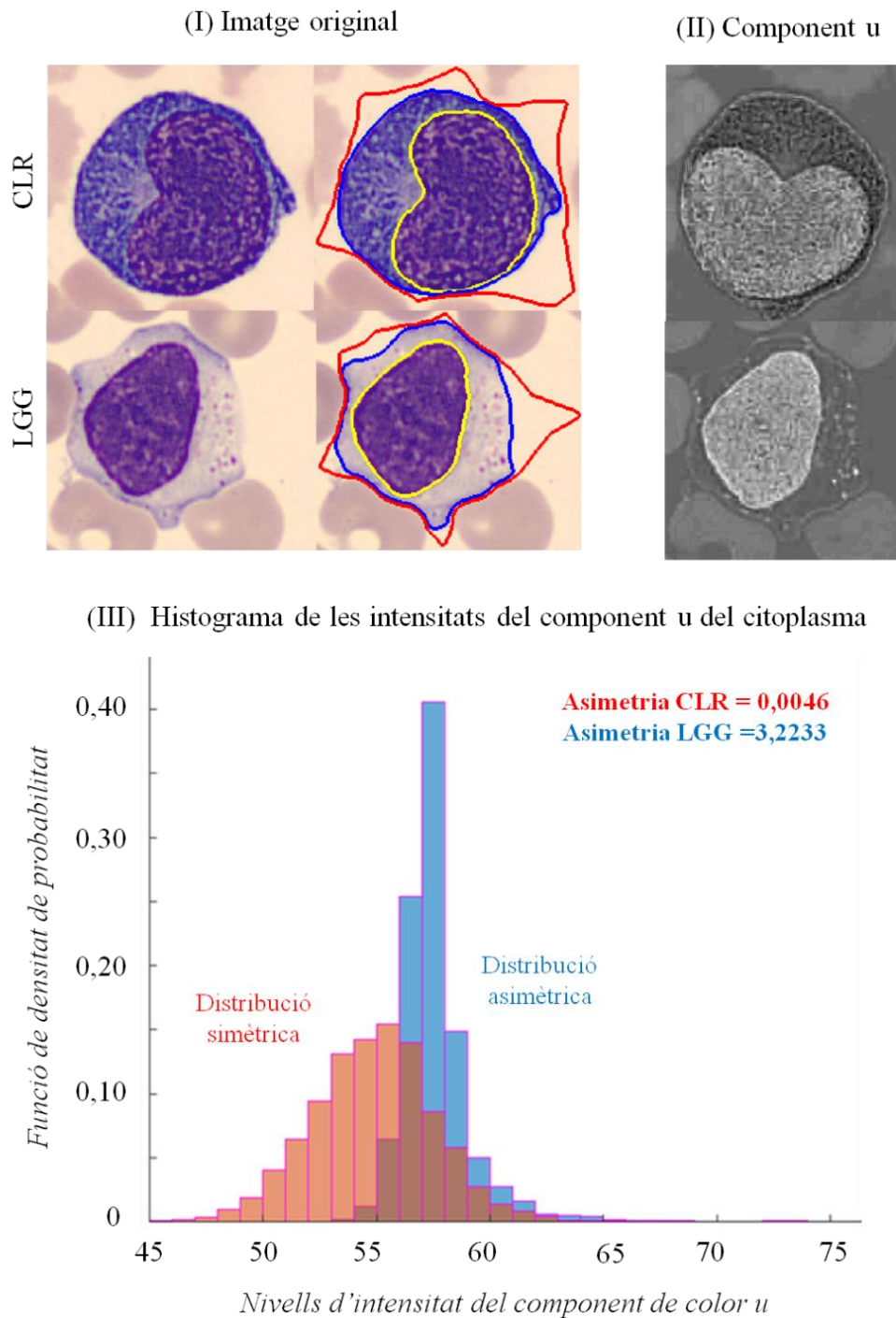


Figura 5.9: Anàlisi d'imatges de l'asimetria del component de color u de la regió del citoplasma.

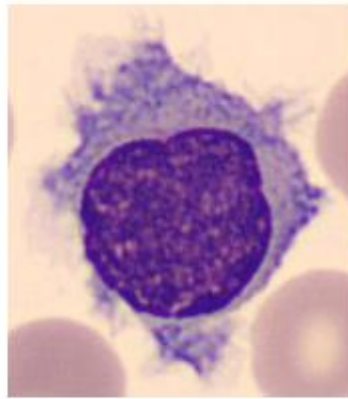
Com es pot veure, la CLR de la Figura 5.9 no mostra granulació al seu citoplasma i la distribució dels nivells d'intensitat dels píxels es troba restringida entre els de valors 46 i 64, obtenint una distribució simètrica amb un valor d'asimetria propera a zero (0,0046). En canvi, si ens fixem en el LGG, al tenir grànuls al seu citoplasma, aquests es corresponen amb els nivells d'intensitat més elevats, és a dir, els píxels més brillants. La presència d'aquests últims provoca que la distribució de l'histograma sigui asimètrica, deixant una cua a la dreta (valor d'asimetria positiu de 3,2233).

Pel que fa a la quantificació de les vellositats citoplasmàtiques exhibides per les cèl·lules limfoides de l'HCL i el LZME, en treballs anteriors del grup de recerca (29,59) es va demostrar la utilitat del descriptor “*hairiness*” per a la detecció de les vellositats citoplasmàtiques. De fet, aquest descriptor va ser desenvolupat per Alférez *et al.* (29) com a descriptor del perfil citoplasmàtic per determinar la presència o absència de vellositats citoplasmàtiques en els limfòcits vellosos de la tricoleucèmia, tal i com s'ha detallat al Capítol 3 secció 3.4.1.5.

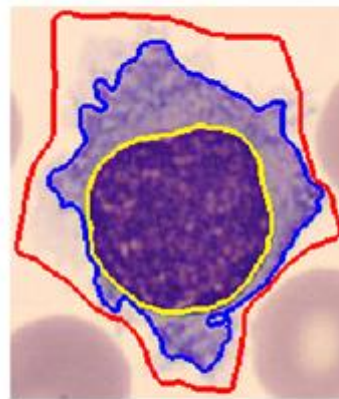
A la Figura 5.10 es mostra una imatge original d'un tricoleucòcit obtinguda amb el sistema DM96 (I), la seva segmentació (II) i l'histograma del component verd de la regió externa de la cèl·lula (III), que està delimitada entre el contorn extern (vermell) i el contorn de la cèl·lula (blau). Com es pot observar, l'histograma mostra una distribució bimodal: el pic de major intensitat de verd correspon al fons de la regió externa (valors d'intensitat de verd entre 223 i 230), mentre que l'altre pic situat al voltant de 210, correspon a les projeccions citoplasmàtiques. Una cèl·lula sense projeccions citoplasmàtiques mostra una distribució unimodal. La suma de píxels de les intensitats entre 70 i 215 (llindars establerts per S. Alférez pel sistema DM96) és de 26,57 μm^2 o 2.693 píxels.

En la present recerca es demostra que el *hairiness* exhibeix una gran especificitat per als limfòcits vellosos de l'HCL i el LZME, ja que aquests dos grups presenten els valors més alts (6,4 i 21,2 μm^2 , respectivament) i mostren diferències significatives en comparació amb els restants grups cel·lulars analitzats ($p < 0,0001$, Taula 5.6 i Fig. 5.11).

I) Imatge original DM96



II) Segmentació



III) Histograma del component verd de la regió externa de la cèl·lula

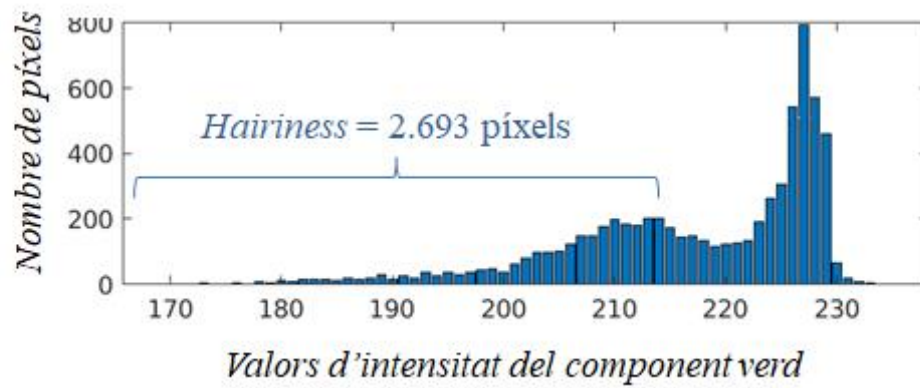


Figura 5.10: Càlcul del descriptor *hairiness* per a un tricoleucòcit.

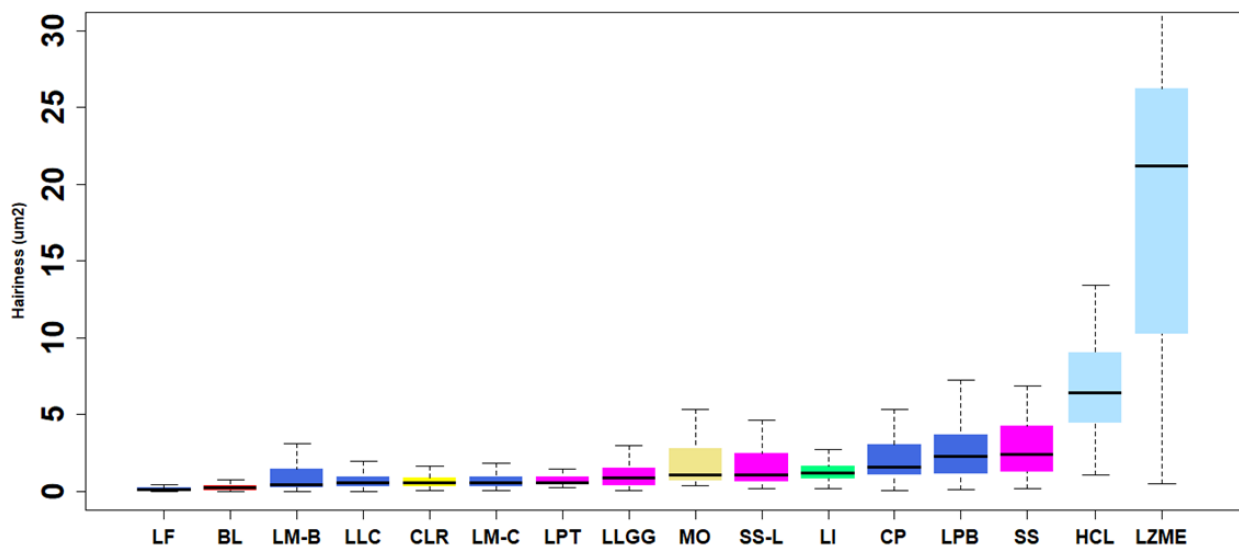


Figura 5.11: Valors del *hairiness* calculats per a tots els grups de cèl·lules del conjunt 2 d'imatges.

5.3.2 Avaluació dels descriptors quantitius per a la classificació de frotis

En aquesta segona part de l'estudi, per a cada descriptor avaluat, es detallen els pacients inclosos, els valors de tall discriminador obtinguts així com el rendiment de la classificació dels frotis. Els resultats d'aquesta segona part es resumeixen a les Taules 5.7 i 5.8 i a les Figures 5.12 – 5.15.

A la Taula 5.7 es mostren els valors de tall obtinguts a partir de l'anàlisi ROC sobre el primer conjunt d'imatges.

Taula 5.7: Valors de tall (cut-off), àrea sota la corba (AUC), sensibilitat (Sn) i especificitat (Sp) obtinguts per a cada descriptor mitjançant l'anàlisi ROC.

Característica citològica	Descriptor	Cut-off	AUC	Sn	Sp
Cromatina madura i condensada	<i>Correlació del cian del nucli</i>	78×10^{-2}	0,850	83,0 %	73,4 %
Cromatina cerebriforme	<i>DE de la corba granulomètrica del cian del nucli</i>	$26,1 \times 10^{-3}$	0,969	94,0 %	87,6 %
Granulació citoplasmàtica	<i>Asimetria de l'histograma del component u del citoplasma</i>	1,07	0,999	98,0 %	100 %
Velositats citoplasmàtiques	<i>Hairiness</i>	$3,28 \mu\text{m}^2$ (332 píxels)	0,856	76,6 %	100 %

A la Taula 5.8 es mostra, per cada descriptor, els seus valors calculats de les diferents imatges incloses així com el percentatge d'imatges positives.

Taula 5.8: Nombre total d'imatges corresponents a cada diagnòstic, mediana, rang interquartílic (RIQ) i imatges positives per a cada descriptor quantitatiu específic obtingudes a l'avaluació dels frotis.

Descriptor cromatina madura i condensada ($\times 10^{-2}$)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges LLC	2.277	79,39	1,99	1.725 (76 %)
Imatges no LLC	1.557	76,17	3,54	-
Descriptor cromatina cerebriforme ($\times 10^{-3}$)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges SS	125	20,96	4,75	124 (99 %)
Imatges no SS	840	27,57	2,10	-
Descriptor granulació citoplasmàtica				
	Imatges	Mediana	RIQ	Imatges positives
Imatges LLGG	292	1,27	1,28	199 (68 %)
Imatges no LLGG	1.173	0,16	0,32	-
Descriptor velositats citoplasmàtiques (μm^2)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges HCL/LZME	387	10,06	9,78	352 (91 %)
Imatges no HCL/LZME	1.572	1,09	1,99	-

HCL, tricoleucèmia; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulats; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

El descriptor de la *correlació del component cian del nucli* ha estat capaç de diferenciar la cromatina madura i condensada i s'ha calculat en frotis de dos grups diferents de nous pacients. El primer està format per 95 pacients amb LLC mentre que el segon inclou 52 pacients, els frotis dels quals contenen limfòcits amb cromatina madura però no condensada (14 LF, 4 SS i 34 individus sans).

El valor de tall òptim obtingut ha estat de 78×10^{-2} (veure Taula 5.7), entenent que una cèl·lula limfoide serà considerada com que conté cromatina madura i condensada quan el valor de *correlació del component cian del nucli* sigui superior a 78×10^{-2} . D'altra banda, l'anàlisi estadística ha proporcionat un llindar del 45 %, de manera que un frotis és considerat positiu per a cromatina madura i condensada de LLC quan el rati entre el número d'imatges amb un valor de *correlació del component cian del nucli* superior al *cut-off* i el número total d'imatges del frotis sigui superior a 0,45.

Un total de 89/95 (93,7%) de frotis de LLC s'han classificat a la classe correcta, així com en 49/52 (94,2%) dels frotis del segon grup (limfòcits sense cromatina madura i condensada). A la Taula 5.8 es mostren els resultats desglossats per imatges i s'observa que el 76 % d'imatges de limfòcits LLC mostren un valor del descriptor superior a 78×10^{-2} .

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 5.12:

- Sensibilitat: 92,6 %
- Especificitat: 92,3 %
- VPP: 95,7 %
- VPN: 87,3 %
- Exactitud: 92,5 %

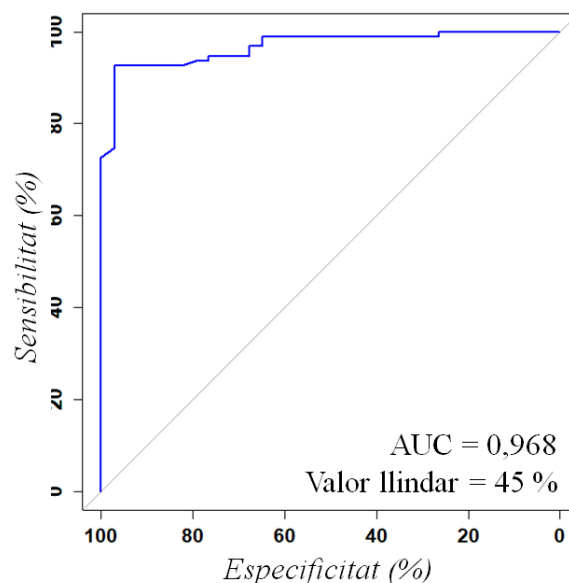


Figura 5.12: Corba ROC de la *correlació del component cian del nucli*.

La *desviació estàndard (DE) de la corba granulomètrica del component cian del nucli* és un descriptor de textura que com hem vist anteriorment, ha demostrat ser útil en la identificació de la cromatina cerebriforme. A l'avaluació dels frotis, aquest descriptor s'ha calculat en frotis d'un primer conjunt format per dos nous pacients amb SS i un segon conjunt format per limfòcits normals. S'han escollit els limfòcits normals perquè, a l'igual que les cèl·lules de Sézary, també mostren una cromatina madura.

El valor de tall òptim obtingut ha estat de $26,1 \times 10^{-3}$ (veure Taula 5.7), el que significa que un nucli cel·lular amb un valor de *DE de la corba granulomètrica del component cian del nucli* inferior, contindrà cromatina cerebriforme. En aquest cas, el llindar que ha proporcionat l'anàlisi estadística ha estat del 78 %. Per tant, un frotis es considera positiu per cromatina cerebriforme quan el rati entre el número d'imatges amb el descriptor per sota de $26,1 \times 10^{-3}$ i el número total d'imatges del frotis sigui major a 0,78.

L'avaluació ha mostrat que tots els frotis que contien imatges de cèl·lules limfoides amb cromatina cerebriforme s'han classificat correctament com que pertanyien a pacients amb SS. A més a més, tots els frotis que contien limfòcits normals han estat classificats a la categoria veritable. A la Taula 5.8 es mostren els resultats desglossats per imatges i s'observa que el 99 % d'imatges de cèl·lules de Sézary mostren un valor del descriptor inferior a $26,1 \times 10^{-3}$.

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 5.13:

- Sensibilitat: 100%
- Especificitat: 100%
- VPP: 100%
- VPN: 100%
- Exactitud: 100%

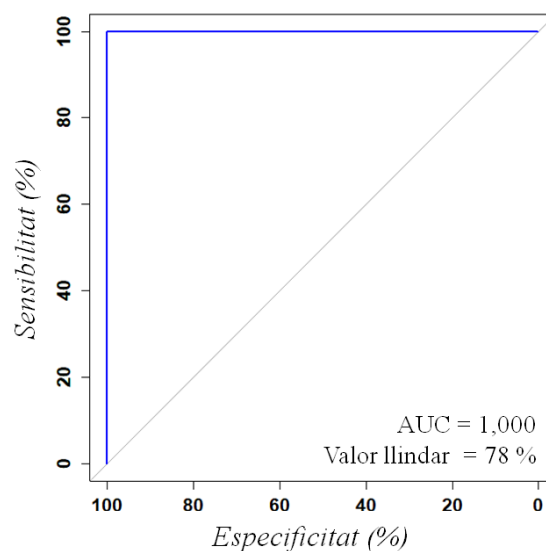


Figura 5.13: Corba ROC de la *DE de la corba granulomètrica del component cian del nucli*.

L'asimetria de l'histograma del component u del citoplasma és capaç de detectar els grànuls citoplasmàtics de color lilós i ha estat avaluat en frotis d'un primer grup de 10 nous pacients amb LLGG i d'un segon grup de 45 pacients, dels quals 11 mostraven limfòcits vellosos (2 HCL i 9 LZME) i els restants limfòcits normals (34 individus sans).

El valor de tall òptim obtingut ha estat d' 1,07 (veure Taula 5.7), el que significa que un citoplasma amb un valor superior d'asimetria de l'histograma del component u del citoplasma, contindrà grànuls al seu citoplasma. Per tant, un frotis es considera positiu per granulació citoplasmàtica quan el rati entre el número d'imatges amb el descriptor per sobre d' 1,07 i el número total d'imatges del frotis sigui major a 0,30.

Un total de 10/10 frotis que contenien limfòcits grans granulars s'han classificat com a pacients amb LLGG. A més a més, totes les imatges dels 45 frotis corresponents a l'altre grup de pacients amb múltiples patologies, han estat correctament classificades com a no LLGG. En relació a les imatges, el 68 % d'imatges de limfòcits grans granulars mostren un valor del descriptor superior a 1,07.

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 5.14:

- Sensibilitat: 100 %
- Especificitat: 100 %
- VPP: 100 %
- VPN: 100 %
- Exactitud: 100 %

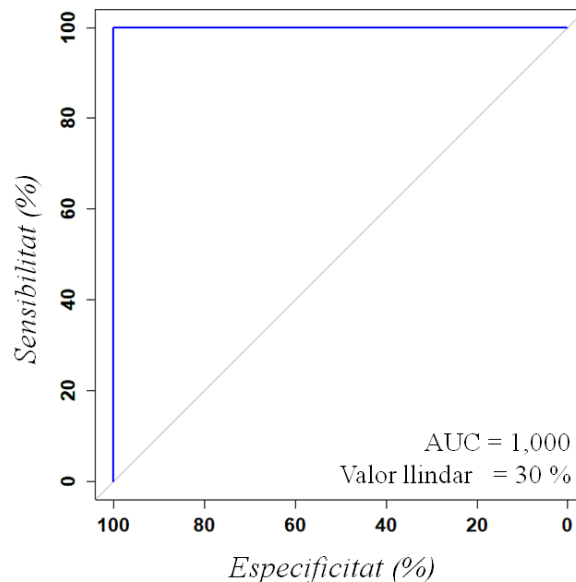


Figura 5.14: Corba ROC de l'asimetria de l'histograma del component u del citoplasma.

El descriptor *hairiness* s'ha calculat en frotis de SP que pertanyien a un primer grup de 13 pacients nous, en els quals el diagnòstic era HCL (2) o LZME (11). El segon grup de frotis, format per 63 pacients, contenien limfòcits reactius (18 pacients), limfòcits grans granulars (11 pacients amb LLGG) i limfòcits normals (34 individus sans) perquè tenen una mida similar a la dels limfòcits vellosos.

El valor de tall òptim obtingut per diferenciar ambdues categories (vellós i no vellós) és de $3,28 \mu\text{m}^2$ (o 332 píxels) (veure Taula 5.7), el que significa que cèl·lules amb un valor de *hairiness* per sobre, són considerades com a LV. Un frotis es considera positiu per LV prenent com a valor llindar el 62,8 % de les imatges reconegudes correctament en la categoria veritable.

Utilitzant aquest descriptor, un total de 12 frotis que contenen LV s'han classificat a la classe correcta (92,3 %). A més a més, els valors de *hairiness* en les imatges d'un total de 62/63 (98,4%) frotis del segon grup sense contenir limfòcits vellosos, han estat inferiors al *cut-off*. En relació a les imatges, el 91 % d'imatges de LV mostren un valor del descriptor superior a $3,28 \mu\text{m}^2$ (o 332 píxels).

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 5.15:

- Sensibilitat: 93,3 %
- Especificitat: 100 %
- VPP: 100 %
- VPN: 98,5 %
- Exactitud: 98,7%

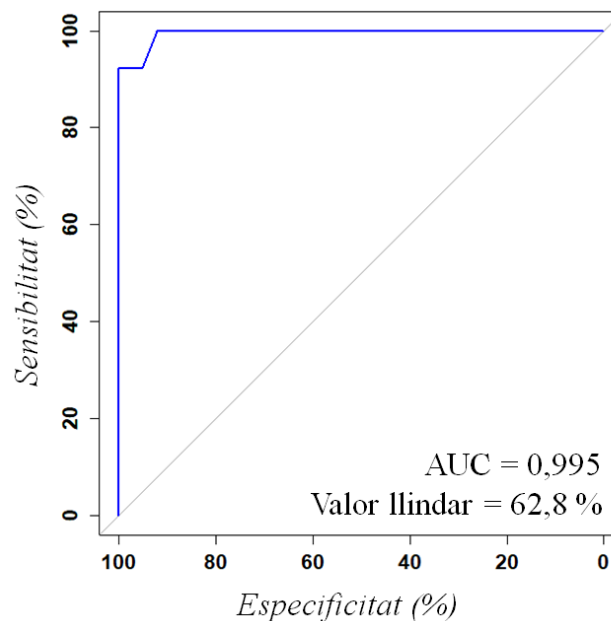


Figura 5.15: Corba ROC del *hairiness*.

5.4 Discussió

Freqüentment el frotis de sang perifèrica proporciona les primeres evidències morfològiques de diagnòstic, per tant, encara segueix sent una important eina de diagnòstic (39). Una mesura reproducible, precisa i objectiva de la morfologia cel·lular és beneficiosa per evitar variacions subjectives i entre diferents observadors (131). Els paràmetres citològics de cèl·lules individuals constitueixen un objectiu factible per a ésser mesurat de forma quantitativa mitjançant l'anàlisi d'imatges (2), però actualment manquen escales numèriques objectives.

La recerca presentada en aquest Capítol 5 va un pas més enllà en comparació amb el nostre estudi anterior (126) on analitzàvem els 20 descriptors més rellevants obtinguts per un procediment de *machine learning* per a la diferenciació morfològica de 12 grups de cèl·lules limfoides. Aquest treball també utilitza *machine learning* per tal d'obtenir una gran quantitat de descriptors potencialment útils. Tot i així, en comptes de realitzar una anàlisi *a posteriori*, en el treball presentat en aquest Capítol 5 s'ha buscat específicament els millors descriptors per a la detecció d'anomalies específiques de cèl·lules limfoides.

En aquest estudi s'ha utilitzat l'anàlisi d'imatges per aconseguir una avaluació morfològica objectiva de cèl·lules limfoides anormals de SP. S'han inclòs monòcits i 15 cèl·lules limfoides diferents, juntament amb quatre característiques citològiques qualitatives per a trobar descriptors quantitius i assignar-los valors numèrics. En una segona etapa, s'ha realitzat una avaluació de la capacitat discriminatòria dels quatre descriptors que són específics d'un determinat tipus cel·lular limfoide utilitzant imatges de sang perifèrica obtingudes a partir de frotis individuals de nous pacients.

En aquest treball s'han extret un total de 2.678 descriptors de color i textura. Fins al nostre coneixement, aquest elevat nombre i la diversitat de descriptors extrets d'imatges de limfòcits anormals encara no s'ha realitzat a la literatura. La majoria dels estudis recents s'han centrat en les cèl·lules blàstiques mieloides i limfoides (8,17,132) i hi ha un buit a la literatura pel que fa a cèl·lules limfoides anormals.

Els pocs estudis realitzats en neoplàsies limfoides han considerat només un nombre reduït i en la majoria només descriptors geomètrics han estat implicats. Greenspan *et al.* (133) utilitza tècniques d'imatges espectrals aplicades al nucli per a diferenciar cèl·lules limfoides reactives de limfòcits de la LLC i d'altres neoplàsies limfoides. Tanmateix, aquest treball només inclou 10 imatges per grup i les diferents neoplàsies limfoides involucrades no s'especifiquen. Jahanmehr *et al.* (2) calcula l'àrea cel·lular, el diàmetre cel·lular, l'àrea del citoplasma, l'àrea nuclear, la relació nucli/citoplasma i la densitat nuclear, però l'estudi només inclou tres neoplàsies limfoides: LLC, LCM i LPB.

L'aportació de Benattar *et al.* (31) representa el primer intent de mesurar algunes característiques morfològiques comunes d'una manera estandaritzada per a proporcionar criteris objectius per caracteritzar malalties i ser útils per al diagnòstic. A l'estudi es proposa un sistema de puntuació per distingir neoplàsies limfoides de cèl·lules B a partir de set paràmetres cel·lulars: forma nuclear, àrea cel·lular i forma, rati

nucli/citoplasma, rati nuclear vermell/blau, rati citoplasmàtic verd/blau i detecció de nuclèol. El treball inclou 87 frotis de SP de neoplàsies limfoides de cèl·lules B i T i sis individus sans. Per a cada frotis de SP, s'analitzen 30 imatges i demostren la utilitat del sistema de puntuació, tot i que cal confirmar-lo amb un mostreig més gran.

Angulo *et al.* (9,10) extreu diversos paràmetres quantitius de limfòcits, com ara mides nuclears i cel·lulars, rati nucli/citoplasma, excentricitat nuclear, densitat de la cromatina, forma nuclear, nuclèol, basofília, granulació i forma citoplasmàtica. L'estudi considera limfòcits anormals de sis neoplàsies limfoides i la quantitat d'imatges utilitzades no s'especifica. D'altra banda el treball no es va completar amb estudis posteriors.

Amb el mateix propòsit de quantificació de components cel·lulars, també s'han utilitzat els paràmetres dels analitzadors automàtics d'Hematologia. Silva *et al.* (134) avaluen els índexs de volum limfocitari i la conductivitat en un analitzador d'Hematologia en malalties virals, LLC i altres limfòcits anormals. L'estudi conclou que aquests índexs podrien ser útils per a la diferenciació de cèl·lules de la LLC, de CLR o limfòcits anormals d'altres neoplàsies limfoides.

El present estudi ha inclòs un conjunt d'imatges més gran en comparació amb publicacions anteriors i ha calculat quatre descriptors quantitius, incloent no només descriptors geomètrics sinó també de color i textura, ja que el patró de densitat de la cromatina nuclear és una característica molt important per al diagnòstic morfològic i difícil de quantificar. A més a més, s'han caracteritzat 15 cèl·lules limfoides diferents (13 de les quals són malignes) i s'han presentat descriptors específics i sensibles per a determinades característiques nuclears i citoplasmàtiques.

La distribució de la cromatina reflexa l'organització nuclear de l'ADN, la qual conté informació important sobre el diagnòstic i el pronòstic. Aquest treball ha destacat que la caracterització de la distribució de la cromatina dels nuclis cel·lulars mitjançant descriptors quantitius és molt rellevant per al diagnòstic morfològic. Aquestes mesures reflecteixen com la cromatina es compacta al nucli cel·lular amb diferents variacions relacionades amb la malaltia neoplàsica subjacent.

Des de la seva introducció per part de Haralick (88), la matriu de co-ocurrència del nivell de gris (GLCM) ha estat àmpliament utilitzada com a mesura de textura en imatges mèdiques, com ara imatges d'ultrasons en neoplàsies sòlides (15) i imatges de medul·la òssia per distingir quatre etapes en les cèl·lules precursors d'eritròcits (16). No obstant això, pocs estudis relacionats amb la GLCM s'han realitzat utilitzant imatges de SP i principalment els que s'han dut a terme només han diferenciat entre els leucòcits normals (14,18) i les cèl·lules blàstiques limfoides (17). Sabino *et al.* (14) calcula cinc atributs texturals basats en la GLCM (energia, entropia, entre altres) per diferenciar entre els leucòcits normals (cinc subtipus) i les cèl·lules limfoides anormals de la LLC. Recentment, Kono *et al.* (18) ha aplicat la GLCM per a diferenciar sis leucòcits de SP, incloent-hi les bandes.

L'estudi que s'ha presentat en aquets Capítol ha aplicat la GLCM per analitzar les característiques de textura de 13 cèl·lules neoplàsiques que circulen a SP. En particular, la cromatina madura i condensada s'ha quantificat a partir d'un descriptor basat en la GLCM. La *correlació del cian del nucli* ha mostrat ser específica i sensible per a la mesura de la cromatina madura i condensada de les cèl·lules de la LLC. En la mesura del nostre coneixement, no existeixen estudis previs que hagin utilitzat les probabilitats de co-ocurrència per millorar la diferenciació de 16 cèl·lules limfoides, incloent 12 limfòcits anormals, cèl·lules blàstiques limfoides i cèl·lules limfoides reactives.

La granulometria també ha estat utilitzada a la literatura per a l'anàlisi de la textura de limfòcits anormals (3,10). La metodologia per analitzar automàticament la textura de la cromatina proposada per Angulo *et al.* (10) es basa en funcions granulomètriques calculades sobre el component de la luminància. En relació amb la cromatina cerebriforme de les cèl·lules de Sézary, l'estudi presentat en aquest Capítol 5 demostra que la *DE de la corba granulomètrica del component cian del nucli*, basada en la granulometria, exhibeix una gran especificitat per a les cèl·lules de Sézary i pot proporcionar una mesura per avaluar la textura nuclear cerebriforme. Fins al coneixement de l'autor, no s'han publicat prèviament descriptors quantitius per a detectar patrons de cromatina madura i condensada i cerebriforme.

L'avaluació de la limfocitosi hauria d'incloure la presència d'inclusions com vacuoles o grànuls (135). Els descriptors de color es troben implicats en la quantificació de la basofília citoplasmàtica i la granulació citoplasmàtica. L'*asimetria de l'histograma del component u del citoplasma* ha demostrat ser un descriptor de color específic dels grànuls azuròfils de les cèl·lules limfoides anormals de la LLGG. Angulo *et al.* (10) mostra resultats similars però utilitzant el component verd, encara que no defineix valors llindars. A més a més, el descriptor *hairiness* ha mostrat ser útil per a la descripció de les projeccions citoplasmàtiques.

L'avaluació dels descriptors quantitius en aquest estudi s'ha realitzat de manera que s'analitzen frotis individuals per a ser classificats en funció del valor del descriptor. Aquesta anàlisi és factible a la pràctica mitjançant un procediment seqüencial algorítmic implementat en un ordinador. Per a una determinada imatge cel·lular o un frotis de sang que conté un conjunt d'imatges, calen els següents passos:

- (1) segmentació de les regions d'interès de la cèl·lula: nucli, citoplasma i regió externa
- (2) càlcul del descriptor quantitiu i comparació amb els valors de referència
- (3) decisió sobre la classificació de la cèl·lula individual o la totalitat del frotis

La principal contribució d'aquest treball és la identificació i anàlisi de quatre nous descriptors que quantifiquen les característiques morfològiques específiques de les anomalies nuclears i citoplasmàtiques relacionades amb patrons de cromatina i granulació citoplasmàtica o vellositats que són indicadors de neoplàsies limfoides específiques. Aquestes característiques poden ser útils no només en l'etapa de diagnòstic, sinó també en la detecció de limfòcits anormals residuals en pacients que han completat el seu tractament. Cal esmentar que la *correlació del cian del nucli* i l'*asimetria de l'histograma del component u*

del citoplasma van aparèixer dins dels 140 descriptors més rellevants per a la diferenciació morfològica dels 12 grups de limfoides a l'estudi detallat al Capítol 4 (126).

Dos dels descriptors proposats en aquest treball impliquen l'ús de dues eines matemàtiques com la corba granulomètrica i la GLCM. Pot ser que siguin termes poc familiars pels patòlegs clínics, però admeten un cert grau d'interpretació morfològica com es discuteix a (127,128). La troballa de descriptors quantitius per a les característiques qualitatives morfològiques mitjançant l'anàlisi d'imatges pot proporcionar informació rellevant sobre propietats de les cèl·lules sanguínies anormals. Els descriptors presentats en aquest estudi podrien incorporar-se als nous analitzadors hematològics basats en morfologia automàtica d'imatges digitals de frotis de SP (136,137). Similar als índexs proporcionats pels analitzadors hematològics automatitzats que indiquen canvis en la mida dels eritròcits, es podrien introduir altres mesures específiques de cèl·lules limfoides anormals, com els descriptors presentats en aquest estudi, en aquests nous analitzadors per optimitzar la morfologia a través de l'anàlisi d'imatges de la sang (128).

Una limitació de la metodologia presentada en aquest estudi està relacionada amb la tinció del frotis de sang perifèrica. En general, les tecnologies basades en l'anàlisi d'imatges requereixen una estandardització del procediment de tinció per tal de garantir un frotis d'alta qualitat i minimitzar la diferència en les imatges adquirides de diferents fonts.

En resum, la identificació de les cèl·lules limfoides anormals considerades en aquest estudi és factible mitjançant descriptors quantitius obtinguts mitjançant l'anàlisi d'imatges. L'aplicabilitat dels descriptors descrits aquí pot millorar la reproductibilitat i l'objectivitat de l'avaluació morfològica de les cèl·lules limfoides, contribuint al diagnòstic en una etapa precoç, així com la seva detecció després del tractament corresponent.

5.5 Conclusions

A diferència de l'objectiu del Capítol 4 on es pretenia analitzar i interpretar els descriptors més rellevants per a la classificació de 12 cèl·lules limfoides, el present Capítol 5 aborda el problema d'obtenir descriptors quantitius específics en un entorn on el punt de partida són les anomalies cel·lulars específiques donades *a priori*, que estan associades a cèl·lules i malalties específiques. En aquest sentit, l'objectiu d'aquest Capítol 5 és buscar descriptors per a identificar als limfòcits anormals que circulen a sang perifèrica a la leucèmia limfàtica crònica, a la tricoleucèmia, al limfoma de la zona marginal esplènica, a la síndrome de Sézary i a la leucèmia de limfòcits grans granulars. Ambdós capítols han utilitzat la mateixa font d'adquisició d'imatges (sistema DM96) i la mateixa metodologia d'extracció dels descriptors.

Un total de quatre descriptors (de 2.678 analitzats) s'han proposat per mesurar quatre característiques citològiques: cromatina madura i condensada, cromatina cerebriforme, granulació citoplasmàtica i vellositats citoplasmàtiques i s'han establert escales numèriques per la seva quantificació.

La detecció i quantificació de la cromatina madura i condensada característica de les cèl·lules de la leucèmia limfàtica crònica ha estat possible mitjançant un descriptor basat en la matriu de co-ocurrència del nivell de gris, la *correlació del cian del nucli*. Aquest descriptor ha mostrat ser específic (especificitat del 92,3 %) i sensible (sensibilitat del 92,6 %) per a la detecció dels limfòcits anormals de la leucèmia limfàtica crònica.

S'ha detectat i quantificat la cromatina cerebriforme característica de les cèl·lules de Sézary (tant la variant clàssica com la Lutzner) mitjançant un descriptor granulomètric: la *desviació estàndard de la corba granulomètrica del component cian del nucli*, la qual exhibeix valors d'especificitat i sensibilitat del 100 %.

El descriptor estadístic de primer ordre denominat *asimetria de l'histograma del component u del citoplasma* ha demostrat ser útil per detectar els grànuls azuròfils característics de les cèl·lules limfoides anormals de la leucèmia de limfòcits grans granulars. La sensibilitat i l'especificitat ha estat del 100 %.

S'ha demostrat que el descriptor *hairiness* permet quantificar el perfil del citoplasma de les cèl·lules limfoides velloses de la tricoleucèmia i del limfoma de la zona marginal esplènica. La sensibilitat ha estat del 93,3 % i l'especificitat del 100 %.

S'ha demostrat que els quatre descriptors obtinguts per a la detecció de característiques morfològiques específiques de les mencionades cèl·lules limfoides neoplàsiques B i T anteriorment, són eficients (exactitud superior al 90 %) per al reconeixement d'aquestes cèl·lules en el frotis sanguini, pel que podrien contribuir al diagnòstic d'aquestes entitats. Dels 210 frotis de nous pacients utilitzats en l'avaluació, el 96 % s'han reconegut correctament.

Es conclou, per tant, que una avaluació basada en la morfologia més objectiva de l'anàlisi de limfòcits anormals a través de l'anàlisi d'imatges és factible amb descriptors quantitius. L'aplicabilitat dels descriptors descrits aquí pot millorar la reproductibilitat i l'objectivitat de l'avaluació morfològica dels limfòcits anormals gràcies al descobriment de descriptors específics de la granulació citoplasmàtica o de l'organització de la cromatina.

Amb l'objectiu d'analitzar si els descriptors trobats i descrits al Capítol 5 poden ser utilitzats per quantificar les mateixes característiques citològiques quan les imatges són adquirides amb un microscopi convencional, al Capítol 7 es calculen aquests quatre descriptors utilitzant 9.000 imatges adquirides amb una altra font (sistema BX43) i a l'hora se n'avalua la seva eficiència diagnòstica.

Capítol 6

6 Descriptors quantitius per a la diferenciació morfològica de cèl·lules limfoides anormals a partir d'imatges de sang perifèrica adquirides amb un microscopi convencional (sistema BX43)

Resum

L'objectiu d'aquest Capítol 6 és reproduir els resultats obtinguts al Capítol 4 amb una altra font d'adquisició d'imatges, com pot ésser un microscopi convencional, per tal de demostrar que els descriptors quantitius obtinguts mitjançant l'anàlisi d'imatges poden discriminar entre les cèl·lules limfoides normals, reactives, blàstiques i anormals de sang perifèrica independentment del sistema d'adquisició d'imatges que s'utilitzi. Per a dur-ho a terme, s'han utilitzat 8.463 imatges digitals adquirides manualment amb el microscopi Olympus® BX43 i la càmera digital DP73. Les imatges pertanyen a 12 grups cel·lulars limfoides corresponents a 192 pacients i inclouen nou neoplàsies limfoides: leucèmia limfàtica crònica, leucèmia prolimfocítica tipus B, tricoleucèmia, limfoma de la zona marginal esplènic, limfoma de cèl·lules del mantell, limfoma fol·licular, leucèmia prolimfocítica tipus T, leucèmia de limfòcits grans granulars tipus T i la síndrome de Sézary. Posteriorment, per a cada imatge s'han extret els 20 descriptors presentats al Capítol 4. Els valors obtinguts dels descriptors amb el sistema BX43 es correlacionen en la seva majoria amb els obtinguts amb el sistema DM96. D'altra banda, la majoria de descriptors presenten diferències significatives: 20/20 si comparem cèl·lules blàstiques amb les limfoides reactives, 18/20 per les cèl·lules de Sézary i els prolimfòcits T, 16/20 si comparem les cèl·lules del limfoma fol·licular amb les del limfoma de cèl·lules del mantell, 14/20 si comparem els prolimfòcits B amb els T i 14/20 si comparem els limfòcits normals amb els de la leucèmia limfàtica crònica. Mitjançant la nova formulació del descriptor *hairiness*, s'aconsegueix una bona especificitat pels limfòcits vellosos, com ja succeïa amb el sistema DM96. Per tant, els descriptors presentats al Capítol 4 permeten discriminar entre els diferents tipus de cèl·lules limfoides independentment del sistema d'adquisició d'imatges utilitzat.

6.1 Introducció

El present Capítol pretén demostrar que els resultats presentats al Capítol 4 amb el sistema DM96 són reproduïbles en imatges digitals obtingudes mitjançant un microscopi convencional, com és el microscopi Olympus BX43 amb la càmera digital DP73. Específicament el treball es basa en els punts següents:

- 1) utilitza un significatiu nombre d'imatges digitals adquirides de forma manual,
- 2) inclou fins a 12 tipus diferents de cèl·lules limfoides, sent nou anormals,
- 3) compara el poder discriminant dels descriptors per a la diferenciació de cèl·lules limfoides normals, reactivas, blàstiques i anormals amb l'obtingut amb el sistema DM96.

Per tal de facilitar la lectura independent del present Capítol, es descriuran alguns materials que s'han descrit prèviament al Capítol 4.

6.2 Material i mètodes

Aquest treball s'ha realitzat en sis etapes que es descriuen a continuació:

A la Figura 6.1 es mostra l'esquema d'aquest Capítol 6, el qual segueix les mateixes etapes descrites al Capítol 4. La principal diferència és el sistema d'adquisició d'imatges i la resolució de les imatges.

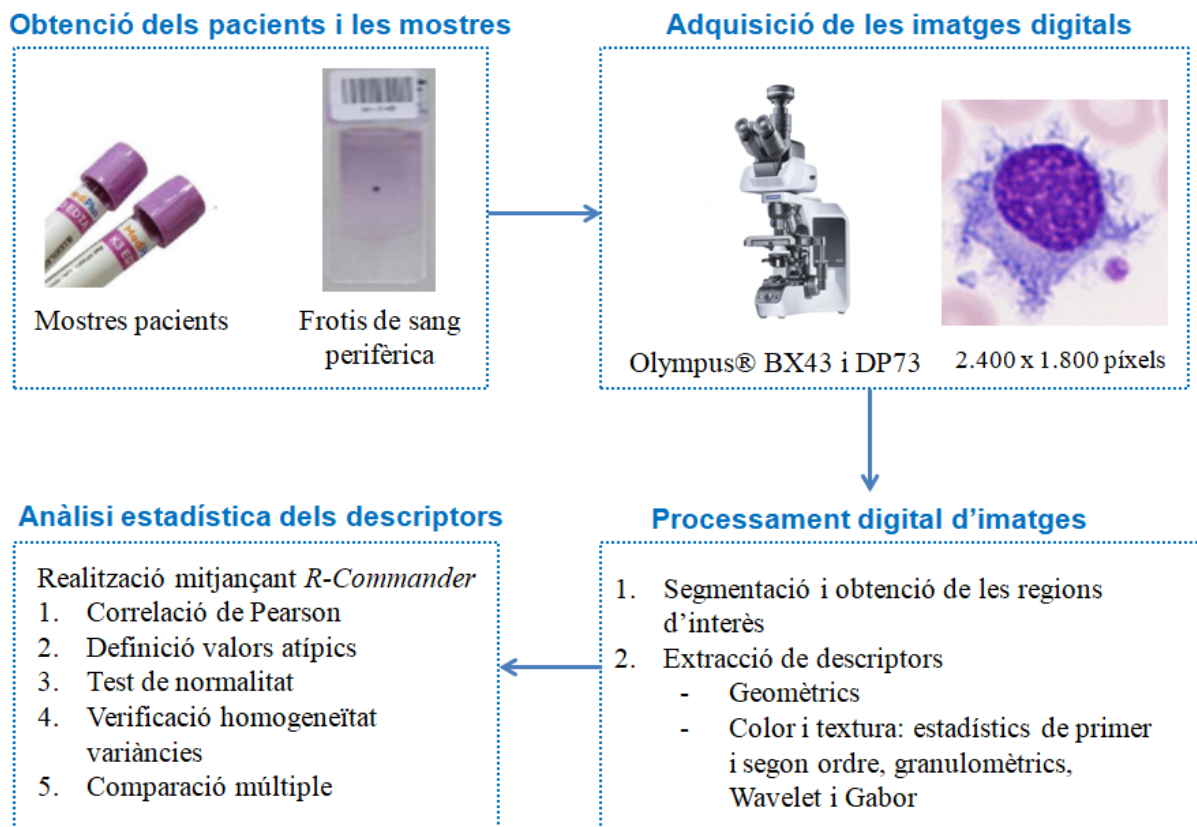


Figura 6.1: Esquema del Capítol 6.

A continuació s'explica cada una de les etapes.

6.2.1 Obtenció dels pacients i les mostres

Un total de 8.463 imatges digitals de cèl·lules de sang perifèrica corresponents a 192 pacients dels següents grups de cèl·lules limfoides van ser incloses en aquest estudi (Taula 6.1): limfòcits normals (LI) d'individus sans; cèl·lules limfoides anormals (LA) de pacients amb leucèmia limfàtica crònica (LLC), leucèmia prolimfocítica tipus B (LPB), tricoleucèmia (HCL), limfoma de la zona marginal esplènica (LZME), leucèmia de cèl·lules del mantell (LCM), limfoma fol·licular (LF), leucèmia prolimfocítica tipus T (LPT), leucèmia de limfòcits grans granulars tipus T (LLGG) i síndrome de Sézary (SS), així com cèl·lules blàstiques limfoides (BL) de pacients amb leucèmies/limfomes limfoblàstiques, cèl·lules limfoides reactives (CLR) de pacients amb infeccions virals o d'altres etiologies (Fig. 6.2). Els diagnòstics malignes van ser confirmats mitjançant troballes clíniques i morfològiques així com l'immunofenotip de les cèl·lules

limfoides analitzades per citometria de flux (veure Capítol 4 i secció 4.2.1 per més detall) i altres proves complementàries seguint la classificació de l'OMS 2016 de les neoplàsies limfoides (38).

Taula 6.1: Distribució dels diferents grups de cèl·lules limfoides, el nombre de pacients i imatges incloses a l'estudi.

Grup cel·lular	Pacients, n	Imatges, n
Limfòcits (LI)	38	989
Leucèmia limfàtica crònica (LLC)	32	863
Tricoleucèmia (HCL)	10	411
Limfoma de la zona marginal esplènica (LZME)	12	885
Limfoma fol·licular (LF)	12	938
Leucèmia de cèl·lules del mantell (LCM)	12	986
Leucèmia prolimfocítica B (LPB)	3	510
Leucèmia prolimfocítica T (LPT)	1	100
Síndrome de Sézary (SS)	7	817
Leucèmia de limfòcits grans granulars (LLGG)	15	561
Cèl·lula limfoide reactiva (CLR)	42	877
Blasts limfoides (BL)	8	526
TOTAL	192	8.463

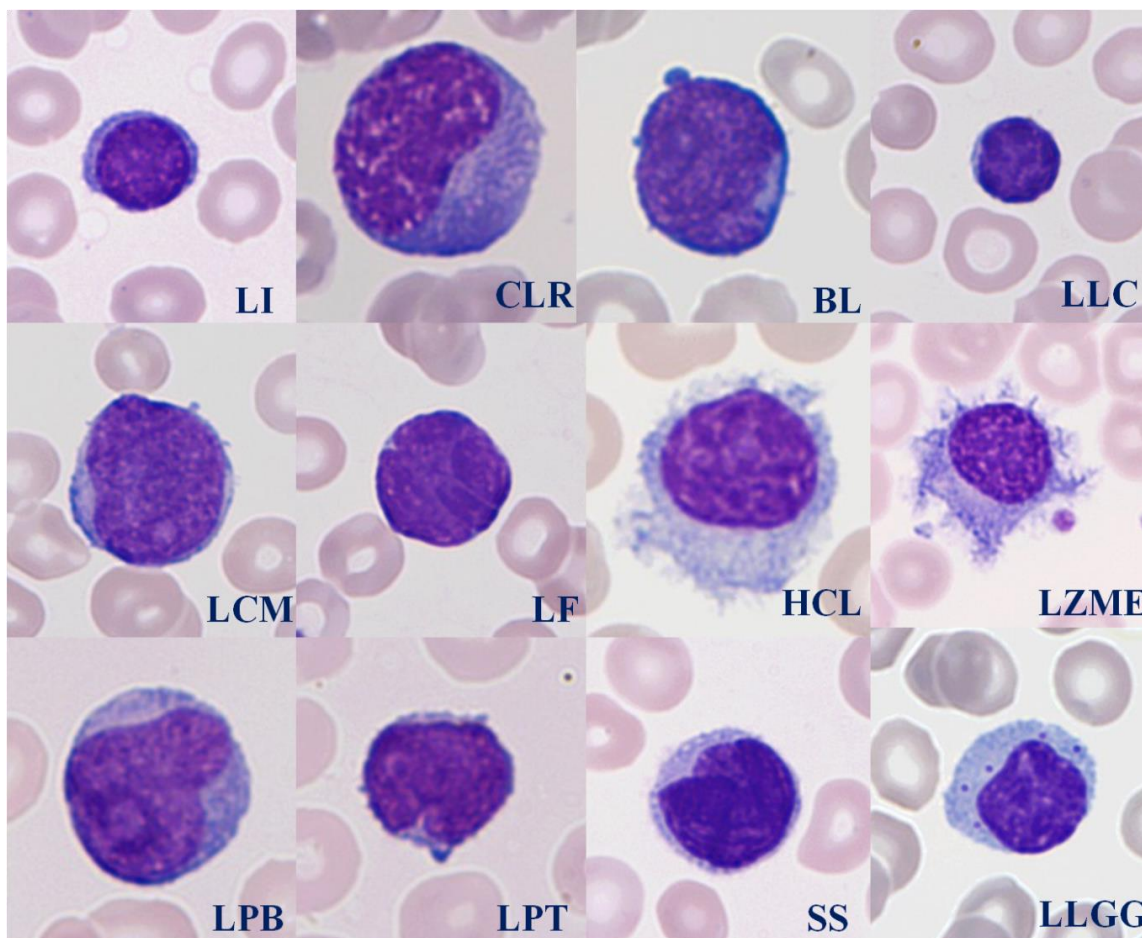


Figura 6.2: Exemples de cada un dels tipus de cèl·lules limfoides incloses a l'estudi. Imatges obtingudes amb el microscopi Olympus BX43 a x1.000 augments.

De la mateixa manera que en els estudis anteriors, els frotis de sang perifèrica es van obtenir de la rutina diària del Laboratori Core de l'Hospital Clínic de Barcelona. Per més detall, Capítol 4 i secció 4.2.1.

6.2.2 Adquisició de les imatges digitals

La qualitat dels frotis de sang perifèrica així com la morfologia cel·lular de les imatges van ser valorades per les facultatius especialistes (LP, AM) abans de realitzar l'estudi de les imatges. Cada imatge cel·lular mostrava una resolució de 2.400 x 1.800 píxels i es van adquirir manualment a 16 bits mitjançant el microscopi Olympus® BX43 i la càmera digital DP73 de CoDAlab. Un total de 8.946 imatges de cèl·lules limfoides van ser adquirides seguint la distribució que es mostra a la Taula 6.1. Tots els processos desenvolupats sobre les imatges es van realitzar i implementar a una estació de treball Dell Precision 5810 amb processador Intel® Xeon® E5-1620 v3 4 x 3.50 GHz, 16 GB de RAM, una targeta gràfica NVIDIA Quadro K2200 i MATLAB® com a software científic.

6.2.3 Processament de les imatges digitals

Les imatges originals (16 bits, 2.400 x 1.800 píxels) es van transformar a vuit bits i la seva mida es va reduir a 600 x 450 píxels perquè la segmentació, basada en *clustering*, no tingui problemes i funcioni correctament. A continuació es va seguir el mateix procediment automàtic de segmentació basat en *color clustering* i morfologia matemàtica desenvolupat pel nostre grup de recerca (19–21), per tal de separar les tres regions d'interès principals: el nucli, la cèl·lula sencera i la zona perifèrica al voltant de la cèl·lula. La quarta regió d'interès s'obté pel citoplasma calculant la diferència entre la regió de la cèl·lula i la del nucli.

Un cop obtingudes les diferents regions d'interès, es van extreure els 20 descriptors presentats al Capítol 4 de les imatges adquirides mitjançant el sistema BX43. Per a més detall, consultar l'apartat 4.2 *Material i mètodes* del Capítol 4 (Taula 4.3).

6.2.4 Anàlisi estadística dels descriptors

En primer lloc es va aplicar el test de correlació de Pearson per tal de veure si existia correlació dels descriptors entre ambdós sistemes (DM96 i BX43) i en cas d'existir correlació, poder comparar el poder discriminant dels descriptors. Posteriorment, es va dur a terme la mateixa anàlisi estadística realitzada al Capítol 4 amb el software *R-Commander* (93). Per més detall, veure el Capítol 4 i la secció 4.2.5.

6.3 Resultats

Els 20 descriptors més rellevants per a la discriminació dels 12 tipus de cèl·lules limfoides obtinguts amb el sistema DM96 s'han calculat en aquest Capítol 6 amb les noves imatges adquirides manualment amb el microscopi BX43 i la càmera digital DP73 integrada.

Cap dels 20 descriptors han seguit una distribució normal i tots ells han presentat variàncies heterogènies. Tots els 20 descriptors han mostrat una bona correlació amb els dos sistemes d'adquisició d'imatges. Consultar l'Annex A secció 9.1 per a veure els resultats dels tests de correlació de Pearson. Pel que fa als resultats de l'anàlisi estadística descriptiva amb els corresponents diagrames de caixes i els de la comparació múltiple per cada un dels 20 descriptors també es poden consultar a l'Annex A secció 9.1.

A continuació, es comparen els resultats obtinguts entre ambdós sistemes: DM96 (Capítol 4) i BX43 (Capítol 6). A la Taula 6.2 es mostra la significació estadística obtinguda de cada descriptor i sistema d'adquisició d'imatges per a les següents parelles de subconjunts de cèl·lules limfoides anormals, la discriminació de les quals a sang perifèrica per morfologia sol ser difícil: limfòcits normals – LLC, HCL – LZME, LCM – LF, LPB – LPT, LPT – SS, CLR – BL.

Com succeïa amb el sistema DM96, la majoria dels 20 descriptors mostren diferències significatives a l'hora de diferenciar les parelles d'interès amb el sistema BX43. Concretament, la parella CLR (benigne) – BL (maligne) és la que es pot diferenciar amb tots els descriptors (20/20) extrets amb el sistema BX43, mentre que amb el DM96 són 19/20. Per a la diferenciació dels limfòcits normals dels limfòcits anormals de la LLC, s'obtenen resultats similars: 15/20 descriptors mostren diferències significatives amb el sistema DM96 mentre que amb el BX43 són 14/20. Resultats similars s'observen per a la diferenciació entre els prolimfòcits B i T.

En relació als prolimfòcits T i les cèl·lules de Sézary, un total de 18/20 descriptors són significatius amb el sistema BX43, mentre que amb el DM96 són 17/20. Pel que fa a la diferenciació entre limfòcits anormals del LCM i el LF, amb el sistema DM96 19/20 descriptors són significatius mentre que amb el sistema BX43 són 16/20. Per últim, mentre que amb el sistema DM96 16/20 descriptors mostren diferències significatives entre els dos subconjunts de limfòcits vellosos (HCL i LZME), amb el sistema BX43 només són significatius la meitat (10/20).

Taula 6.2: p valors obtinguts mitjançant el test de Kruskal-Wallis-Dunn i l'ajust de Bonferroni. Comparació múltiple dels 20 descriptors obtinguts mitjançant el sistema DM96 i el BX43 per a sis parelles de subconjunts de cèl·lules limfoides anormals. Un p<0,05 s'ha considerat significatiu. NS, no significació.

Descriptor	LI vs LLC		HCL vs LZME		LCM vs LF		LPB vs LPT		LPT vs SS		CLR vs BL	
	DM96	BX43	DM96	BX43	DM96	BX43	DM96	BX43	DM96	BX43	DM96	BX43
1.Relació nucli/citoplasma	< 2 ⁻¹⁶		<0,00001	NS	< 2 ⁻¹⁶		< 2 ⁻¹⁶		0,001	<0,05	< 2 ⁻¹⁶	
2.Perímetre nuclear	NS		NS		< 2 ⁻¹⁶		< 2 ⁻¹⁶		NS	<0,00001	<0,00001	<0,001
3.Diàmetre equivalent	NS	<0,00001	NS		< 2 ⁻¹⁶		< 2 ⁻¹⁶		NS	<0,00001	<0,00001	< 2 ⁻¹⁶
4.Curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula	0,02	< 2 ⁻¹⁶	< 2 ⁻¹⁶		< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	NS	<0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	
5.Asimetria de l'histograma del blau de la cèl·lula	< 0,00001	NS	< 2 ⁻¹⁶	NS	NS		< 2 ⁻¹⁶		NS		< 2 ⁻¹⁶	
6.Information measure of correlation I del magenta de la cèl·lula	NS	< 2 ⁻¹⁶	<0,00001		< 2 ⁻¹⁶		NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶		< 2 ⁻¹⁶	
7.Desviació estàndard de l'histograma del blau de la cèl·lula	< 0,00001	< 2 ⁻¹⁶	<0,00001		< 2 ⁻¹⁶		<0,00001		< 2 ⁻¹⁶		< 2 ⁻¹⁶	
8.Entropia de l'histograma del magenta de la cèl·lula	< 0,00001	NS	NS		< 2 ⁻¹⁶		< 2 ⁻¹⁶		< 2 ⁻¹⁶	<0,00001	< 2 ⁻¹⁶	
9.Asimetria de l'histograma del verd de la cèl·lula	NS	<0,00001	< 0,00001	NS	< 2 ⁻¹⁶		< 2 ⁻¹⁶		<0,00001	<0,05	< 2 ⁻¹⁶	
10.Mitjana de la corba pseudo-granulomètrica del negre del citoplasma	< 2 ⁻¹⁶	NS	NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶		NS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	<0,001	< 2 ⁻¹⁶	
11.Mitjana de l'histograma del blau del nucli	< 2 ⁻¹⁶	<0,00001	<0,00001	NS	< 2 ⁻¹⁶		< 2 ⁻¹⁶		<0,00001		< 2 ⁻¹⁶	
12.Cluster shade de la saturació de la cèl·lula	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	<0,00001	< 2 ⁻¹⁶		NS		< 2 ⁻¹⁶		< 2 ⁻¹⁶	
13.Mitjana de l'histograma del magenta del citoplasma	< 2e ⁻¹⁶	NS	0,0001	<0,00001	< 2 ⁻¹⁶		NS	<0,00001	< 2 ⁻¹⁶	NS	< 2 ⁻¹⁶	
14.Homogeneïtat del magenta de la cèl·lula	< 0,00001	< 2 ⁻¹⁶	<0,00001		< 2 ⁻¹⁶		< 2 ⁻¹⁶	<0,001	<0,00001		< 2 ⁻¹⁶	<0,00001
15.Curtosi de l'histograma de la saturació de la cèl·lula	NS	< 2 ⁻¹⁶	<0,00001	NS	< 2 ⁻¹⁶		< 2 ⁻¹⁶		< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	
16.Information measure of correlation I del cian del nucli	< 2 ⁻¹⁶		<0,00001	< 0,01	0,001	NS	<0,00001	NS	<0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	
17.Hairiness	< 2 ⁻¹⁶	<0,00001	0,0001	NS	< 2 ⁻¹⁶		<0,00001	NS	<0,00001	< 2 ⁻¹⁶	NS	0,002
18.Desviació estàndard de l'histograma de la saturació de la cèl·lula	< 2 ⁻¹⁶		0,008	< 0,00001	< 2 ⁻¹⁶		< 2 ⁻¹⁶	<0,00001	< 2 ⁻¹⁶		< 2 ⁻¹⁶	
19.Circularitat nuclear	< 2 ⁻¹⁶		<0,00001	< 0,05	< 2 ⁻¹⁶	NS	<0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,01	NS	<0,00001
20.Homogeneïtat del negre del citoplasma	< 2 ⁻¹⁶		<0,00001	NS	< 2 ⁻¹⁶		0,005	< 2 ⁻¹⁶	<0,00001	< 0,001	< 2 ⁻¹⁶	<0,001
Parelles significatives	15/20	14/20	16/20	10/20	19/20	16/20	16/20	14/20	17/20	18/20	18/20	20/20

En relació als descriptors que mostren significació estadística quan es fan comparacions entre les cèl·lules limfoides reactives amb cadascun dels nou subconjunts de cèl·lules limfoides anormals incloses a l'estudi, amb el sistema DM96 hem observat que això succeïa en 5/20 descriptors: el 7è, 8è, 11è, 16è i 18è. Per a més detall, consultar la Taula 4.5 de l'apartat 4.3 Resultats del Capítol 4. D'aquests cinc descriptors, tres es repeteixen pel sistema BX43. A la Taula 6.3 s'inclouen els descriptors que mostren significació estadística quan es fan comparacions entre les CLR amb cadascun dels nou subconjunts de cèl·lules limfoides anormals. En total són sis: dos descriptors geomètrics (*perímetre nuclear* i *diàmetre equivalent de la cèl·lula*) i quatre de color i textura (el 5è, 7è, 11è i 18è més rellevants).

Taula 6.3: p valors dels descriptors geomètrics i de color i textura que mostren significació estadística quan es comparen els seus valors de les cèl·lules limfoides reactives amb cadascun dels nous subconjunts de limfòcits anormals.

Grup cel·lular	Perímetre nuclear (2n)	Diàmetre equivalent de la cèl·lula (3r)	Asimetria de l'histograma del blau de la cèl·lula (5è)	Desviació estàndard del blau de la cèl·lula (7è)	Mitjana del blau del nucli (11è)	Desviació estàndard de la saturació de la cèl·lula (18è)
LLC	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶
HCL	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LZME	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001
LCM	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LF	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LPB	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LPT	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
LLGG	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶
SS	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 2 ⁻¹⁶	< 0,00001	< 2 ⁻¹⁶	< 2 ⁻¹⁶

HCL, tricoleucèmia; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulars; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

Pel que fa a la resta dels descriptors geomètrics (cinc en total), en el present treball s'ha reproduït l'elevada variabilitat en la *relació N/C* en el LCM a causa del pleomorfisme característic d'aquestes cèl·lules limfoides. Les dues cèl·lules limfoides velloses (HCL i LZME) i les CLR mostren els valors més baixos de *relació N/C*, mentre que les cèl·lules limfoides anormals del LF presenten els més alts. Aquests resultats concorden amb els obtinguts amb el sistema DM96. Com s'ha mencionat al Capítol 4, treballs anteriors del nostre grup de recerca també han trobat que la *relació N/C* és el descriptor més rellevant en la diferenciació entre subconjunts de limfòcits anormals (31) i també entre cèl·lules blàstiques limfoides en comparació amb les cèl·lules limfoides reactives (123).

D'altra banda, els valors del *perímetre nuclear* presenten una gran variabilitat en el LCM i la SS. Això és degut a l'heterogeneïtat morfològica en el primer i a la inclusió de les cèl·lules de Sézary petites (variant Lutzner) en la segona. Aquests resultats coincideixen amb els obtinguts amb el sistema DM96. En relació al *hairiness*, la quantificació del perfil de la regió citoplasmàtica ha permès discriminar els limfòcits vellosos d'altres limfòcits anormals, com ja succeïa amb el sistema DM96 però ha estat necessària una reformulació del descriptor, tal com s'explica en el següent subapartat. Si analitzem els descriptors de color i textura, tots ells han mostrat una capacitat discriminatòria similar a l'observada amb el sistema DM96 (veure Taula 6.2).

6.3.1 Reformulació del descriptor *hairiness*

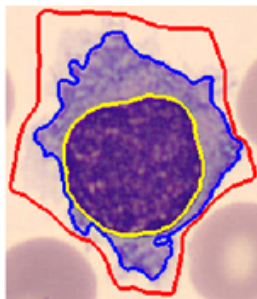
Aquest descriptor va ser creat específicament per a les imatges adquirides amb el sistema DM96, tal i com s'ha detallat al Capítol 3 secció 3.4.1.5. És per aquest motiu que si es calcula per les imatges del sistema BX43 sense cap modificació, no mostra l'especificitat per les cèl·lules velloses de l'HCL i el LZME que s'havia observat pel cas del sistema DM96. A continuació s'expliquen les principals diferències entre ambdues imatges i com s'han abordat.

En primer lloc, existeix una diferència de mides entre les cèl·lules de les imatges adquirides amb ambdós sistemes. El camp de les imatges adquirides amb el sistema BX43 és més gran i la cèl·lula es veu més petita en comparació amb les imatges del sistema DM96. Degut a la diferència de mides, el descriptor *hairiness* ja de per si no és comparable perquè està mesurant una àrea.

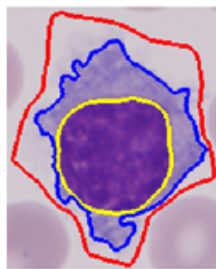
En segon lloc, els límits establerts pels valors d'intensitat de verd en la formulació del descriptor *hairiness* van ser escollits a partir d'histogrames d'imatges adquirides amb el sistema DM96. Concretament, es va triar el rang entre 70 i 215 d'intensitat de verd per delimitar les vellositats. Si comparem els histogrames del component verd de la regió externa d'una mateixa cèl·lula adquirida amb els dos sistemes, s'observa que el rang de valors d'intensitat de verd es desplaça cap a l'esquerra en el cas de les imatges del sistema BX43. A la Figura 6.3 es mostra un exemple.

Com s'observa a la Figura 6.3, en ambdós histogrammes s'observa una distribució bimodal, on el pic més alt correspon als píxels més brillants, els quals corresponen al fons de la imatge, mentre que el pic secundari correspon a les vellositats. Si ens fixem en el rang de valors de les intensitats de verd de la regió externa de la cèl·lula, s'observa que l'histograma del sistema DM96 està desplaçat cap a la dreta (arriba fins a 230). Això indica que les imatges del sistema DM96 són més brillants i, per tant, els píxels de la regió externa de la cèl·lula prenen valors més elevats d'intensitat de verd.

I) Imatge original



Sistema DM96



Sistema BX43

II) Histograma del component verd de la regió externa de la cèl·lula

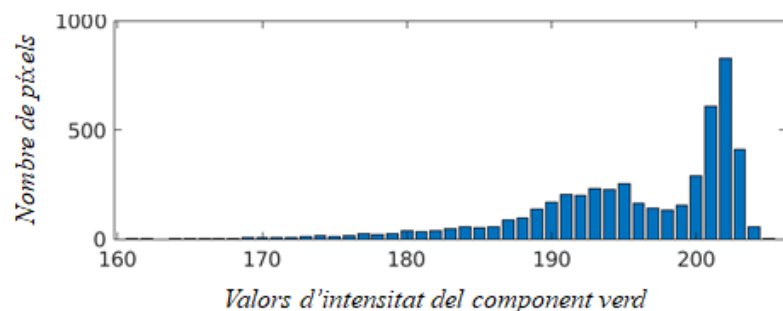
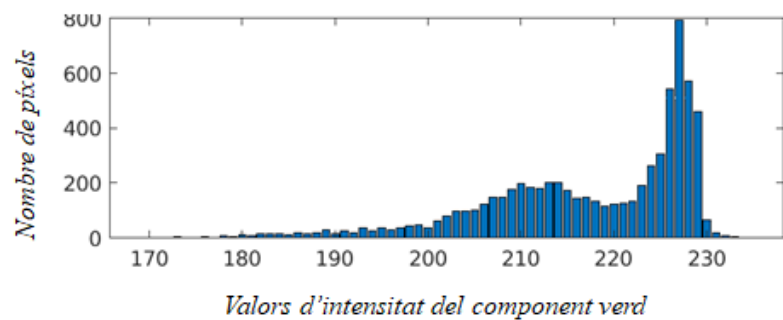


Figura 6.3: Comparació del descriptor *hairiness* per la mateixa imatge (tricoleucòcit) adquirida amb els dos sistemes d'adquisició d'imatges (BX43 i DM96).

A la Figura 6.4 es mostra una cèl·lula sense prolongacions citoplasmàtiques (limfòcit anormal de la leucèmia limfàtica crònica) i el corresponent histograma del component verd de la regió externa de la cèl·lula. La distribució en aquest cas és unimodal i els píxels amb una intensitat de verd compresa entre 200 i 220, corresponen al fons i no a les vellositats. Per tant, el descriptor *hairiness* està comptant els píxels del fons de la imatge i per això erròniament totes les cèl·lules mostren valors elevats de *hairiness* i aquest

no és específic dels limfòcits vellosos. En definitiva, cal ajustar els llindars del descriptor *hairiness* descrits per S. Alférez (70 i 215) per a les imatges del sistema BX43.

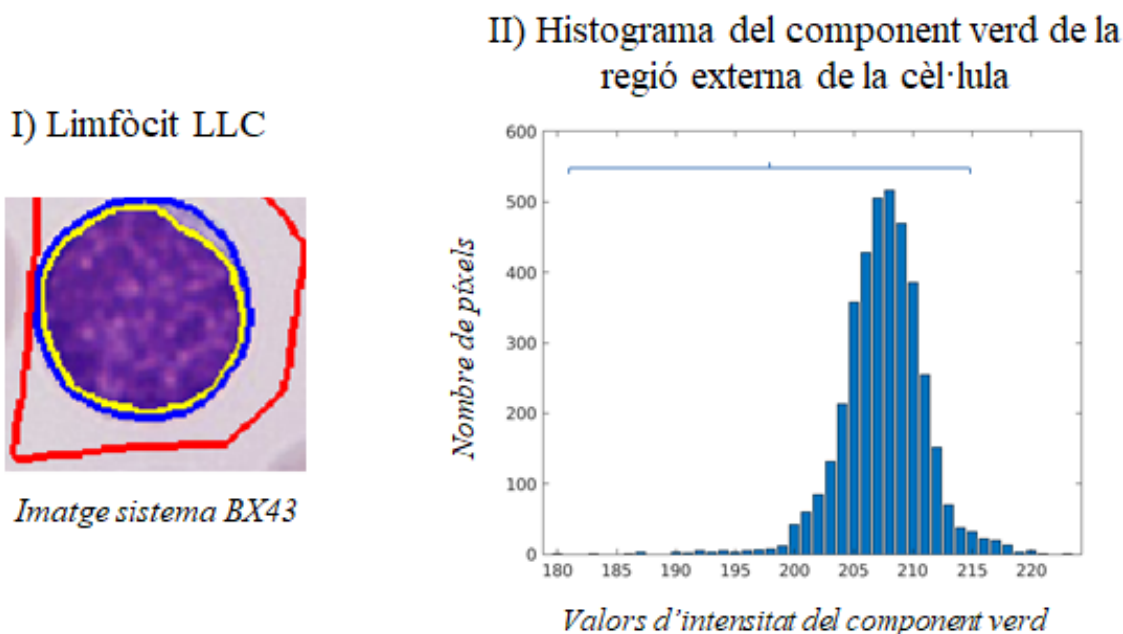


Figura 6.4: Càlcul del descriptor *hairiness* per a una imatge d'un limfòcit anormal de la leucèmia limfàtica crònica (LLC) adquirida amb el sistema BX43 i el corresponent histograma del component verd de la regió externa de la cèl·lula.

A continuació es detalla la nova formulació utilitzada a la present Tesi Doctoral per tal d'ajustar el descriptor *hairiness* a les imatges del sistema BX43:

$$Hairiness_{BX43} = \frac{Hairiness_{DM96} \text{ però restant 20 unitats als llindars mínim i màxim}}{\text{Àrea de la cèl·lula}}$$

La divisió entre l'àrea de la cèl·lula permet que els valors del descriptor *hairiness* siguin semblants (normalitzats) pel que fa a la mida. La resta de 20 unitats als llindars mínim i màxim s'obté de calcular en una mostra d'imatges la diferència mitjana de les intensitats de verd i permet que el nou llindar sigui l'adient per tallar la part dreta de l'histograma i no s'agafin píxels del fons de la imatge per al càlcul del descriptor *hairiness*. En aquesta formulació, el descriptor *Hairiness* és adimensional.

6.4 Discussió

L'objectiu d'aquest Capítol 6 era utilitzar una nova font d'adquisició d'imatges (microscopi BX43) i analitzar si els 20 descriptors presentats al Capítol 4 són reproduïbles per a la diferenciació de 12 tipus diferents de cèl·lules limfoides, incloent cèl·lules limfoides normals, reactives, blàstiques i anormals.

D'altra banda, dues de les tres propietats que esperàvem que tinguessin els descriptors del Capítol 4, s'han plantejat en el present Capítol: (i) formulació quantitativa i (ii) eficiència en la discriminació entre un nombre significatiu de diferents subconjunts de cèl·lules limfoides. La tercera propietat, que tinguessin una interpretació qualitativa en termes de característiques morfològiques, ja s'ha realitzat al Capítol 4 de manera que no s'ha repetit en el present Capítol.

Pel que fa a l'estat de l'art de descriptors utilitzats en treballs anteriors per a la diferenciació cel·lular limfoide, als capítols anteriors ja se n'ha fet una extensa explicació. Per tant, a la present discussió no se'n farà referència i se centrarà a explicar els resultats obtinguts.

En primer lloc, s'han correlacionat els valors de les 20 variables quantitatives (descriptors) obtingudes entre ambdós sistemes d'adquisició d'imatges (sistema DM96 i BX43) per cada un dels 12 grups de cèl·lules limfoides estudiades: nou limfòcits anormals (incloent sis neoplàsies limfoides B i tres de tipus T), cèl·lules limfoides reactives, blasts limfoides i limfòcits normals.

Tots els descriptors han mostrat una bona correlació entre ambdós sistemes d'adquisició d'imatges. Mitjançant una nova formulació del descriptor *hairiness*, la qual normalitza la mida de la cèl·lula i adapta els llindars d'intensitat de verd, s'aconsegueix una bona especificitat pels limfòcits vellosos, com ja succeïa amb el sistema DM96. A partir de l'histograma es pot identificar la presència o absència de vellositats citoplasmàtiques: quan la distribució és bimodal, la cèl·lula contindrà vellositats (Figura 6.3) i si és unimodal, la cèl·lula no en mostrarà (Figura 6.4).

Tots els descriptors (20/20) han presentat diferències significatives quan es comparen cèl·lules blàstiques amb cèl·lules limfoides reactives, 18 de 20 descriptors per diferenciar les cèl·lules de Sézary dels prolimfòcits T, 16 de 20 si es comparen les cèl·lules limfoides anormals del limfoma fol·licular amb les cèl·lules del mantell, 14 de 20 per a diferenciar entre els prolimfòcits B i els T i 14 de 20 si es comparen els limfòcits normals amb els de la leucèmia limfàtica crònica. A més a més, sis descriptors quantitius de color i textura han mostrat ser significatius per a discriminar les cèl·lules limfoides reactives presents en infeccions dels limfòcits anormals observats quan hi ha una neoplàsia limfoide (veure Taula 6.3).

Per tant, les noves variables citològiques basades en descriptors geomètrics i de color i textura presentades al Capítol 4 han contribuït a una valoració morfològica igual de precisa que amb el sistema DM96, ja que permeten discriminar entre els diferents tipus de cèl·lules limfoides independentment del sistema d'adquisició d'imatges.

De la mateixa manera que s'ha explicat al Capítol 4, l'aplicabilitat d'aquesta investigació rau en el potencial de reduir la subjectivitat associada a la morfologia de sang perifèrica a qualsevol laboratori clínic, però destacant que no és indispensable disposar del sistema DM96. L'imprescindible és obtenir un frotis de sang perifèrica de bona qualitat.

6.5 Conclusions

Aquest Capítol s'ha centrat en dos aspectes principals: la correlació dels descriptors entre ambdues fonts d'adquisició d'imatges (DM96 i BX43) i l'avaluació de la capacitat discriminatòria dels descriptors aplicats a les imatges obtingudes amb el microscopi convencional (BX43). Els dos aspectes es troben vinculats consecutivament per acabar obtenint noves variables per al reconeixement de les cèl·lules limfoides anormals objecte d'estudi.

Els resultats obtinguts mostren que el conjunt dels 20 descriptors extrets, tant de descriptors geomètrics com de color i textura, ha estat capaç de descriure la mida, la forma, la textura i el color dels diferents tipus de cèl·lules limfoides, com ja s'aconseguia amb el sistema DM96. Tots els valors dels descriptors han mostrat bona correlació entre ambdós sistemes d'adquisició d'imatges (BX43 i DM96). S'ha proposat una nova formulació del descriptor *hairiness*, la qual normalitza la mida de la cèl·lula entre sistemes d'adquisició d'imatges, adapta els llindars d'intensitat de verd i aconsegueix mostrar una bona especificitat pels limfòcits vellosos, com ja succeïa amb el sistema DM96. D'altra banda, aquests 20 descriptors han mostrat una bona capacitat discriminatòria entre les diferents cèl·lules limfoides anormals a partir d'imatges de sang perifèrica adquirides mitjançant un microscopi convencional.

Seguint amb el propòsit de la *Tesi Doctoral* de la caracterització morfològica dels diferents grups cel·lulars limfoides, el proper Capítol 7 utilitza la mateixa font d'adquisició d'imatges (el sistema BX43) però el punt de partida són els descriptors obtinguts al Capítol 5, els quals pretenen contribuir a la identificació objectiva dels limfòcits de la leucèmia limfàtica crònica, la síndrome de Sézary, els limfòcits grans granulars i els limfòcits vellosos circulants a sang perifèrica.

Capítol 7

7 Descriptors quantitius citològics per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, les cèl·lules de Sézary, els limfòcits vellosos i els limfòcits grans granulars mitjançant l'anàlisi d'imatges utilitzant el sistema BX43

Resum

L'objectiu de l'estudi que es presenta en aquest Capítol 7 és reproduir els resultats obtinguts amb els quatre descriptors quantitius citològics presentats al treball del Capítol 5 amb el microscopi convencional (sistema BX43) per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, la tricoleucèmia i el limfoma de la zona marginal esplènica, la síndrome de Sézary i la leucèmia de limfòcits grans granulars. En primer lloc, s'han adquirit 8.946 imatges mitjançant el microscopi Olympus® BX43 i la càmera digital DP73 corresponents a limfòcits normals, cèl·lules limfoides reactives, blàstiques i 12 tipus de limfòcits B o T anormals. De cada una de les imatges, s'han extret els quatre descriptors proposats al Capítol 5 per a mesurar les següents característiques citològiques: la cromatina madura i condensada, la cromatina cerebriforme, la granulació citoplasmàtica i les vellositats citoplasmàtiques. En una segona etapa, s'han realitzat correlacions de Pearson dels quatre descriptors entre ambdós sistemes d'adquisició (DM96 i BX43) per tal de verificar que els descriptors quantitius eren comparables per a cada un dels 15 tipus de cèl·lules limfoides incloses a l'estudi. Per últim, s'ha realitzat una avaluació de l'eficiència diagnòstica dels descriptors per a la classificació d'un frotis en una única característica cel·lular utilitzant els valors llinars obtinguts al Capítol 5 i s'han inclòs un total de 119 frotis de sang perifèrica de diferents pacients. La *correlació del cian del nucli* identifica la cromatina madura i condensada de la leucèmia limfàtica crònica amb una sensibilitat del 72 % i una especificitat del 84 % i la *desviació estàndard de la corba granulomètrica del cian del nucli* és específica (100 %) per a la cromatina cerebriforme. *L'asimetria de l'histograma del component de color u de la regió del citoplasma* mostra una sensibilitat i especificitat per sobre del 92 % per a la detecció de la granulació citoplasmàtica. El descriptor *hairiness* ha mostrat una especificitat del 79 % per als limfòcits vellosos. Pel que fa als frotis utilitzats en l'avaluació de l'exactitud dels quatre descriptors específics, el 89 % s'han classificat correctament.

7.1 Introducció

Tradicionalment, l'anàlisi morfològica del frotis de sang perifèrica s'ha realitzat de forma manual mitjançant un microscopi i malgrat el seu paper crucial en el diagnòstic, a la literatura no existeixen valors objectius per a definir variables citològiques, ja que les característiques citològiques distintives de les cèl·lules limfoides anormals són subjectives i es troben influenciades per la interpretació del citòleg (2).

En aquest sentit, l'estudi presentat al Capítol 5 ha proporcionat escales numèriques objectives de 15 tipus diferents de cèl·lules limfoides incloent nou anormals mitjançant l'ús de la microscòpia digital i l'anàlisi d'imatges. L'estudi també ha presentat quatre descriptors sensibles i específics per a les següents característiques citològiques: cromatina madura i condensada de la leucèmia limfàtica crònica, cromatina cerebriforme de les cèl·lules de Sézary, granulació citoplasmàtica dels limfòcits grans granulars i vellositats citoplasmàtiques dels tricoleucòcits i els limfòcits vellosos del limfoma de la zona marginal esplènica.

L'objectiu de l'estudi que es presenta en aquest Capítol 7 és reproduir els resultats obtinguts amb els quatre descriptors quantitius citològics presentats al treball del Capítol 5 amb les imatges adquirides amb el microscopi convencional (sistema BX43) per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, la tricoleucèmia i el limfoma de la zona marginal esplènica, la síndrome de Sézary i la leucèmia de limfòcits grans granulars.

7.2 Material i mètodes

Aquest estudi consta de dues etapes que es descriuen a continuació.

7.2.1 Càlculs dels descriptors quantitius citològics i les corresponents escales numèriques

La Figura 7.1 detalla els diferents processos inclosos en la primera etapa, la qual conté quatre subetapes.

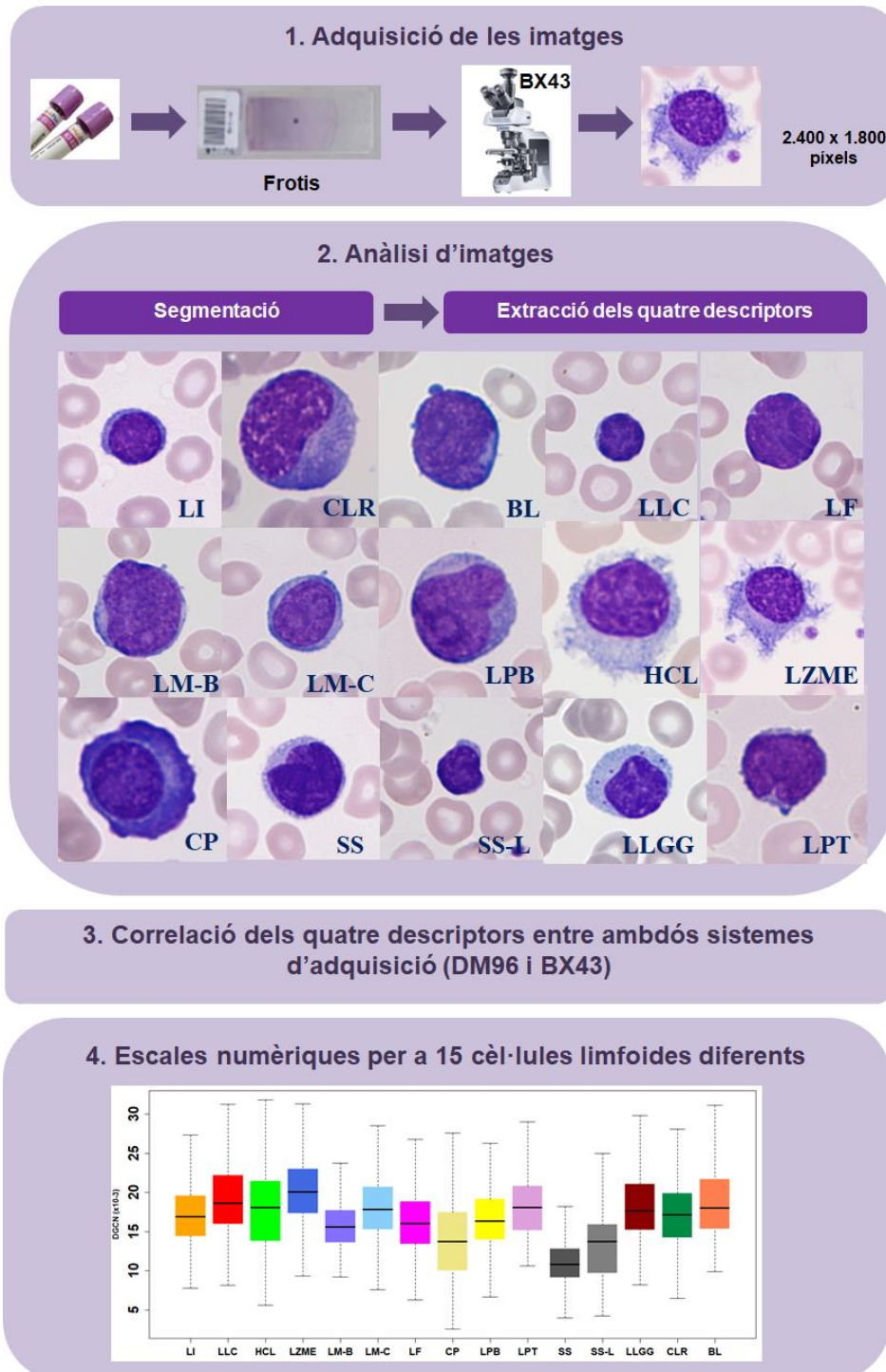


Figura 7.1 Esquema de la primera etapa de l'estudi.

La primera subetapa consisteix en l'adquisició manual de les imatges de cèl·lules limfoides de sang perifèrica mitjançant el microscopi BX43 i la càmera digital DP73 integrada. A continuació es realitza l'anàlisi d'imatges que consisteix en fer un preprocessament, una segmentació i finalment, l'extracció dels quatre descriptors d'interès. Posteriorment i abans de realitzar l'anàlisi estadística descriptiva per a l'establiment d'escala numèriques, s'ha aplicat el test de correlació de Pearson, com ja s'ha fet al treball del Capítol 6 amb les imatges adquirides amb el sistema BX43.

A continuació es detallen les subetapes:

1) Preparació de les mostres de sang i adquisició de les imatges

Un total de 208 frotis de sang perifèrica corresponents a 208 pacients contenint els següents grups de cèl·lules limfoides van ser inclosos en aquest estudi (veure Taula 7.1): limfòcits normals (LI) d'individus sans; cèl·lules limfoides anormals (LA) de pacients amb leucèmia limfàtica crònica (LLC), leucèmia prolimfocítica tipus B (LPB), tricoleucèmia (HCL), limfoma de la zona marginal esplènica (LZME), leucèmia de les cèl·lules del mantell (LCM) amb morfologia blàstica (LM-B) o clàssica (LM-C), limfoma fol·licular (LF), leucèmia prolimfocítica tipus T (LPT), leucèmia de limfòcits grans granulars tipus T (LLGG) i síndrome de Sézary incloent la variant clàssica (SS) i la variant Lutzner de cèl·lules petites (SS-L). D'altra banda també es van incloure cèl·lules blàstiques limfoides (BL) de pacients amb leucèmies/limfomes limfoblàstiques i cèl·lules limfoides reactives (CLR) de pacients amb infeccions virals o d'altres etiologies (Fig.7.1, segona subetapa).

Taula 7.1: Tipus de cèl·lula, nombre (n) de pacients i imatges utilitzades a l'estudi.

Tipus de cèl·lula	Pacients, n	Imatges, n
Limfòcits	38	989
Limfòcit anormal de la leucèmia limfàtica crònica	32	863
Limfòcit vellós de la tricoleucèmia	10	411
Limfòcit vellós del limfoma de la zona marginal esplènica	12	885
Limfòcit anormal del limfoma fol·licular	12	938
Cèl·lula del limfoma del mantell morfologia blàstica	3	195
Cèl·lula del limfoma del mantell morfologia clàssica	9	791
Cèl·lula plasmàtica	16	483
Prolimfòcit B	3	510
Prolimfòcit T	1	100
Cèl·lula de Sézary variant clàssica	3	440
Cèl·lula de Sézary variant Lutzner o petita	4	377
Limfòcit gran granular	15	561
Cèl·lula limfoide reactiva	42	877
Blast limfoide	8	526
TOTAL	208	8.946

Els diagnòstics malignes van ser confirmats mitjançant troballes clíniques i morfològiques així com l'immunofenotip de les cèl·lules limfoides analitzades per citometria de flux i altres proves complementàries seguint la classificació de l'OMS 2017 de les neoplàsies limfoides (38). Els frotis de sang perifèrica es van obtenir de la rutina diària del Laboratori Core de l'Hospital Clínic de Barcelona. Es van realitzar a partir de sang venosa recollida en EDTA i utilitzant l'extensor-tenyidor automàtic SP1000i (Sysmex, Kobe, Japó) i seguint el procediment de tinció de May Grünwald-Giemsa durant les quatre primeres hores des de l'extracció sanguínia.

La qualitat dels frotis de sang perifèrica així com la morfologia cel·lular de les imatges van ser valorades per les facultatius especialistes (LP, AM) abans de realitzar l'estudi de les imatges. Cada imatge cel·lular tenia una resolució de 2.400 x 1.800 píxels i van ser adquirides manualment en 16 bits mitjançant el microscopi Olympus® BX43 i la càmera digital DP73. Un total de 8.946 imatges de cèl·lules limfoides van ser recollides i distribuïdes com es mostra a la Taula 7.1. Tots els processos desenvolupats sobre les imatges es van realitzar i implementar a una estació de treball Dell Precision 5810 amb processador Intel® Xeon® E5-1620 v3 4 x 3.50 GHz, 16 GB de RAM, una targeta gràfica NVIDIA Quadro K2200 i MATLAB® com a software científic.

2) Anàlisi d'imatges

Les imatges originals (16 bits, 2.400 x 1.800 píxels) es van transformar a vuit bits i la seva mida es va reduir a 600 x 450 píxels com al Capítol anterior. A continuació es va seguir el mateix procediment automàtic de segmentació basat en *color clustering* i morfologia matemàtica desenvolupat pel nostre grup de recerca (19–21), per tal de separar les tres regions d'interès principals: el nucli, la cèl·lula sencera i la regió externa de la cèl·lula. La quarta regió d'interès s'obté pel citoplasma calculant la diferència entre la regió de la cèl·lula i la del nucli.

En el treball que es presenta en aquest Capítol, s'han extret els quatre descriptors presentats al Capítol 5 (veure Taula 7.2) de les imatges adquirides mitjançant el sistema BX43. Per a més informació, consultar la secció 5.2.1 Material i mètodes del Capítol 5. Pel descriptor *hairiness* s'ha utilitzat la nova formulació presentada al Capítol 6.

Taula 7.2: Descriptors quantitius proposats per a la quantificació de quatre característiques citològiques.

Característica citològica	Descriptor quantitiu
Cromatina madura i condensada	<i>Correlació del cian del nucli</i>
Cromatina cerebriforme	<i>Desviació estàndard de la corba granulomètrica del cian del nucli</i>
Granulació citoplasmàtica	<i>Asimetria de l'histograma del component u del citoplasma</i>
Velositats citoplasmàtiques	<i>Hairiness</i>

3) Correlacions de Pearson

Per tal de poder realitzar l'anàlisi estadística dels descriptors així com l'avaluació de la seva eficiència diagnòstica (segona part de l'estudi), abans es van correlacionar els valors obtinguts del sistema BX43 amb els presentats amb el DM96 (Capítol 5) per cada un dels descriptors considerant cada tipus de cèl·lula limfoide mitjançant el test de correlació de Pearson.

4) Establiment d'escala numèriques objectives

En aquest estudi s'han aplicat els mateixos tests estadístics sobre els quatre descriptors que els presentats al Capítol 5 amb el software *R-Commander* (93). Per més detall, veure Capítol 5 i secció 5.2.1. Els valors numèrics per a cada descriptor i per a cada un dels 15 tipus de cèl·lules limfoides s'han calculat segons el valor de la mediana i el rang interquartílic, el qual es calcula de la diferència entre els quartils superior i inferior.

7.2.2 Avaluació dels descriptors quantitatius per a la classificació de frotis

La segona etapa de l'estudi (Figura 7.2) va consistir en avaluar el rendiment dels descriptors presentats al Capítol 5 per a la classificació de frotis en funció dels següents subtipus cel·lulars limfoides:

- 1) Cromatina madura i condensada per la LLC
- 2) Cromatina cerebriforme per a les cèl·lules de Sézary
- 3) Granulació azuròfila citoplasmàtica per als limfòcits grans granulars
- 4) Vellositats citoplasmàtiques per l'HCL i el LZME

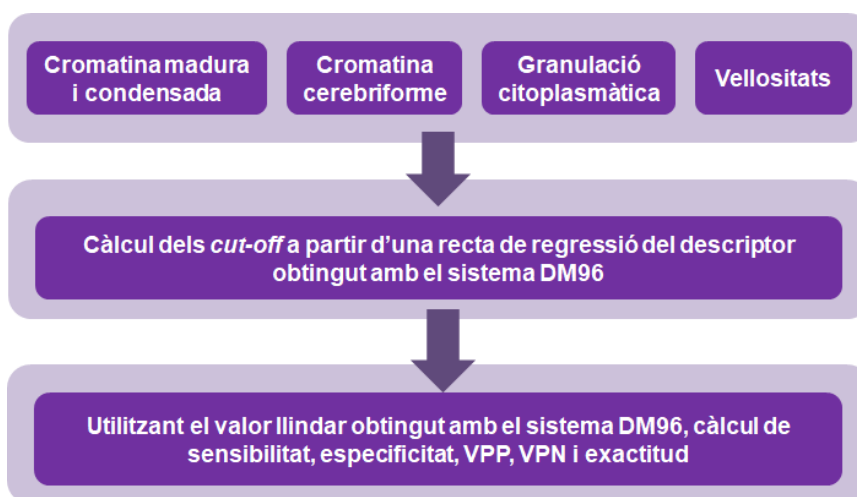


Figura 7.2: Esquema de la segona etapa de l'estudi.

En primer lloc, per a l'avaluació de cada un dels descriptors, es van calcular els valors llindars o *cut-offs* mitjançant el càlcul d'una recta de regressió a partir dels *cut-offs* obtinguts amb el sistema DM96 per tal d'obtenir el valor llindar equivalent pel sistema BX43. Com s'ha explicat al Capítol 5, si el valor del descriptor obtingut era superior o inferior al valor de tall, es pot concloure que la cèl·lula de la imatge mostra o no la característica citològica corresponent. És a dir, permet diferenciar entre dues categories, generalment positiva (p. ex. cromatina cerebriforme) i negativa (p. ex. no cerebriforme).

A continuació, els valors de tall obtinguts van ser utilitzats en una prova d'avaluació on el frotis de sang perifèrica s'ha pres com a unitat d'anàlisi (130). Es van incloure un total de 174 frotis que contenen de 12 a 56 imatges cel·lulars cadascun, incloent un total de 6.969 imatges (veure Taula 7.3).

Taula 7.3: Nombre total d'imatges adquirides amb el microscopi BX43 corresponents a cada diagnòstic i descriptor

Descriptor cromatina madura i condensada	
	Imatges
Imatges LLC	863
Imatges no LLC	1.728
Descriptor cromatina cerebriforme	
	Imatges
Imatges SS	218
Imatges no SS	989
Descriptor granulació citoplasmàtica	
	Imatges
Imatges LLGG	435
Imatges no LLGG	1.430
Descriptor vellositats citoplasmàtiques	
	Imatges
Imatges HCL/LZME	321
Imatges no HCL/LZME	985

HCL, tricoleucèmia; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulars; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

Seguint el mateix procediment realitzat amb les imatges del sistema DM96, per a cada característica (p. ex., la cromatina cerebriforme), dues citòlogues es van seleccionar dos grups d'imatges de frotis obtinguts de pacients amb diagnòstics confirmats:

- (1) *Positiu*: quan les imatges cel·lulars exhibien la morfologia específica de la malaltia hematològica (p. ex. cèl·lules de Sézary)
- (2) *Negatiu*: quan les imatges cel·lulars no mostraven la morfologia específica característica (p. ex. limfòcits vellosos).

A continuació, es va realitzar una prova de classificació binària cega per a la totalitat del frotis seguint la metodologia utilitzada al Capítol 5 amb les imatges adquirides amb el sistema DM96. La totalitat del frotis es va classificar com a positiu quan el rati d'imatges cel·lulars individuals classificades com a positives era superior a un valor límit preestablert (per exemple, classificat com a frotis de SS perquè el rati de cèl·lules positives era superior al llindar preestablert). El valor llindar utilitzat ha estat el mateix que el presentat al Capítol 5 per tal de poder comparar els resultats. Amb aquest valor llindar es van calcular els següents paràmetres: sensibilitat, especificitat, valor predictiu negatiu (VPN), valor predictiu positiu (VPP) i exactitud. Consultar Capítol 5 i secció 5.2.2 per més detall.

7.3 Resultats

Aquest apartat es presenta d'acord amb les dues etapes detallades anteriorment.

7.3.1 Correlació dels descriptors i les corresponents escales numèriques

Els resultats d'aquesta primera part es resumeixen a les Taules 7.4 i 7.5 i a les Figures 7.3 – 7.5. Si recordem del Capítol 5, el dos descriptors quantitius relacionats amb la distribució de la cromatina de la LLC i les cèl·lules de Sézary obtinguts van ser la “*correlació del cian del nucli*” i la “*desviació estàndard (DE) de la corba granulomètrica del cian del nucli*”, respectivament. El descriptor quantitiu trobat per a la identificació dels grànuls citoplasmàtics azuròfils va ser l’“*asimetria de l’histograma del component u del citoplasma*” i el descriptor per a la detecció de les vellositats citoplasmàtiques, el descriptor geomètric “*hairiness*”.

Tots els descriptors mencionats han mostrat una bona correlació. Consultar l’Annex A secció 9.1.17 per a veure els resultats dels tests de correlació de Pearson per al descriptor *hairiness*. A la Taula 7.4 es mostren els coeficients de correlació de Pearson per als quatre descriptors (*correlació del cian del nucli*, *DE de la corba granulomètrica del cian del nucli*, *asimetria de l’histograma del component u del citoplasma* i *hairiness*) obtinguts amb els dos sistemes d’adquisició d’imatges (DM96 i BX43). Per veure més en detall els resultats de correlació en funció de cada tipus de cèl·lula limfoide, consultar Annex A secció 9.2 (Fig. 9.22 i Taula 9.4).

Taula 7.4: Resultats de la correlació de Pearson dels tres descriptors entre el sistema DM96 i el BX43. Es mostra el coeficient R de Pearson, la recta d’ajust, l’interval de confiança del 95 % i el p valor.

Descriptor	Coefficient R de Pearson	Recta d’ajust entre sistema BX43 i DM96	Interval de confiança del 95 %	p valor
<i>Correlació del cian del nucli ($\times 10^{-2}$)</i>	0,901	$y = 1,058x - 2,498$	0,896 - 0,905	$< 2^{-16}$
<i>DE de la corba granulomètrica del cian del nucli ($\times 10^{-3}$)</i>	0,810	$y = 0,782x - 4,902$	0,802 - 0,818	$< 2^{-16}$
<i>Asimetria de l’histograma del component u del citoplasma</i>	0,779	$y = 0,584x - 0,409$	0,770 - 0,787	$< 2^{-16}$
<i>Hairiness</i>	0,871	$y = 0,0000371x - 0,00236$	0,862 - 0,879	$< 2^{-16}$

La Taula 7.5 mostra les medianes i rangs interquartílics dels quatre descriptors corresponents al conjunt d'imatges adquirides amb el sistema BX43, separades per a cada tipus cel·lular limfoide objecte d'estudi.

Taula 7.5: Valors dels quatre descriptors que han permès la quantificació de les corresponents característiques citològiques qualitatives en el conjunt d'imatges adquirides amb el sistema BX43. Els valors mostren la mediana i el rang interquartílic entre parèntesi.

Característica qualitativa	Cromatina madura i condensada	Cromatina cerebriforme	Granulació citoplasmàtica	Velositats citoplasmàtiques
Regió d'interès	Nucli		Citoplasma	Regió externa
Grup cel·lular	<i>Correlació del cian del nucli ($\times 10^{-2}$)</i>	<i>DE de la corba granulomètrica del cian del nucli ($\times 10^{-3}$)</i>	<i>Asimetria de l'histograma del component u del citoplasma</i>	<i>Hairiness ($\times 10^{-3}$)</i>
Limfòcit	78,4 (3,4)	16,9 (5,2)	- 0,10 (0,42)	2,3 (3,0)
Limfòcit anormal de la leucèmia limfàtica crònica	80,6 (3,4)	18,7 (6,2)	- 0,14 (0,43)	1,0 (2,1)
Limfòcit vellós de la tricoleucèmia	78,1 (3,8)	18,1 (7,6)	- 0,21 (0,60)	13,5 (27,0)
Limfòcit vellós del limfoma de la zona marginal esplènic	78,8 (3,4)	20,1 (5,7)	- 0,43 (0,36)	17,9 (34,0)
Limfòcit anormal del limfoma fol·licular	77,8 (4,3)	16,0 (5,4)	0,00 (0,00)	0,0 (0,0)
Cèl·lula del limfoma del mantell morfologia blàstica	77,6 (4,7)	15,6 (4,1)	- 0,22 (0,81)	0,7 (3,1)
Cèl·lula del limfoma del mantell morfologia clàssica	78,3 (3,9)	17,8 (5,4)	- 0,22 (0,50)	0,8 (1,8)
Cèl·lula plasmàtica	78,0 (5,4)	13,8 (7,5)	- 0,13 (0,59)	1,5 (3,1)
Prolimfòcit B	76,9 (5,8)	16,4 (5,1)	- 0,37 (0,42)	2,1 (4,0)
Prolimfòcit T	77,0 (3,5)	18,1 (5,4)	0,00 (0,48)	3,3 (4,6)
Cèl·lula de Sézary variant clàssica	79,9 (3,4)	10,8 (3,7)	- 0,31 (0,48)	1,7 (3,5)
Cèl·lula de Sézary variant Lutzner o petita	80,2 (4,3)	13,7 (6,3)	- 0,41 (0,53)	2,8 (6,2)
Limfòcit gran granular	79,6 (3,9)	17,7 (5,9)	0,28 (0,82)	0,5 (1,3)
Cèl·lula limfoide reactiva	80,2 (4,1)	17,1 (5,7)	- 0,37 (0,93)	0,6 (1,6)
Blast limfoide	76,2 (4,7)	18,0 (6,3)	- 0,15 (0,57)	0,3 (0,7)

La *correlació del component cian del nucli (CCN)* ha permès la quantificació de la cromatina madura i condensada coneguda clàssicament com *grûmelée* i generalment observada al nucli dels limfòcits de la LLC, com s'ha detallat al Capítol 5. La *correlació del component cian del nucli* és un descriptor estadístic de segon ordre, calculat a partir de la GLCM. En concret mesura com es correlaciona un píxel amb el seu veí i en aquest cas la regió d'interès és el nucli.

A la Figura 7.3 es poden veure els diagrames de caixes de la *correlació del component cian del nucli* per a tots els grups de 15 cèl·lules del conjunt d'imatges adquirides amb el sistema BX43 i els valors quantitius corresponents es detallen a la Taula 7.5. La *correlació del component cian del nucli* mostra diferències significatives ($p < 0,005$) entre la LLC ($80,6 \times 10^{-2}$) i amb cadascun dels 14 grups de cèl·lules restants com el LF ($77,8 \times 10^{-2}$), el LM-C ($78,3 \times 10^{-2}$) i els limfòcits normals ($78,4 \times 10^{-2}$), a excepció de les CLR ($80,2 \times 10^{-2}$) i la variant de cèl·lules petites de la síndrome de Sézary ($80,2 \times 10^{-2}$).

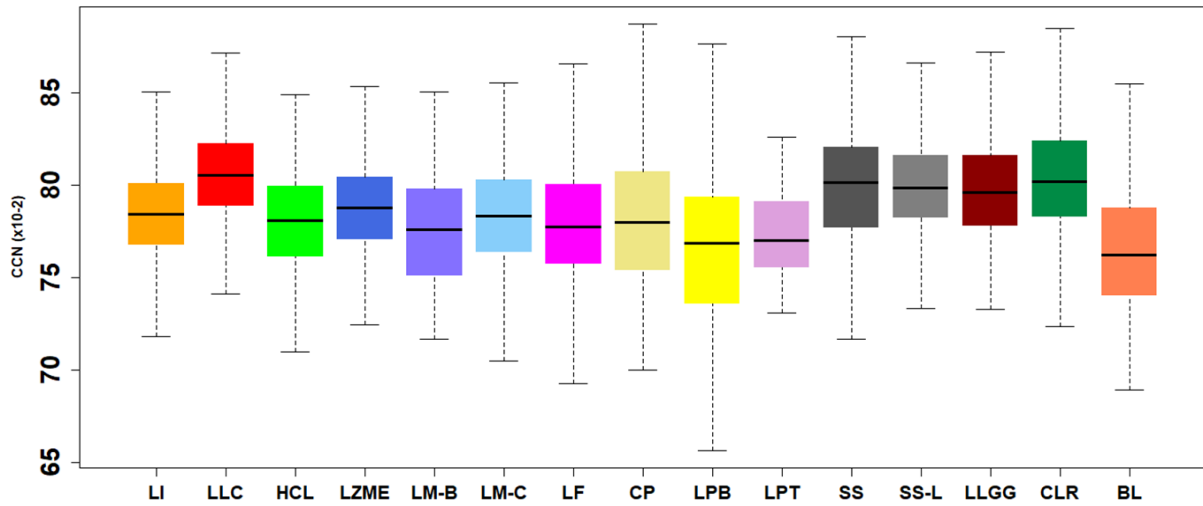


Figura 7.3: Diagrames de caixes obtinguts per a la *correlació del component cian del nucli (CCN)* per als 15 grups de cèl·lules limfoides.

La *DE* de la corba granulomètrica del component cian del nucli, la qual és un descriptor granulomètric, al Capítol 5 va resultar ser un descriptor de color i textura específic per a la quantificació de la cromatina cerebriforme, característica del nucli de les cèl·lules de Sézary. A la Taula 7.5 se'n mostren els valors quantitius. A la Figura 7.4 es poden veure els corresponents diagrames de caixes per als grups de 15 cèl·lules limfoides del conjunt d'imatges adquirides amb el sistema BX43.

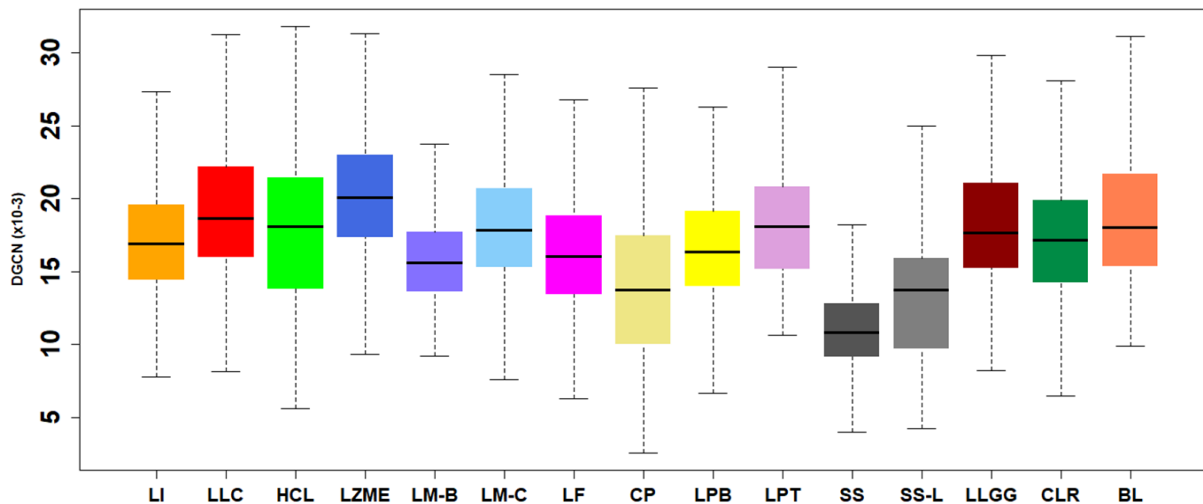


Figura 7.4: Diagrames de caixes obtinguts per a la *desviació estàndard de la corba granulomètrica del component cian del nucli (DGCN)* per als 15 grups de cèl·lules limfoides.

Els valors de les medianes de la *DE de la corba granulomètrica del component cian del nucli* per a les dues classes de cèl·lules de Sézary, SS ($10,8 \times 10^{-3}$) i SS-L ($13,7 \times 10^{-3}$), són significativament inferiors respecte a tots els grups cel·lulars restants com l'HCL ($18,1 \times 10^{-3}$), el LM-B ($15,6 \times 10^{-3}$), els prolimfòcits T ($18,1 \times 10^{-3}$) i el LF ($16,0 \times 10^{-3}$) ($p < 0,01$). Només les cèl·lules plasmàtiques ($13,7 \times 10^{-3}$) no han presentat valors de *DE de la corba granulomètrica del component cian del nucli* significativament diferents en comparació amb la SS-L.

Referent a les característiques citoplasmàtiques, l'*asimetria de l'histograma del component u del citoplasma* ha mostrat una gran especificitat per a la mesura de la granulació citoplasmàtica (explicat al Capítol 5). Aquest descriptor és un paràmetre estadístic de primer ordre que té en compte els píxels violeta del component de color *u* (de l'espai de color *Luv*).

La Figura 7.5 mostra els corresponents diagrames de caixes per als 15 grups cel·lulars del conjunt d'imatges adquirides amb el sistema BX43 i la Taula 7.5 en detalla els seus valors. Es pot veure que la LLGG mostra els valors significativament més alts d'*asimetria de l'histograma del component u del citoplasma* (0,28) quan es compara entre els 14 grups restants ($p < 0,001$).

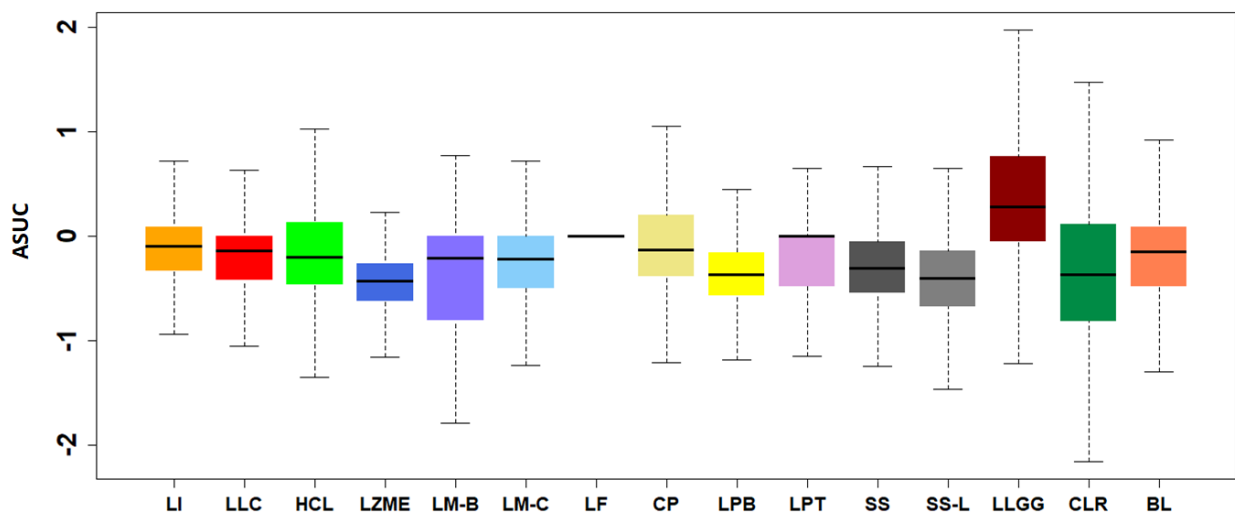


Figura 7.5: Diagrames de caixes obtinguts per a l'*asimetria de l'histograma del component u del citoplasma* (ASUC) per als 15 grups de cèl·lules limfoides.

La Figura 7.6 mostra els diagrames de caixes per als 15 grups cel·lulars del conjunt d'imatges adquirides amb el sistema BX43 i la Taula 7.5 en detalla els seus valors. Com es pot observar, l'HCL i el LZME són els grups que presenten els valors més alts (0,0135 i 0,0179, respectivament) i mostren diferències significatives en comparació amb els 14 restants grups cel·lulars analitzats ($p < 0,0001$, Fig. 7.6). Recordem que el descriptor hairiness en aquest cas no té unitats d'acord amb la seva reformulació per adaptar-ho al sistema BX43 (secció 6.3.1 del Capítol 6).

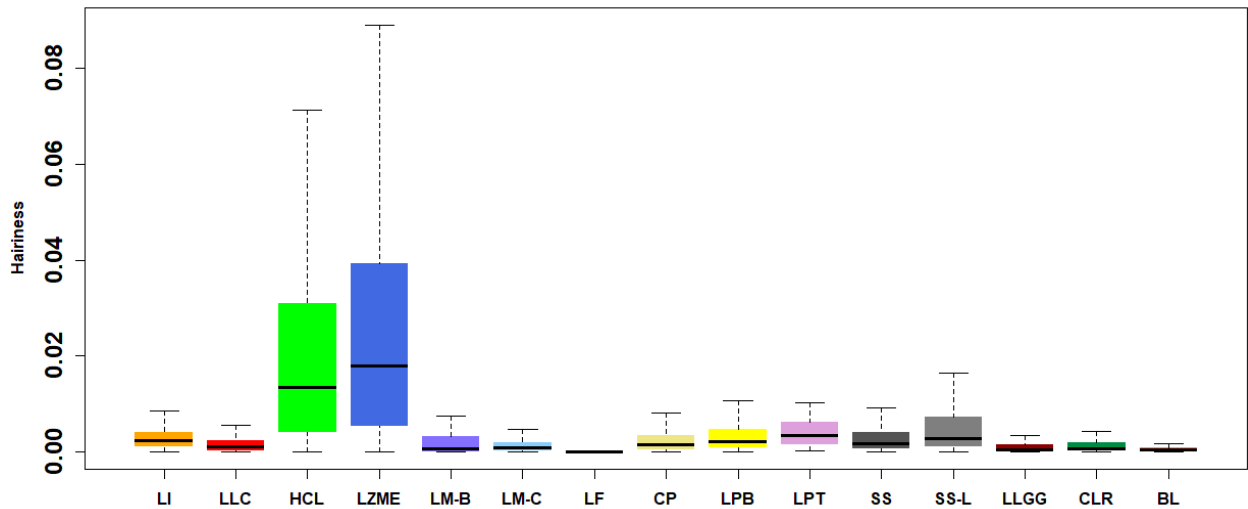


Figura 7.6: Diagrames de caixes obtinguts per al *hairiness* per als 15 grups de cèl·lules limfoides.

7.3.2 Avaluació dels descriptors quantitatius per a la classificació de frotis

En aquesta segona part de l'estudi, per a cada un dels quatre descriptors avaluats, es detallen els pacients inclosos, els valors de tall discriminador obtinguts així com el rendiment de la classificació dels frotis. Els resultats d'aquesta segona part es resumeixen a les Taules 7.6 i 7.7 i a les Figures 7.7 – 7.10.

Tal i com s'ha detallat a l'apartat de Material i mètodes, per a l'avaluació de cada un dels descriptors, es va partir del valor llindar obtingut amb el sistema DM96 utilitzant corbes ROC i mitjançant la recta d'ajust calculada a l'anterior apartat (veure Taula 7.4), s'ha obtingut el *cut-off* equivalent pel sistema BX43 (veure Taula 7.6).

Taula 7.6: Recta d'ajust, cut-offs d'ambdós sistemes d'adquisició d'imatges i valor llindar del sistema DM96 utilitzats en l'avaluació de l'eficiència diagnòstica dels tres descriptors.

Descriptor	Recta d'ajust	Cut-off DM96	Cut-off BX43	Llindar DM96
Correlació del cian del nucli ($\times 10^{-2}$)	$y = 1,058x - 2,498$	78	80	45 %
DE de la corba granulomètrica del cian del nucli ($\times 10^{-3}$)	$y = 0,782x - 4,902$	26,1	15,5	78 %
Asimetria de l'histograma del component u del citoplasma	$y = 0,584x - 0,409$	1,07	0,22	30 %
Hairiness	$y = 0,0000371x - 0,00236$	3,28 μm^2 (332 píxels)	0,0107	63 %

A la Taula 7.7 es detalla, per a cada descriptor, els seus valors calculats de les diferents imatges incloses així com el percentatge d'imatges positives.

Taula 7.7: Nombre total d'imatges corresponents a cada diagnòstic, mediana, rang interquartílic (RIQ) i imatges positives per a cada descriptor obtingudes a l'avaluació de la classificació de frotis.

Descriptor cromatina madura i condensada ($\times 10^{-2}$)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges LLC	863	80,56	3,43	503 (58 %)
Imatges no LLC	1.728	78,88	4,07	-
Descriptor cromatina cerebriforme ($\times 10^{-3}$)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges SS	218	12,34	6,67	158 (72 %)
Imatges no SS	989	17,38	6,21	-
Descriptor granulació citoplasmàtica				
	Imatges	Mediana	RIQ	Imatges positives
Imatges LLGG	435	0,25	0,57	255 (57 %)
Imatges no LLGG	1.430	-0,14	0,47	-
Descriptor vellositats citoplasmàtiques ($\times 10^{-3}$)				
	Imatges	Mediana	RIQ	Imatges positives
Imatges HCL/LZME	321	15,0	31,00	225 (70 %)
Imatges no HCL/LZME	985	1,1	2,10	-

LLC, leucèmia limfàtica crònica; SS, Síndrome de Sézary; HCL, tricoleucèmia; LZME, limfoma de la zona marginal esplènica; LLGG, leucèmia de limfòcits grans granulars.

El descriptor de la *correlació del cian del nucli* s'ha calculat en frotis de dos grups diferents. El primer està format per 32 pacients amb LLC i el segon per 57 pacients, en els quals els frotis contenen limfòcits amb cromatina madura però no condensada (38 individus sans, 12 LF, 4 SS-L i 3 SS).

El valor de tall òptim obtingut mitjançant la recta d'ajust és de 80×10^{-2} (Taula 7.6), entenent que una cèl·lula limfoide serà considerada com que conté cromatina madura i condensada quan el valor de la *correlació del cian del nucli* sigui superior a 80×10^{-2} . D'altra banda, un frotis serà considerat positiu per a cromatina madura i condensada quan el rati entre el número d'imatges amb un valor de *correlació del cian del nucli* superior al *cut-off* i el número total d'imatges sigui superior a 0,45 (valor llindar del 45 % obtingut amb el sistema DM96, Taula 7.6).

Un total de 23/32 (71,9 %) de frotis de LLC s'han classificat a la classe correcta, així com en 48/57 (84,2 %) dels frotis del segon grup (limfòcits sense cromatina condensada). A la Taula 7.7 es mostren els resultats desglossats per imatges i s'observa que el 58 % d'imatges de limfòcits LLC mostren un valor del descriptor superior a 80×10^{-2} .

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 7.67:

- Sensibilitat: 71,9 %
- Especificitat: 84,2 %
- VPP: 71,9 %
- VPN: 84,2 %
- Exactitud: 79,8 %

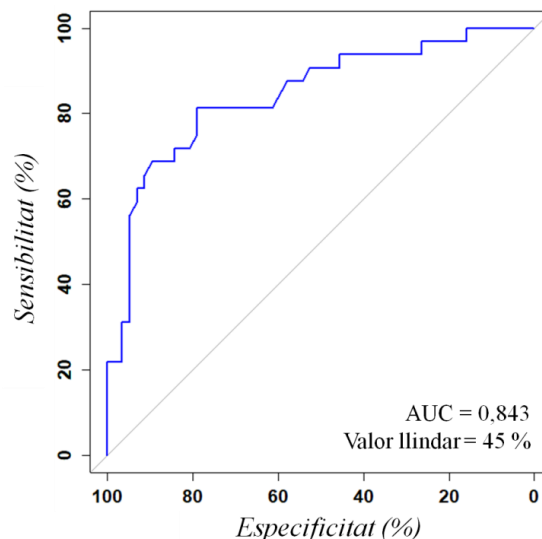


Figura 7.7: Corba ROC de la *correlació del cian del nucli*.

La *DE de la corba granulomètrica del cian del nucli* s'ha calculat en frotis d'un primer conjunt format per set pacients amb SS i d'un segon grup de 38 individus sans. La selecció dels limfòcits normals és degut a que també mostren una cromatina madura com les cèl·lules de Sézary.

El valor de tall òptim obtingut mitjançant la recta d'ajust obtingut és de $15,5 \times 10^{-3}$ (Taula 7.6), el que significa que les cèl·lules amb un valor de *DE de la corba granulomètrica del cian del nucli* inferior, contenen nuclis amb cromatina cerebriforme. Un frotis es considera positiu per a cromatina cerebriforme quan el rati entre el número d'imatges amb el descriptor per sota de $15,5 \times 10^{-3}$ i el número total d'imatges del frotis sigui major a 0,78 (valor llindar del 78 % obtingut amb el sistema DM96, Taula 7.6).

Un total de 4/7 (57,1 %) dels frotis amb imatges de cèl·lules de Sézary s'han classificat correctament com que pertanyen a pacient amb SS. A més a més, tots els frotis del segon grup (38/38) s'han classificat a la categoria veritable. A la Taula 7.7 es mostren els resultats desglossats per imatges i s'observa que el 72 % d'imatges de cèl·lules de Sézary mostren un valor del descriptor inferior a $15,5 \times 10^{-3}$.

Els resultats de l'avaluació de l'eficiència diagnòstica es mostren a continuació i la corresponent corba ROC a la Fig. 7.8:

- Sensibilitat: 57,1 %
- Especificitat: 100 %
- VPP: 100 %
- VPN: 92,7 %
- Exactitud: 93,3 %

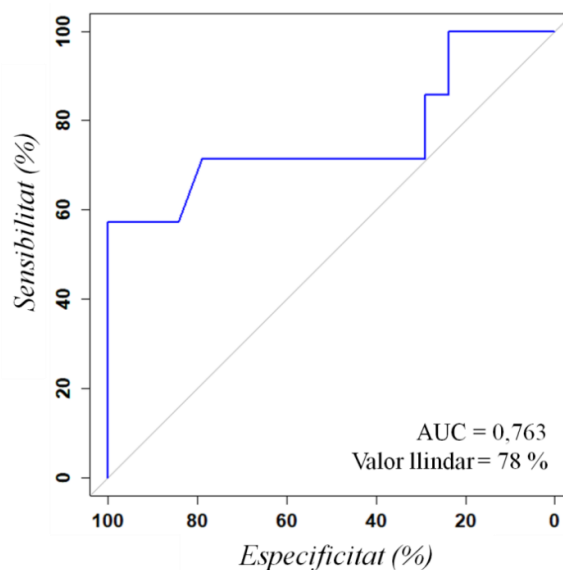


Figura 7.8: Corba ROC de la *DE de la corba granulomètrica del cian del nucli*.

L'asimetria de l'histograma del component u del citoplasma s'ha avaluat en frotis d'un primer grup de 15 pacients amb LLGG i d'un segon grup de 53 pacients, els quals normalment no contenen granulació azuròfila al citoplasma (5 HCL, 10 LZME i 38 individus sans).

El valor de tall òptim obtingut mitjançant la recta d'ajust obtingut és de 0,22 (Taula 7.6), el que significa que les cèl·lules limfoides amb un valor superior d'asimetria de l'histograma del component u del citoplasma, contindran grànuls en el seu citoplasma. Per tant, un frotis es considera positiu quan el rati entre el número d'imatges amb el descriptor per sobre de 0,22 i el número total d'imatges del frotis sigui major a 0,30 (valor llindar del 30 % obtingut amb el sistema DM96, Taula 7.6).

Un total de 14/15 frotis (93,3 %) de pacients amb LLGG s'han classificat com que contenien limfòcits grans granulars i 49/53 (92,5 %) dels frotis corresponents a l'altre grup de pacients amb múltiples patologies, han estat correctament classificats (com a no LLGG). A la Taula 7.7 es mostren els resultats desglossats per imatges i s'observa que el 57 % d'imatges de limfòcits grans granulars mostren un valor del descriptor superior a 0,22.

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 7.9:

- Sensibilitat: 93,3 %
- Especificitat: 92,5 %
- VPP: 77,8 %
- VPN: 98,0 %
- Exactitud: 92,6 %

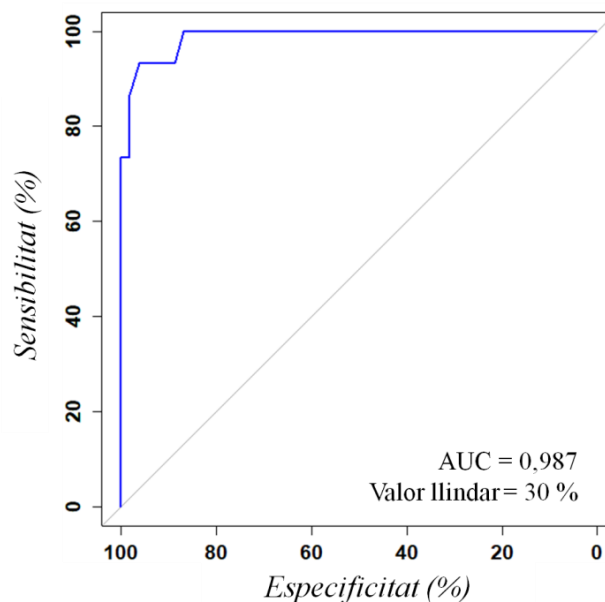


Figura 7.9: Corba ROC de l'asimetria de l'histograma del component u del citoplasma.

El descriptor *hairiness* s'ha calculat en frotis de sang perifèrica que pertanyien a un primer grup de 14 pacients, en els quals el diagnòstic era HCL (7) o LZME (7). El segon grup de frotis, format per 41 pacients, contenien limfòcits reactius (10 pacients), limfòcits grans granulars (15 pacients amb LLGG) i limfòcits normals (16 individus sans). S'han escollit aquests grups cel·lulars perquè tenen una mida similar a la dels limfòcits vellosos.

El valor de tall òptim obtingut mitjançant la recta d'ajust obtingut és de $10,7 \times 10^{-3}$ (Taula 7.6), el que significa que cèl·lules amb un valor de *hairiness* per sobre, són considerades com a limfòcits vellosos. Per tant, un frotis es considera positiu quan el rati entre el número d'imatges amb el descriptor per sobre de $10,7 \times 10^{-3}$ i el número total d'imatges del frotis sigui major a 0,63 (valor llindar del 63 % obtingut amb el sistema DM96, Taula 7.6).

Utilitzant aquest descriptor, un total d'11 frotis que contenen limfòcits vellosos s'han classificat a la classe correcta (78,6 %). A més a més, els valors de *hairiness* en el 100 % de les imatges del segon grup sense contenir limfòcits vellosos (985), han estat inferiors al *cut-off*. En relació a les imatges, el 70,1 % d'imatges de limfòcits vellosos mostren un valor del descriptor superior a $10,7 \times 10^{-3}$.

Els resultats de l'avaluació es mostren a continuació i la corresponent corba ROC a la Fig. 7.10:

- Sensibilitat: 78,6 %
- Especificitat: 100 %
- VPP: 100 %
- VPB: 93,2 %
- Exactitud: 94,5%

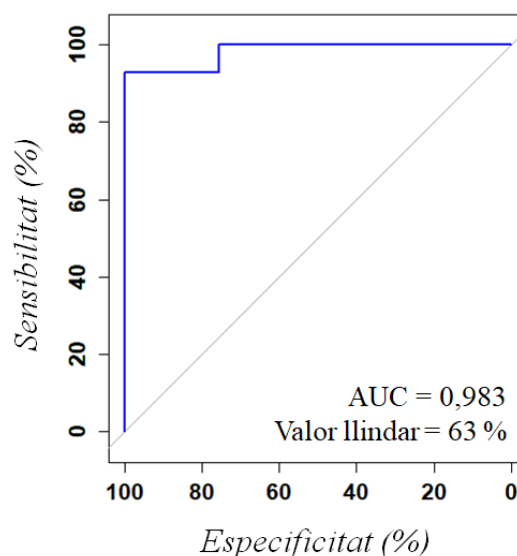


Figura 7.10: Corba ROC del descriptor *hairiness*.

7.4 Discussió

Com s'ha comentat anteriorment, el frotis de sang perifèrica proporciona les primeres evidències morfològiques de diagnòstic, per tant, encara segueix sent una important eina de diagnòstic (39). Una mesura reproducible, precisa i objectiva de la morfologia cel·lular és beneficiosa per evitar variacions subjectives i entre diferents observadors (131). D'altra banda, les cèl·lules limfoides anormals són les més difícils de reconèixer només per característiques morfològiques (4) i no tots els laboratoris disposen del sistema DM96.

A l'estudi presentat en aquest Capítol 7, s'ha utilitzat la metodologia introduïda al Capítol 5 basada en l'anàlisi d'imatges per reproduir els resultats obtinguts amb els quatre descriptors quantitius citològics presentats al treball del Capítol 5 amb el microscopi convencional (sistema BX43).

El present estudi ha inclòs un conjunt de 8.946 imatges adquirides mitjançant el microscopi Olympus® BX43 i la càmera digital DP73 corresponents a limfòcits normals, cèl·lules limfoides reactives i blàstiques i 12 tipus de limfòcits B o T anormals (veure Taula 7.1). De cada una de les imatges, s'han extret els quatre descriptors per a mesurar les següents característiques citològiques: la cromatina madura i condensada, la cromatina cerebriforme, la granulació citoplasmàtica i les vellositats citoplasmàtiques.

Dels quatre descriptors següents, 1) la *correlació del cian del nucli*, 2) la *desviació estàndard de la corba granulomètrica del cian del nucli*, 3) l'*asimetria de l'histograma del component de color u de la regió del citoplasma* i 4) *Hairiness*, s'han obtingut noves escales numèriques corresponents a les imatges del sistema BX43 (Taula 7.5 i Figures 7.3-7.6). Es pot observar que tot i que varia els rangs de valors, els resultats obtinguts són similars respecte als obtinguts utilitzant les imatges adquirides al DM96.

Si comparem els resultats de l'avaluació de l'eficiència diagnòstica dels quatre descriptors entre ambdós sistemes d'adquisició d'imatges (DM96 i BX43), prenent els mateixos *cut-offs* i valors llindars, observem que la *correlació del cian del nucli* (descriptor cromatina madura i condensada) amb el sistema DM96 mostrava una sensibilitat i especificitat superior al 92 % mentre que amb el sistema BX43, la sensibilitat ha estat del 72 % i l'especificitat del 84 %.

Pel que fa a l'*asimetria de l'histograma del component de color u de la regió del citoplasma* (descriptor granulació citoplasmàtica), aquesta mostra una sensibilitat i especificitat similars ja que ambdues es troben per sobre del 92 % en ambdues metodologies.

En relació a la *desviació estàndard de la corba granulomètrica del cian del nucli*, descriptor de la cromatina cerebriforme, aquesta amb les imatges adquirides amb el sistema DM96 presentava una sensibilitat i una especificitat del 100 % tot i que només es van incloure dos pacients amb la síndrome de Sézary. Per contra, amb el sistema BX43 s'han inclòs fins a set pacients de manera que la sensibilitat ha disminuït al 57 % tot i que l'especificitat s'ha mantingut en el 100 %.

Per últim, el descriptor *hairiness* ha mostrat una elevada especificitat (100 %) pels limfòcits vellosos de frotis d'HCL i LZME, com ja succeïa amb el sistema DM96.

El descriptor de la cromatina madura i condensada es basa en el càlcul de la GLCM mentre que el de la granulació azuròfila citoplasmàtica es fonamenta en l'histograma del component de color, de manera que la variació en la mida de la cèl·lula pot explicar les diferències en la sensibilitat i especificitat dels descriptors sobre les imatges adquirides amb el sistema BX43. En comparació amb les imatges adquirides amb el sistema DM96, les cèl·lules de les imatges adquirides amb el BX43 presenten valors inferiors de mida. Com s'ha comentat a la discussió del Capítol 6, el descriptor *hairiness* s'ha hagut de reformular per tal d'ajustar el seu càlcul a les imatges obtingudes amb el sistema BX43 i mitjançant la normalització de la mida, s'han pogut tenir resultats discriminants per les cèl·lules limfoïdes velloses molt similars.

Fins al coneixement de l'autor, no s'han publicat prèviament descriptors quantitius per a detectar patrons de cromatina madura i condensada i cerebriforme. El Capítol 5 i el present estudi representen una primera aproximació per a la seva quantificació amb dos sistemes d'adquisició diferents.

En resum, l'anàlisi de la morfologia cel·lular és un assumpte clau per a la identificació i la classificació d'anomalies i ha rebut molta atenció recentment amb creixents demandes tant en la bioinformàtica com en les aplicacions biomèdiques. En el present estudi, s'han reproduït els resultats obtinguts amb els quatre descriptors quantitius citològics presentats al treball del Capítol 5 amb les imatges adquirides amb el microscopi convencional (sistema BX43). Concretament, els descriptors per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica, la síndrome de Sézary, la leucèmia de limfòcits grans granulars, la tricoleucèmia i el limfoma de la zona marginal esplènic.

7.5 Conclusions

El treball actual és el resultat d'aplicar la mateixa metodologia presentada al Capítol 5 però utilitzant una nova font d'adquisició d'imatges, el sistema BX43, com al Capítol 6. A l'estudi s'han utilitzat 8.946 imatges adquirides mitjançant el microscopi Olympus® BX43 i la càmera digital DP73 corresponents a limfòcits normals, cèl·lules limfoides reactives i blàstiques i 12 tipus de limfòcits B o T anormals.

De cada una de les imatges, s'han extret els quatre descriptors proposats al Capítol 5 per a mesurar les següents característiques citològiques: la cromatina madura i condensada, la cromatina cerebriforme, la granulació citoplasmàtica i les vellositats citoplasmàtiques.

En una segona etapa, els valors dels quatre descriptors s'han correlacionat entre ambdós sistemes d'adquisició (DM96 i BX43) i per últim, s'ha realitzat una avaluació de l'eficiència diagnòstica dels quatre descriptors (*correlació del cian del nucli*, *desviació estàndard de la corba granulomètrica del cian del nucli*, *asimetria de l'histograma del component de color u de la regió del citoplasma* i *hairiness*) aplicats a les imatges adquirides amb el microscopi convencional utilitzant els valors l·lindars obtinguts al Capítol 5 i s'han inclòs un total de 174 frotis de sang perifèrica de nous pacients.

Sobre les imatges adquirides amb el sistema BX43, la *correlació del cian del nucli* per a la detecció de la cromatina madura i condensada dels limfòcits de la leucèmia limfàtica crònica ha presentat una sensibilitat del 72 % i una especificitat del 84 %. La *desviació estàndard de la corba granulomètrica del component cian del nucli* mostra una sensibilitat del 57 % i una especificitat del 100 % per a la detecció de la cromatina cerebriforme característica de les cèl·lules de Sézary sobre les imatges adquirides amb el sistema BX43.

Sobre les imatges adquirides amb el sistema BX43 i utilitzant el descriptor denominat *asimetria de l'histograma del component u del citoplasma* per la detecció dels grànuls azuròfils al citoplasma dels limfòcits grans i granulars, s'ha obtingut una sensibilitat i especificitat superior al 92 %. Mitjançant una nova formulació del descriptor *hairiness* per a les imatges del sistema BX43, la qual normalitza la mida de la cèl·lula i adapta els l·lindars d'intensitat de verd de la regió externa de la cèl·lula, s'aconsegueix una elevada especificitat (100 %) pels limfòcits vellosos, com ja succeïa amb el sistema DM96.

Es conclou, per tant, que una avaluació basada en la morfologia més objectiva de l'anàlisi de limfòcits anormals a través de l'anàlisi d'imatges és factible amb descriptors quantitius. L'aplicabilitat dels descriptors descrits aquí pot millorar la reproductibilitat i l'objectivitat de l'avaluació morfològica dels limfòcits anormals gràcies a l'obtenció dels nous descriptors específics de la granulació citoplasmàtica o de l'organització de la cromatina.

Capítol 8

8 Conclusions i contribucions de la present Tesi Doctoral. Perspectives de futur

Les conclusions derivades d'aquesta investigació es resumeixen a la secció 8.1. Les contribucions principals de la present Tesi es destaquen a la secció 8.2. Finalment, algunes de les perspectives de futur sorgides de la present recerca es descriuen a la secció 8.3.

8.1 Conclusions

A la present *Tesi Doctoral* s'han obtingut nous descriptors quantitativs mitjançant processament digital d'imatges i aprenentatge automàtic per a la diferenciació de cèl·lules limfoïdes normals, reactives, anormals i blàstiques de sang perifèrica, contribuint a una anàlisi més objectiva de la citologia sanguínia.

A continuació s'anomenen les principals conclusions derivades de la present recerca, les quals es divideixen en dos subapartats corresponents als objectius específics proposats al Capítol 1 i que han donat lloc a dues publicacions en revistes de caire internacional. Finalment, un tercer subapartat recull les conclusions de la comparació dels resultats obtinguts amb els dos sistemes d'adquisició d'imatges.

8.1.1 Conclusions derivades de la primera publicació

Les conclusions obtingudes són les següents:

1. S'han identificat els 20 descriptors geomètrics, de color i de textura més eficients per a la diferenciació entre nou grups de cèl·lules limfoïdes anormals circulant a sang perifèrica de pacients amb diferents tipus de neoplàsies limfoïdes B i T, blasts limfoïdes, cèl·lules limfoïdes reactives i limfòcits normals.
2. Els descriptors obtinguts tenen formulacions quantitatives explícites i admeten interpretacions qualitatives en termes de característiques morfològiques visuals.
3. Els descriptors més rellevants per al reconeixement de les diferents cèl·lules limfoïdes incloses a l'estudi són geomètrics i, en ordre d'importància, són els següents: 1) la *relació nucli/citoplasma*, 2) el *perímetre del nucli* i 3) el *diàmetre de la cèl·lula*.
4. Dels 20 descriptors obtinguts, cinc són geomètrics i 15 de color i textura, dels quals 13 són estadístics i dos són granulomètrics. Només tres espais de color (CMGN, RGB i HSV) i sis components de color (el blau, el verd, el magenta, el negre, el cian i la saturació) dels 19 utilitzats es troben involucrats en la formulació dels descriptors de color i de textura.
5. Un total de 19 descriptors identificats són significatius per a discriminar els limfòcits anormals del limfoma fol·licular amb expressió perifèrica, respecte als del limfoma del mantell.
6. Un total de 18 descriptors resulten significatius per a diferenciar els limfoblasts de les cèl·lules limfoïdes reactives.
7. Un total de 17 descriptors mostren diferències significatives entre els prolimfòcits T i les cèl·lules de Sézary. Aquesta diferenciació és particularment rellevant donat que ambdues són cèl·lules anormals de neoplàsies limfoïdes T que comparteixen similituds morfològiques.

8. Un total dels 15 descriptors són significativament discriminants per a la diferenciació dels limfòcits anormals de la leucèmia limfàtica crònica respecte als limfòcits normals.
9. Un total de 16 descriptors permeten diferenciar els limfòcits amb prolongacions citoplasmàtiques de la tricoleucèmia respecte als del limfoma de la zona marginal esplènica leucemitzat i entre prolimfòcits de tipus B i T.
10. Els cinc descriptors de color i textura següents mostren significació estadística quan es fan comparacions entre les cèl·lules limfoides reactives (benignes) amb cadascun dels diferents tipus de limfòcits anormals (malignes): 1) *Desviació estàndard de l'histograma del component blau de la cèl·lula*, 2) *Entropia del component magenta de la cèl·lula*, 3) *Mitjana de l'histograma del component blau del nucli*, 4) *Information measure of correlation 1 del component cian del nucli* i 5) *Desviació estàndard del component de la saturació de la cèl·lula*.

8.1.2 Conclusions derivades de la segona publicació

El segon article s'ha basat en cercar descriptors quantitius citològics per a la detecció d'anomalies morfològiques específiques de cèl·lules anormals de determinades neoplàsies limfoides. Aquests descriptors s'han buscat tant a nivell del nucli, ja que el patró de densitat de la cromatina nuclear és una característica molt important per al diagnòstic morfològic i difícil de quantificar, com a nivell del citoplasma.

Respecte el nucli, s'han buscat descriptors quantitius per a la cromatina madura i condensada dels limfòcits anormals de la leucèmia limfàtica crònica o la cromatina cerebriforme de la síndrome de Sézary. A nivell del citoplasma, s'ha estudiat la detecció de les prolongacions citoplasmàtiques dels limfòcits de la tricoleucèmia i del limfoma de la zona marginal esplènica amb expressió a sang perifèrica, i la granulació azuròfila característica dels limfòcits de la leucèmia de limfòcits grans granulars.

Les conclusions són les següents:

1. Un total de quatre descriptors (de 2.678 analitzats) s'han proposat per mesurar quatre característiques citològiques: cromatina madura i condensada, cromatina cerebriforme, granulació citoplasmàtica i vellositats citoplasmàtiques i s'han establert escales numèriques objectives per la seva quantificació.
2. La detecció i quantificació de la cromatina madura i condensada característica de les cèl·lules de la leucèmia limfàtica crònica ha estat possible mitjançant un descriptor basat en la matriu de co-ocurrència del nivell de gris, la *correlació del cian del nucli*. Aquest descriptor ha mostrat ser específic (especificitat del 92,3 %) i sensible (sensibilitat del 92,6 %) per a la detecció dels limfòcits anormals de la leucèmia limfàtica crònica independentment del sistema d'adquisició de les imatges.

3. S'ha detectat i quantificat la cromatina cerebriforme característica de les cèl·lules de Sézary (tant la variant clàssica com la Lutzner) mitjançant un descriptor granulomètric: la *desviació estàndard de la corba granulomètrica del component cian del nucli*, la qual exhibeix valors d'especificitat i sensibilitat del 100 %.
4. El descriptor estadístic de primer ordre denominat *asimetria de l'histograma del component u del citoplasma* ha demostrat ser útil per detectar els grànuls azuròfils característics de les cèl·lules limfoides anormals de la leucèmia de limfòcits grans granulars. La sensibilitat i l'especificitat ha estat del 100 %.
5. S'ha demostrat que el descriptor *hairiness* permet quantificar el perfil del citoplasma de les cèl·lules limfoides velloses de la tricoleucèmia i del limfoma de la zona marginal esplènic mostrant una molt bona sensibilitat i especificitat sobre les imatges adquirides amb el CellaVision DM96.
6. S'ha demostrat que els quatre descriptors obtinguts per a la detecció de característiques morfològiques específiques de determinades cèl·lules limfoides neoplàsiques B i T són eficients (exactitud superior al 90 %) per al reconeixement d'aquestes cèl·lules en el frotis sanguini, pel que podrien contribuir al diagnòstic d'aquestes entitats. Dels 210 frotis de nous pacients utilitzats en l'avaluació, el 96 % s'han reconegut correctament.

8.1.3 Conclusions derivades de l'estudi comparatiu dels dos sistemes d'adquisició d'imatges

1. Tant en el cas de les imatges obtingudes amb el sistema CellaVision DM96 com amb el sistema BX43, els 20 descriptors identificats en aquesta Tesi permeten discriminar entre els diferents tipus de cèl·lules limfoides incloses a l'estudi.
2. Sobre les imatges adquirides amb el sistema BX43, la *correlació del cian del nucli* per a la detecció de la cromatina madura i condensada dels limfòcits de la leucèmia limfàtica crònica ha presentat una sensibilitat del 72 % i una especificitat del 84 %.
3. La *desviació estàndard de la corba granulomètrica del component cian del nucli* mostra una sensibilitat del 57 % i una especificitat del 100 % per a la detecció de la cromatina cerebriforme característica de les cèl·lules de Sézary sobre les imatges adquirides amb el sistema BX43.
4. Sobre les imatges adquirides amb el sistema BX43, i utilitzant el descriptor denominat *asimetria de l'histograma del component u del citoplasma* per a la detecció dels grànuls azuròfils al citoplasma dels limfòcits grans i granulars, s'ha obtingut una sensibilitat i especificitat superior al 92 %.

5. Mitjançant una nova formulació del descriptor *hairiness* per a les imatges del sistema BX43, la qual normalitza la mida de la cèl·lula i adapta els llindars d'intensitat de verd de la regió externa de la cèl·lula, s'aconsegueix una elevada especificitat (100 %) pels limfòcits vellosos, com ja succeïa amb el sistema DM96.

8.1.4 Consideracions finals

Com a resum final podem dir que els resultats de la recerca realitzada en la present Tesi proporcionen objectivitat en l'avaluació morfològica de cèl·lules limfoides normals, reactives (infeccions), blàstiques (leucèmies limfoides agudes) i anormals (neoplàsies limfoides B o T). La diferenciació d'aquest elevat nombre de cèl·lules limfoides mitjançant descriptors quantitius, segons el nostre coneixement, no ha estat publicat prèviament a la literatura. D'altra banda, l'obtenció de descriptors quantitius per cèl·lules limfoides anormals específiques de determinades neoplàsies limfoides B i T amb expressió a sang perifèrica és una novetat que no ha estat reportada prèviament i que podria contribuir a facilitar la detecció d'aquestes malalties.

8.2 Contribucions de la tesi

8.2.1 Articles

1. Puigví L, Merino A, Alférez S, Acevedo A, Rodellar J. *New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood*. J Clin Pathol. 2017;70:1038-1048. doi: 10.1136/jclinpath-2017-204389.
2. Puigví L, Merino A, Alférez S, Boldú L, Acevedo A, Rodellar J. *Quantitative cytologic descriptors to differentiate CLL, Sézary, granular and villous lymphocytes through image analysis*. Am J Clin Pathol. 2019;5;152:74-85. doi: 10.1093/ajcp/aqz025.

8.2.2 Conferències i comunicacions

1. L. Puigví, A. Merino, S. Alférez, A. Acevedo, J. Rodellar – *Analysis of the most relevant quantitative features for the automatic differentiation of normal, reactive, abnormal lymphoid cell images and blast cell images from peripheral blood* - International Society of Laboratory of Hematology (ISLH) – Milà (Itàlia), maig 2016. International Journal of Laboratory Hematology, Issue S2, vol 38, pàg 86. Wiley. doi: 10.1111/ijlh.12567. **Comunicació oral.**
2. S. Alférez, A. Merino, A. Acevedo, L. Puigví, J. Rodellar – *Automatic classification of normal, reactive lymphocytes, abnormal lymphoid cells and blast cells* - International Society of Laboratory of Hematology (ISLH) – Milà (Itàlia), maig 2016. International Journal of Laboratory Hematology, Issue S2, vol 38, pàg 6-7. Wiley. doi: 10.1111/ijlh.12567. **Comunicació oral.**
3. L. Puigví, A. Merino, A. Acevedo, S. Alférez, J. Rodellar – *Cuantificación de características citológicas para el reconocimiento automático de células linfoides normales, reactivas, anormales y básicas de sangre periférica* – Libro de Comunicaciones del X Congreso Nacional del Laboratorio Clínico, pp. 418, ISBN: 978-84-617-4651-4. Saragossa, octubre 2016. **Comunicació.**
4. L. Puigví, A. Merino, S. Alférez, A. Acevedo, J. Rodellar – *Analysis of the most relevant quantitative features for the automatic differentiation of normal, reactive, abnormal lymphoid cell images and blast cell images from peripheral blood* – Ponent convidada al International Council for Standardization in Hematology (ICSH) a Lucerna, 26 de octubre de 2016. **Conferència.**
5. L. Puigví, A. Merino, S. Alférez, A. Acevedo, J. Rodellar – *Definition of quantitative variables for the morphological differentiation of malignant lymphoid cells in peripheral blood based on image analysis* – International Symposium on Biomedical Imaging (ISBI) – Melbourne (Austràlia), abril 2017. **Comunicació.**
6. L. Puigví, A. Acevedo, A. Merino – *Contribution of Advia 2120i blood count parameters and morphologic flags to the lymphoid neoplasms detection* – International Society of Laboratory of

- Hematology (ISLH) – Honolulu (Hawaii), maig 2017. International Journal of Laboratory Hematology, Issue S2, vol 39, pàg 94. Wiley. doi: 10.1111/ijlh.12698. **Comunicació.**
7. L. Puigví, A. Merino, S. Alférez, A. Acevedo, J. Rodellar – *New quantitative features obtained by image analysis to standardize morphological assessment of abnormal lymphoid cells in peripheral blood* – International Society of Laboratory of Hematology (ISLH) – Honolulu (Hawaii), maig 2017. International Journal of Laboratory Hematology, Issue S2, vol 39, pàg 95. Wiley. doi: 10.1111/ijlh.12698. **Comunicació.**
 8. A. Merino, S. Alférez, A. Acevedo, L. Puigví, J. Rodellar – *Detection of the presence of blasts, atypical promyelocytes, abnormal and reactive lymphoid cells in peripheral blood using image analysis* – International Society of Laboratory of Hematology (ISLH) – Honolulu (Hawaii), maig 2017. International Journal of Laboratory Hematology, Issue S2, vol 39, pàg 74-75. Wiley. doi: 10.1111/ijlh.12698. **Comunicació.**
 9. S. Alférez, A. Merino, L. Puigví, A. Acevedo, J. Rodellar – *Automatic cell recognition of abnormal cells in peripheral blood using digital images obtained by a conventional microscope* – International Society of Laboratory of Hematology (ISLH) – Honolulu (Hawaii), maig 2017. International Journal of Laboratory Hematology, Issue S2, vol 39, pàg 11. Wiley. doi: 10.1111/ijlh.12698. **Comunicació.**
 10. L. Puigví, A. Merino, S. Alférez, A. Acevedo, J. Rodellar – *Detección de nuevas características cuantitativas mediante procesamiento digital de imágenes de células linfoides de sangre periférica utilizando microscopía convencional* – Libro de Comunicaciones del XI Congreso Nacional del Laboratorio Clínico, 2017, pp. 466, ISBN: 978-84-697-6528-9. Màlaga, novembre 2017. **Comunicació.**
 11. L. Puigví, A. Merino, S. Alférez, L. Boldú, A. Acevedo, J. Rodellar – *Avaluació de nous descriptors obtinguts per anàlisi d'imatge per a la diferenciació de cèl·lules limfoides anormals de sang perifèrica* – Associació Catalana de Ciències del Laboratori Clínic (ACCLC) – Reus, març 2018. **Comunicació.**
 12. L. Puigví, A. Merino, S. Alférez, L. Boldú, A. Acevedo, J. Rodellar – *Quantitative features in peripheral blood cell images obtained by different hospitals and acquisition methods: utility for reactive and neoplastic lymphoid cell discrimination* – International Society of Laboratory of Hematology (ISLH) – Brusel·les (Bèlgica), maig 2018. International Journal of Laboratory Hematology, Issue S2, vol 40, pàg 119. Wiley. doi: 10.1111/ijlh.12922. **Comunicació.**
 13. L. Puigví, A. Merino, S. Alférez, L. Boldú, A. Acevedo, J. Rodellar – *Evaluación de características cuantitativas para la detección de granulación citoplasmática y textura de la cromatina en células linfoides anormales de sangre periférica* – Libro de Comunicaciones del XII Congreso Nacional

del Laboratorio Clínico, 2018, pp. 149, ISBN: 978-84-09-05283-7. Bilbao, octubre 2018.

Comunicació.

14. A. Merino, S. Alférez, L. Boldú, A. Molina, L. Puigví, A. Acevedo, J. Rodellar – *A new image-based machine-learning system (CellsiMatic) for the automatic recognition of hematologic neoplasia versus infections in peripheral blood* – International Society of Laboratory of Hematology (ISLH) – Vancouver (Canadà), maig 2019 – International Journal of Laboratory Hematology, Issue S2, vol 41, – pendent de publicació. **Comunicació oral.**

8.2.3 Premis

- Premi Carol Briggs-Smalley Scholarship 2016 per realitzar un projecte d'estandardització relacionat amb el Laboratori d'Hematologia. Entregat per l'International Council for Standardization in Haematology (ICSH), maig 2016.
- Young Investigator Award de la International Society for Laboratory Hematology (ISLH). Sessió plenària del XIX International Symposium on Technological Innovations in Laboratory Hematology. Maig 2016, Milà (Itàlia).

8.3 Perspectives de futur

Els resultats obtinguts en aquesta Tesi donen peu a obrir alguns temes que podrien seguir desenvolupant-se. Alguns d'ells es detallen a continuació:

- Estudi de descriptors per a la diferenciació entre les leucèmies agudes.
- Anàlisi de descriptors per a la detecció d'alteracions eritrocitàries de forma i inclusions.
- Recerca de descriptors quantitativs per a diferenciar entre les cèl·lules limfoides de la leucèmia limfàtica crònica convencional respecte a la variant morfològica atípica, ja que aquesta variant mostra unes alteracions citogenètiques específiques com la trisomia 12, la hiperploïdia i altres alteracions genètiques i immunofenotípiques associades a un pitjor pronòstic.
- Els descriptors identificats en aquesta recerca podrien incorporar-se als nous analitzadors hematològics basats en l'anàlisi d'imatges, ja que permetrien proporcionar alarmes específiques i sensibles per a la detecció de la cromatina madura i condensada, la cromatina cerebriforme, la granulació azuròfila del citoplasma i les vellositats citoplasmàtiques.

Referències

1. Zini G, Bain B, Bettelheim P, Cortez J, d'Onofrio G, Faber E, Haferlach T, Kacirkova P, Lewandowski K, Matutes E, Maynadie M, Meletis J, Petersen BL, Porwit A, Terpos E, Tichelli A, Vallespi T, Woessner S, Bennett J BM. A European consensus report on blood cell identification: terminology utilized and morphological diagnosis concordance among 28 experts from 17 countries within the European Leukemia Net network WP10, on behalf of the ELN Morphology Faculty. *Br J Haematol*. 2010;151:359–64.
2. Jahanmehr S, Rogers M, Zheng J, Lai R, Wang C. Quantitation of cytological parameters of malignant lymphocytes using computerized image analysis. *Int J Lab Hematol*. 2008;30(4):278–85.
3. Alférez S. Methodology for Automatic Classification of Atypical Lymphoid Cells from Peripheral Blood Cell Images. Universitat Politècnica de Catalunya · BarcelonaTech; 2015.
4. Gutiérrez G, Merino A, Domingo A, Jou JM, Reverter JC. EQAS for peripheral blood morphology in Spain: a 6-year experience. *Int J Lab Hematol*. 2008;30(6):460–6.
5. van der Meer W, van Gelder W, de Keijzer R, Willems H. The divergent morphological classification of variant lymphocytes in blood smears. *J Clin Pathol*. 2007;60(7):838–9.
6. Tuzel O, Yang L, Meer P, Foran DJ. Classification of hematologic malignancies using texton signatures. *Pattern Anal Appl PAA*. 2007 Oct 1;10(4):277–90.
7. Yang L, Meer P, Foran DJ. Unsupervised segmentation based on robust estimation and color active contour models. *IEEE Trans Inf Technol Biomed A Publ IEEE Eng Med Biol Soc*. 2005 Sep;9(3):475–86.
8. Yang L, Tuzel O, Chen W, Meer P, Salaru G, Goodell L a., et al. PathMiner: a Web-based tool for computer-assisted diagnostics in pathology. *IEEE Trans Inf Technol Biomed*. 2009 May;13(3):291–9.
9. Angulo J, Flandrin G, Math M, Central L, Necker U. Microscopic image analysis using mathematical morphology: Application to haematological cytology. *Sci Technol Educ Microsc An Overv*. 2003;304–312.
10. Angulo J, Klossa J, Flandrin G. Ontology-based lymphocyte population description using mathematical morphology on colour blood images. *Cell Mol Biol*. 2006;52(6):2–15.
11. Eldar S, Sabo E, Cohen A. Computer- assisted image analysis of small cell lymphoma of the thyroid gland. Comparison of the nuclear parameters of small lymphocytes in lymphomas and Hashimoto's thyroiditis. *Comput Med Imaging GraphComput Med Imaging Graph*. 1998;22:479–88.
12. Lunning M, Zenger V, Dreyfuss R, Stetler-Stevenson M, Rick M, White T, et al. Albumin enhanced morphometric image analysis in CLL. *Cytom B Clin Cytom*. 2004;57(1):7–14.

13. Benattar L, Flandrin G. Morphometric and colorimetric analysis of peripheral blood smears lymphocytes in B-cell disorders: proposal for a scoring system. *Leuk Lymphoma*. 2001;42(1–2):29–40.
14. Sabino D, Dafontouracosta L, Gilrizzatti E, Antoniozago M, da Fontoura Costa L, Gil Rizzatti E, et al. A texture approach to leukocyte recognition. *Real-Time Imaging*. 2004 Aug;10(4):205–16.
15. Yang X, Tridandapani S, Beitler J, Yu D, Yoshida E, Curran W, et al. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an in vivo study of late toxicity. *Med Phys*. 2012;39(9):5732–9.
16. Kono K, Hayata R, Murakami S, Yamamoto M, Kuroki M, Nanato K, et al. Quantitative distinction of the morphological characteristic of erythrocyte precursor cells with texture analysis using gray level co-occurrence matrix. *J Clin Lab Anal*. 2018;32(Jan):e22175.
17. Patel N, Mishra A. Automated Leukaemia Detection Using Microscopic Images. *Procedia Comput Sci*. 2015;58:635–42.
18. Yamamoto M, Kono K, Kuroki M, Murakami S, Hayata R, Nanato K, et al. Quantitative distinction of morphological characteristics of white blood cells in peripheral blood through texture analysis using gray level co-occurrence matrix. *Clin Lab*. 2017;63:1851–68.
19. Sabino DMU, Costa LF, Rizzatti EG, Zago M a. Toward leukocyte recognition using morphometry, texture and color. 2004 2nd IEEE Int Symp Biomed Imaging Macro to Nano (IEEE Cat No 04EX821). 2004;2:121–4.
20. Sabino DM., Lorena AC, de Carvalho ACPLF. Support vector machines applied to white blood cell recognition. In: *Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*. IEEE; 2005. p. 6 pp.
21. Comaniciu D, Meer P, Foran DJ. Image-guided decision support system for pathology. *Mach Vis Appl*. 1999 Dec 16;11(4):213–24.
22. Foran DJ, Comaniciu D, Meer P, Goodell L a. Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. *IEEE Trans Inf Technol Biomed*. 2000 Dec;4(4):265–73.
23. Saeedizadeh Z, Mehri Dehnavi A, Talebi A, Rabbani H, Sarrafzadeh O, Vard A. Automatic recognition of myeloma cells in microscopic images using bottleneck algorithm, modified watershed and SVM classifier. *J Microsc*. 2015;261(1):46–56.
24. Mohapatra S, Patra D. Automated leukemia detection using hausdorff dimension in blood microscopic images. *Interact-2010*. 2010 Dec;64–8.
25. Mohapatra S, Samanta SS, Patra D, Satpathi S. Fuzzy based blood image segmentation for automated leukemia detection. *2011 Int Conf Devices Commun*. 2011 Feb;1–5.
26. Gonzalez J, Olmos I, Altamirano L, Morales BA, Reta C, Galindo MC, et al. Leukemia identification from bone marrow cells images using a machine vision and data mining strategy. *Intell Data Anal* . 2011;15:443–62.
27. Madhloom HT, Kareem SA, Ariffin H. A robust feature extraction and selection method for the recognition of lymphocytes versus acute lymphoblastic leukemia. *2012 Int Conf Adv Comput Sci Appl Technol*. 2012 Nov;330–5.
28. Aimi A, Mashor M, Rosline H. Classification of acute leukaemia cells using multilayer perceptron and simplified fuzzy ARTMAP neural networks. *Int Arab J Inf Technol*. 2013;10(4):356–64.

29. Alférez S, Merino A, Mujica LE, Ruiz M, Bigorra L, Rodellar J. Automatic classification of atypical lymphoid B cells using digital blood image processing. *Int J Lab Hematol.* 2014;36(4):472–80.
30. Alférez S, Merino A, Bigorra L, Mujica L, Ruiz M, Rodellar J. Automatic recognition of atypical lymphoid cells from peripheral blood by digital image analysis. *Am J Clin Pathol.* 2015;143:168–76.
31. Alférez S, Merino A, Bigorra L, Rodellar J. Characterization and automatic screening of reactive and abnormal neoplastic B lymphoid cells from peripheral blood. *Int J Lab Hematol.* 2016;38(2):209–19.
32. Ciesla B. *Hematology in Practice.* FA Davis; 2011.
33. Bain BJ, Clark DM, Lampert IA, Wilkins BS. *Bone Marrow Pathology.* John Wiley & Sons; 2008.
34. Merino A. *Manual de Citología de Sangre Periférica.* Grupo Acción Médica; 2005.
35. Kottke-Marchant K, Davis BH. *Laboratory Hematology Practice.* Kottke-Marchant K, Davis BH, editors. Oxford, UK: Wiley-Blackwell; 2012.
36. Vardiman JW, Thiele J, Arber D a, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood.* 2009 Jul 30;114(5):937–51.
37. Palmer L, Briggs C, McFadden S, Zini G, Burthem J, Rozenberg G, et al. ICSH recommendations for the standardization of nomenclature and grading of peripheral blood cell morphological features. *Int J Lab Hematol.* 2015;37(3):287–303.
38. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* 4th ed. Lloyd R, Osamura R, Klöppel G, Rosai J, editors. France: IARC Press; 2017.
39. Bain BJ. Diagnosis from the blood smear. *N Engl J Med.* 2005;353(5):498–507.
40. Merino A. Signos de alarma en el hemograma y utilidad diagnóstica de la morfología sanguínea. *Rev del Lab Clínico.* 2010 Oct;3(4):151–2.
41. Sinclair GJ. Blood film review by biomedical scientists. *Br J Biomed Sci.* 2005 Jan;62(2):77–80.
42. Ongun G, Halici U. An automated differential blood count system. In: *Engineering in Medicine and Biology Society, 2001 Proceedings of the 23rd Annual International Conference of the IEEE.* 2001. p. 2583–6.
43. Ward PC. The CBC at the turn of the millennium: an overview. *Clin Chem [Internet].* 2000 Aug;46(8 Pt 2):1215–20.
44. Harms H, Aus HM, Haucke M, Gunzer U. Segmentation of stained blood cell images measured at high scanning density with high magnification and high numerical aperture optics. *Cytometry.* 1986 Nov;7(6):522–31. 4
45. Pierre RV. Peripheral blood film review. The demise of the eyecount leukocyte differential. *Clin Lab Med.* 2002;22 (1):279–97.
46. Merino A, Brugués R, García R, Kinder M, Torres F, Escolar G. Estudio comparativo de la morfología de sangre periférica analizada mediante el microscopio y el CellaVision DM96 en enfermedades hematológicas y no hematológicas. *Rev del Lab Clínico.* 2011 Jan;4(1):3–14.

47. Sysmex. Medica EasyCell [Internet]. Available from: <https://www.sysmex.com/us/en/Products/Hematology/CellImageAnalysis/Pages/Medica-EasyCell.aspx>
48. Fraunhofer Institute for Integrated Circuits. HemaCAM.
49. Kratz A, Bengtsson H-I, Casey JE, Keefe JM, Beatrice GH, Grzybek DY, et al. Performance evaluation of the CellaVision DM96 system: WBC differentials by automated digital image analysis supported by an artificial neural network. *Am J Clin Pathol*. 2005 Nov 1;124(5):770–81.
50. Cornet E, Perol J-P, Troussard X. Performance evaluation and relevance of the CellaVision DM96 system in routine analysis and in patients with malignant hematological diseases. *Int J Lab Hematol*. 2008 Dec;30(6):536–42.
51. Briggs C, Longair I, Slavik M, Thwaite K, Mills R, Thavaraja V, et al. Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. *Int J Lab Hematol*. 2009 Feb;31(1):48–60.
52. Bloodhound, Roche Website. cobas m 511.
53. Longo D. Harrison's Hematology and Oncology. McGraw Hill Professional; 2010.
54. Bain BJ, Catovsky D. The leukemic phase of non-Hodgkin's lymphoma. *J Clin Pathol*. 1994;48:289–93.
55. Woessner S, Florensa L. La citología óptica en el diagnóstico hematológico. 4th ed. Médica A, editor. 2000.
56. Bennett JM, Catovsky D DM et al. Proposals for the classification of chronic (mature) B and T lymphoid leukemias. *J Clin Pathol*. 1989;42:567–84.
57. Matutes E, García TJ OM et al. The morphological spectrum of T-prolymphocytic leukemia. *Br J Haematol*. 1986;64:111–24.
58. Riley RS, Ben-Ezra JM, Massey D, Cousar J. The virtual blood film. *Clin Lab Med*. 2002 Mar;22(1):317–45.
59. Alférez S, Merino A, Acevedo A, Puigví L, Rodellar J. Colour Clustering Segmentation Framework for Image Analysis of Malignant Lymphoid Cells in Peripheral Blood. *Med Biol Eng Comput*. 2019;
60. Fox J, Bouchet-Valat M. Rcmdr: R Commander. R package version 2.5-1 [Internet]. 2018. Available from: <http://rcommander.com/>
61. Scotti F. Robust segmentation and measurements techniques of white cells in blood microscope images. 2006 IEEE Instrum Meas Technol Conf Proc. 2006 Dec;(April):43–8.
62. Würflinger T, Stockhausen J, Meyer-Ebrecht D, Böcking A. Automatic coregistration, segmentation and classification for multimodal cytopathology. *Stud Health Technol Inform*. 2003 Jan;95:218–23.
63. Sadeghian F, Seman Z, Ramli AR, Abdul Kahar BH, Saripan M-I. A Framework for White Blood Cell Segmentation in Microscopic Blood Images Using Digital Image Processing. *Biol Proced Online*. 2009 Jun;11(1):196–206.
64. Eom S, Kim S, Shin V, Ahn B. Leukocyte segmentation in blood smear images using region-based active contours. *Adv Concepts Intell Vis*. 2006;867–76.
65. Sabino D, Calado RT, Rizzatti EG. LNCS 4091 - Leukocyte detection using nucleus contour propagation. *Differ Equations*. 2006;389–96.

66. Angulo J, Flandrin G. Automated detection of working area of peripheral blood smears using mathematical morphology. *Anal Cell Pathol.* 2003 Jan;25(1):37–49.
67. Ghosh M, Das D, Chakraborty C, Ray AK. Automated leukocyte recognition using fuzzy divergence. *Micron.* 2010;
68. Sinha N, Ramakrishnan A. Automation of differential blood count. *TENCON 2003 Conf.* 2003;(i).
69. Ross TJ. *Fuzzy Logic with Engineering Applications.* John Wiley & Sons; 2009.
70. Balafar M, Ramli A, Saripan M, Mashohor S. Review of brain MRI image segmentation methods. *Artif Intell Rev.* 2010;33(3):261–74.
71. Dembele D, Kastner P. Fuzzy C-means method for clustering microarray data. *Bioinformatics.* 2003;19(8):973–80.
72. Plissiti M, Nikou C, Charchanti A. Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. *IEEE Trans Inf Technol Biomed.* 2011;15(2):233–41.
73. Otsu N. A Threshold Selection Method from Gray-Level Histograms. In: *IEEE Transactions on Systems, Man, and Cybernetics.* 1979. p. 62–6.
74. Beucher S. The watershed transformation applied to image segmentation. *Scanning Microsc.* 1992;299–314.
75. Gonzalez RC, Woods RE. *Digital Image Processing.* 3rd ed. Prentice Hall; 2007.
76. Mohammed EA, Mohamed MMA, Far BH, Naugler C. Peripheral blood smear image analysis: A comprehensive review. *J Pathol Inform.* 2014;5(1):9.
77. Gelsema ES, Bao HF, Smeulders AWM, Den Harink HC. Application of the method of multiple thresholding to white blood cell classification. *Comput Biol Med.* 1988;18(2):65–74.
78. Madhloom HT, Kareem SA, Ariffin H. An image processing application for the localization and segmentation of lymphoblast cell using peripheral blood images. *J Med Syst.* 2012 Aug [cited 2013 Apr 10];36(4):2149–58.
79. Piuri V, Scotti F. Morphological classification of blood leucocytes by microscope images. 2004 *IEEE Int Conf on Computational Intell Meas Syst Appl 2004 CIMSA.* 2004;(July):103–8.
80. Scotti F. Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images. In: *CIMSA 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2005.* IEEE; 2005. p. 96–101.
81. Katz A. *Image Analysis and Supervised learning in the automated Differentiation of White Blood cells from Microscopic Images.* 2000.
82. Jiang K, Liao QQ-M, Dai SS-Y. A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering. In: *Machine Learning and Cybernetics, 2003 International Conference.* Ieee; 2003. p. 2820–5.
83. Aimi Salihah AN, Mashor MY, Harun NH, Abdullah AA, Rosline H. Improving colour image segmentation on acute myelogenous leukaemia images using contrast enhancement techniques. In: *Biomedical Engineering and Sciences (IECBES), 2010 IEEE EMBS Conference on.* 2010. p. 246–51.
84. Foran DJ, Comaniciu D, Meer P, Goodell LA. Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. *IEEE Trans Inf Technol Biomed.* 2000 Dec;4(4):265–73.

85. Comaniciu D, Meer P, Foran DJ. Image-guided decision support system for pathology. *Mach Vis Appl.* 1999 Dec 16;11(4):213–24.
86. Ramoser H, Laurain V, Bischof H, Ecker R. Leukocyte segmentation and classification in blood-smear images. In: *Engineering in Medicine and Biology Society, 2005 IEEE-EMBS 2005 27th Annual International Conference.* Austria; 2006. p. 3371–4.
87. Nikolaou N, Papamarkos N. Color reduction for complex document images. *Int J Imaging Syst Technol.* 2009;19(1):14–26.
88. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;3(6):610–21.
89. Tuceryan M, Jain AK. Texture analysis. In: *Handbook of Pattern Recognition and Computer.* World Scientific Publishing; 1993. p. 235–76.
90. Haralick RM, Shapiro LG. *Computer and robot vision.* Addison-Wesley Longman Publishing Co., Inc.; 1991.
91. Arivazhagan S, Ganesan L. Texture classification using wavelet transform. *Pattern Recognit Lett.* 2003 Jun;24(9–10):1513–21.
92. Han J, Ma KK. Rotation-invariant and scale-invariant Gabor features for texture image retrieval. *Image Vis Comput.* 2007;25(9):1474–81.
93. R Development Core Team. *R: A language and environment for statistical computing.* [Internet]. Vienna, Austria; 2008. Available from: <http://www.r-project.org>
94. Kuehl RO. *Diseño de experimentos: principios estadísticos para el diseño y análisis de investigaciones.* 2a ed. PARANINFO, S.A. EDICIONES; 2000. 666 p.
95. Martin A, Luna J. *Bioestadística para las ciencias de la salud.* 5a ed. Capitel Ediciones; 2004. 672 p.
96. van Belle G, Fisher L, Heagerty P, Lumley T. *Biostatistics: A Methodology For the Health Sciences.* 2a ed. Wiley-Interscience; 2004. 896 p.
97. Canavos GC. *Probabilidad y Estadística. Aplicaciones y Métodos.* McGraw-Hill; 1984. 651 p.
98. Edgell SE, Noon SM. Effect of violation of normality on the t–test of the correlation coefficient. *Psychol Bull.* 1984;95:576–83.
99. Robin X, Turk N, Hainard A, Tiberti N, Lisacek F, Sanchez J, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
100. Peterson P, McNeill S, Gulati G. Cellular morphologic analysis of peripheral blood. *Lab Hematol Pract.* 2012;
101. Gulati G, Hyun B. Blood smear examination. *Hematol Oncol Clin North Am.* 1994;8(4):631–50.
102. Barnes P, McFadden S, Machin S, Simson E. The international consensus group for hematology review: suggested criteria for action following automated CBC and WBC differential analysis. *Lab Hematol.* 2005;11(2):83–90.
103. de Vet H, Koudstaal J, Kwee W, Willebrand D, Arends J. Efforts to improve interobserver agreement in histopathological grading. *J Clin Epidemiol.* 1995;48(7):869–73.
104. International Society for Laboratory Hematology. *Consensus Guidelines: Positive Smear Findings.* http://www.islh.org/web/consensus_rules.php.

105. Bloom K, Harrington D. Enhanced accuracy and reliability of HER-2/neu immunohistochemical scoring using digital microscopy. *Am J Clin Pathol.* 2004;121(5):620–30.
106. Marchevsky A, Gal A, Shah S, Koss M. Morphometry confirms the presence of considerable nuclear size overlap between “small cells” and “large cells” in high-grade pulmonary neuroendocrine neoplasms. *Am J Clin Pathol.* 2001;116(4):466–72.
107. Benattar L, Flandrin G. Morphometry and quality control for a May-Grunwald Giemsa stained preparation. A 40 centers cooperative study. *Leuk Lymphoma.* 1999;33(5–6):587–91.
108. Angulo J. Morphologie mathématique et indexation d’images couleur: application à la microscopie en biomédecine. École Nationale Supérieure des Mines de Paris; 2003.
109. Mujica LE, Ruiz M, Acho L, Alférez S, Tutivén C, Vidal Y, et al. Texture analysis for wind turbine fault detection. In: *E-Journal of non-destructive testing.* 2016. p. 1–10.
110. Castellano G, Bonilha L, Li L, Cendes F. Texture analysis of medical images. *Clin Radiol.* 2004;59(12):1061–9.
111. Johnson S. *Stephen Johnson on Digital Photography.* O’Reilly Media; 2006. 320 p.
112. Albrechtsen F. *Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices.* Image Processing Laboratory, Department of Informatics, University of Oslo; 1995.
113. Angulo J. A mathematical morphology approach to cell shape analysis. In: Bonilla LL, Moscoso M, Platero G, Vega JM, editors. *Progress in Industrial Mathematics at ECMI 2006.* Berlin, Heidelberg: Springer; 2008. p. 2–6. (Mathematics in Industry; vol. 12).
114. Kruizinga P, Petkov N. Nonlinear Operator for Oriented Texture. *IEEE Trans Image Process.* 1999;8(10):1395–407.
115. Materka A, Strzelecki M. Texture analysis methods - a review. Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels. 1998. p. 9–11.
116. Solomon C, Breckon T. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab.* John Wiley & Sons; 2011.
117. Mehrotra H, Namuduri K, Ranganathan N. Gabor filter-based edge detection. *Pattern Recognit.* 1992;25(12):1479–94.
118. Fleuret F. Fast binary feature selection with conditional mutual information. *J Mach Learn Res.* 2004 Dec;5:1531–55.
119. Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res.* 2012;13(1):27–66.
120. Tukey J. *Exploratory Data Analysis.* Addison-Wesley; 1977.
121. Pohlert T. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR) [Internet]. 2014. Available from: <http://cran.r-project.org/package=PMCMR>.
122. McDonald J. *Handbook of Biological Statistics.* 3rd ed. Baltimore, Maryland.: Sparky House Publishing; 2014.
123. Bigorra L, Merino A, Alférez S, Rodellar J. Feature analysis and automatic identification of leukemic lineage blast cells and reactive lymphoid cells from peripheral blood cell images. *J Clin Lab Anal.* 2017;31(2):e22024.
124. Poostchi M, Silamut K, Maude R, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. *Transl Res.* 2018;194:36–55.

125. Pinco J, Goulart R, Otis C, Garb J, Pantanowitz L. Impact of Digital Image Manipulation in Cytology. *Arch Pathol Lab Med.* 2009;133:57–61.
126. Puigví L, Merino A, Alférez S, Acevedo A, Rodellar J. New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood. *J Clin Pathol.* 2017;70(12):1038–48.
127. Rodellar J, Alférez S, Acevedo A, Molina A, Merino A. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hematol.* 2018;40(Suppl 1):46–53.
128. Merino A, Puigví L, Boldú L, Alférez S, Rodellar J. Optimizing morphology through blood cell image analysis. *Int J Lab Hematol.* 2018;40(Suppl.1):54–61.
129. Alférez S, Merino A, Bigorra L, Mujica LE, Ruiz M, Rodellar J. Method for automatic recognition of neoplastic lymphoid cells using peripheral blood cell images [Abstract]. *Int J Lab Hematol.* 2015 May;37(Supplement S1):1–143.
130. Altman D, Bland J. Statistics Notes. Units of analysis. *BMJ.* 1997;314:1874.
131. Chan S, Leung K, Wong W. An expert system for the detection of cervical cancer cells using knowledge-based image analyzer. *Artif Intell Med.* 1996;8(1):67–90.
132. MoradiAmin M, Memari A, Samadzadehaghdam N, Kermani S, Talebi A. Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis. *Microsc Res Tech.* 2016;79(10):908–16.
133. Greenspan H, Rothmann C, Cycowitz T, Nissan Y, Cohen a M, Malik Z. Classification of lymphoproliferative disorders by spectral imaging of the nucleus. *Histol Histopathol.* 2002 Jan;17(3):767–73.
134. Silva M, Fourcade C, Fartoukh C, Lenormand B, Buchonnet G, Callat M, et al. Lymphocyte volume and conductivity indices of the haematology analyser Coulter® GEN.STM in lymphoproliferative disorders and viral diseases. *Int J Lab Hematol.* 2006;28(1):1–8.
135. Cerny J, Rosmarin A. Why does my patient have leukocytosis? *Hematol Oncol Clin North Am.* 2012;26:303–19.
136. Bruegel M, George T, Feng B, Hilger A, Allen T, Zahniser D, et al. Multicenter evaluation of the cobas m 511 integrated hematology analyzer. *Int J Lab Hematol.* 2018;40:672–82.
137. Roche Website. cobas m 511 [Internet]. Available from: <http://www.cobas.com/home/product/hematology-testing/cobas-m-511.html>
138. Kumar G, Bhatia PK. A Detailed Review of Feature Extraction in Image Processing Systems. 2014 Fourth Int Conf Adv Comput Commun Technol. 2014;5–12.
139. Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. Cengage Learning; 2008.

Glossari

Asimetria L'asimetria o *skewness* descriu el grau de simetria de la distribució d'intensitats de la imatge. Si l'asimetria és negativa, les intensitats dels píxels s'estenen més a l'esquerra de la mitjana (intensitats baixes) que a la dreta (intensitats brillants). Si l'asimetria és positiva, les intensitats dels píxels es troben més cap a la dreta.

Cèl·lula limfoide reactiva Els limfòcits reactius són benignes, de mida gran, tenen un nucli de cromatina laxa i un citoplasma ampli, intensament basòfil, especialment als punts de contacte amb els hematies veïns. És freqüent observar-ne en frotis de pacients que pateixen infeccions víriques o bacterianes.

Corbes granulomètriques La corba granulomètrica és un histograma on es representa la mida dels grànuls i de mida en el qual un valor alt a una mida específica indica la presència de moltes estructures brillants (o estructures fosques) amb una mida similar a la imatge (108).

Curtosi La curtosi és una mesura del grau d'aplanament de l'histograma d'intensitats. En una distribució normal, la curtosi pren un valor de zero.

Descriptors estadístics de primer ordre Aquests descriptors es basen en l'histograma d'una imatge digital en escala de grisos. A partir de la informació estadística sobre la imatge proporcionada per l'histograma, es poden obtenir diversos paràmetres estadístics de primer ordre, com ara: mitjana, desviació estàndard, asimetria, curtosi, energia i entropia (115).

Descriptors estadístics de segon ordre Proporcionen més informació que els descriptors de primer ordre. Aquests descriptors es basen en una probabilitat conjunta que un parell de píxels tenen una combinació particular d'un parell de valors a una distància específica en una direcció determinada (75,88,115).

Entropia L'entropia $H(x)$ quantifica la incertesa present en la distribució d'una variable aleatòria (veure 2.9 i 2.10). Si hi ha poca incertesa sobre el resultat, l'entropia és baixa. Si tots els esdeveniments són igual de probables, és a dir, hi ha la màxima incertesa sobre el resultat, llavors l'entropia és màxima (119).

Energia L'energia és una mesura que mostra com es distribueixen els valors de nivell gris dins de la imatge. Té una relació inversa amb l'entropia, en què l'energia d'una imatge serà la més alta si només té un valor de nivell gris. Com més nivells de grisos estiguin presents en un objecte, menor serà la seva energia. L'energia de l'histograma dona informació de la uniformitat.

Espai de color Un model o espai de color és un mètode per especificar els colors d'una manera estandarditzada i acceptada de forma general. Essencialment consisteix en un sistema de coordenades en tres dimensions (75).

Extracció de descriptors L'objectiu principal de l'extracció de descriptors és obtenir informació dels objectes d'interès de la imatge digital. Es tracta d'una forma de reducció de la dimensionalitat, perquè la imatge està representada per un conjunt de descriptors (75,138).

Frotis de sang perifèrica Extensió de sang tenyida sobre un porta-objectes de vidre que mitjançant la seva observació al microscopi, permet la detecció d'anomalies de les cèl·lules circulants.

Granulometria És un camp de la morfologia matemàtica que tracta de determinar la distribució de la mida de les partícules en una imatge (75,108).

Histograma L'histograma és una forma molt pràctica de descriure quantitativament els píxels d'una imatge i és la base del càlcul dels descriptors de color. És una funció discreta que mostra el nombre de píxels $H(i)$ a la imatge que té un valor d'intensitat de píxels i freqüències (75).

Limfòcit Un dels leucòcits més nombrosos a sang perifèrica (25 – 33 %), de mida petita, al voltant d'uns 7 μm de diàmetre amb un nucli esfèric i dens, que ocupa gairebé la totalitat de la cèl·lula, tot i que un reduït percentatge poden tenir una mida superior (12 – 16 μm), amb un nucli rodó de cromatina condensada o madura i un citoplasma basòfil molt escàs.

Matriu de co-ocurrència del nivell de grisos (GLCM) És un histograma de dues dimensions que especifica la freqüència en què dos nivells de grisos es troben en parelles de píxels separats per una distància i direcció particular. Per tant, considera la relació entre els píxels veïns en la imatge original i representa la probabilitat conjunta $P(i, j)$ que un parell de píxels amb valors d'intensitat d' i i j , respectivament, a una distància d en una direcció particular θ . Aquesta probabilitat es pot calcular com el recompte de freqüències d'ocurrències dividit pel nombre total de píxels veïns. Així, la GLCM considera no només la informació sobre els valors d'intensitat, sinó també la posició dels píxels amb intensitats similars (88).

Neoplàsia limfoide Malaltia neoplàsica derivada dels limfòcits i els seus precursors que es desenvolupa en els òrgans hematopoètics i en el sistema limfàtic. Pot tenir lloc de dues maneres: com una leucèmia (cèl·lules tumorals al moll de l'os i a la sang) i/o un limfoma (massa sòlida de cèl·lules tumorals) (53).

Preprocessament El preprocessament d'imatges suprimeix informació que no és rellevant per a les tasques específiques del processament d'imatges. El preprocessament consisteix en un conjunt d'operacions de processament d'imatges que milloren certs paràmetres de les dades i eliminen cert soroll (139).

Processament d'imatges digitals (PDI) El PDI consisteix a aplicar múltiples algoritmes per processar imatges digitals mitjançant un ordinador (75).

Textura La textura es refereix a propietats que representen la superfície o estructura d'un objecte, però no té una definició precisa a causa de la seva gran variabilitat (139).

9 Annexos

9.1 Annex A

En aquest Annex A es mostren els resultats de la correlació així com l'anàlisi estadística descriptiva dels 20 descriptors presentats als Capítols 4 i 6 entre els dos sistemes d'adquisició d'imatges: DM96 i BX43. A la Taula 9.1 i la Figura 9.1 es mostren els resultats de la correlació de Pearson.

Taula 9.1: Resultats de la correlació de Pearson dels 20 descriptors més rellevants entre el sistema DM96 i el BX43. Es mostra el coeficient R de Pearson, el coeficient de determinació (R^2), l'interval de confiança (IC) del 95 %, el p valor així com la recta d'ajust.

Descriptor	R	R^2	95 % IC	p valor	Recta d'ajust
1.Rati nucli/citoplasma	0,973	0,947	0,972-0,974	$<2^{-16}$	$y = 0,62x + 1,08$
2.Perímetre nuclear	0,950	0,903	0,948-0,952	$<2^{-16}$	$y = 0,83x - 5,46$
3.Diàmetre equivalent de la cèl·lula	0,969	0,938	0,967-0,970	$<2^{-16}$	$y = 0,78x + 4,77$
4.Curtosi de la corba pseudo-granulomètrica del component cian de la cèl·lula	0,865	0,749	0,865-0,870	$<2^{-16}$	$y = 0,46x + 1,1$
5.Asimetria de l'histograma del component blau de la cèl·lula	0,974	0,949	0,973-0,976	$<2^{-16}$	$y = 1,07x + 0,04$
6.Information measure of correlation 1 del component magenta de la cèl·lula	0,799	0,639	0,890-0,898	$<2^{-16}$	$y = 0,77x - 0,26$
7.Desviació estàndard de l'histograma del component blau de la cèl·lula	0,945	0,894	0,943-0,948	$<2^{-16}$	$y = 0,75x + 1,19$
8.Entropia de l'histograma del component magenta de la cèl·lula	0,974	0,948	0,972-0,975	$<2^{-16}$	$y = 0,92x + 4,29$
9.Asimetria de l'histograma del component verd de la cèl·lula	0,929	0,863	0,926-0,932	$<2^{-16}$	$y = 1,53x + 0,09$
10.Mitjana de la corba pseudo-granulomètrica del component negre del citoplasma	0,872	0,761	0,866-0,878	$<2^{-16}$	$y = 4,66x - 0,20$
11.Mitjana de l'histograma del component blau del nucli	0,955	0,998	0,953-0,957	$<2^{-16}$	$y = 1,51x - 51,35$
12.Cluster shade del component de la saturació de la cèl·lula	0,696	0,485	0,684-0,708	$<2^{-16}$	$y = 1,16x - 12,27$
13.Mitjana de l'histograma del component magenta del citoplasma	0,835	0,697	0,828-0,842	$<2^{-16}$	$y = 0,75x + 30,59$
14.Homogeneïtat del component magenta de la cèl·lula	0,872	0,760	0,866-0,877	$<2^{-16}$	$y = 0,4x + 0,55$
15.Curtosi de l'histograma del component de la saturació de la cèl·lula	0,893	0,797	0,889-0,898	$<2^{-16}$	$y = 2,30x + 1,72$
16.Information measure of correlation 1 del component cian del nucli	0,859	0,737	0,853-0,864	$<2^{-16}$	$y = 1,22x + 0,02$
17.Hairiness	0,871	0,759	0,862-0,879	$<2^{-16}$	$y = 0,0000371x - 0,00236$
18.Desviació estàndard de l'histograma del component de la saturació de la cèl·lula	0,945	0,893	0,943-0,947	$<2^{-16}$	$y = 1,03x - 17,96$
19.Circularitat nuclear	0,940	0,884	0,938-0,943	$<2^{-16}$	$y = 0,96x + 0,06$
20.Homogeneïtat del component negre del citoplasma	0,904	0,817	0,899-0,908	$<2^{-16}$	$y = 1,31x - 0,17$

A la Taula 9.2 es detallen els coeficients de determinació per grup cel·lular limfoide i descriptor.

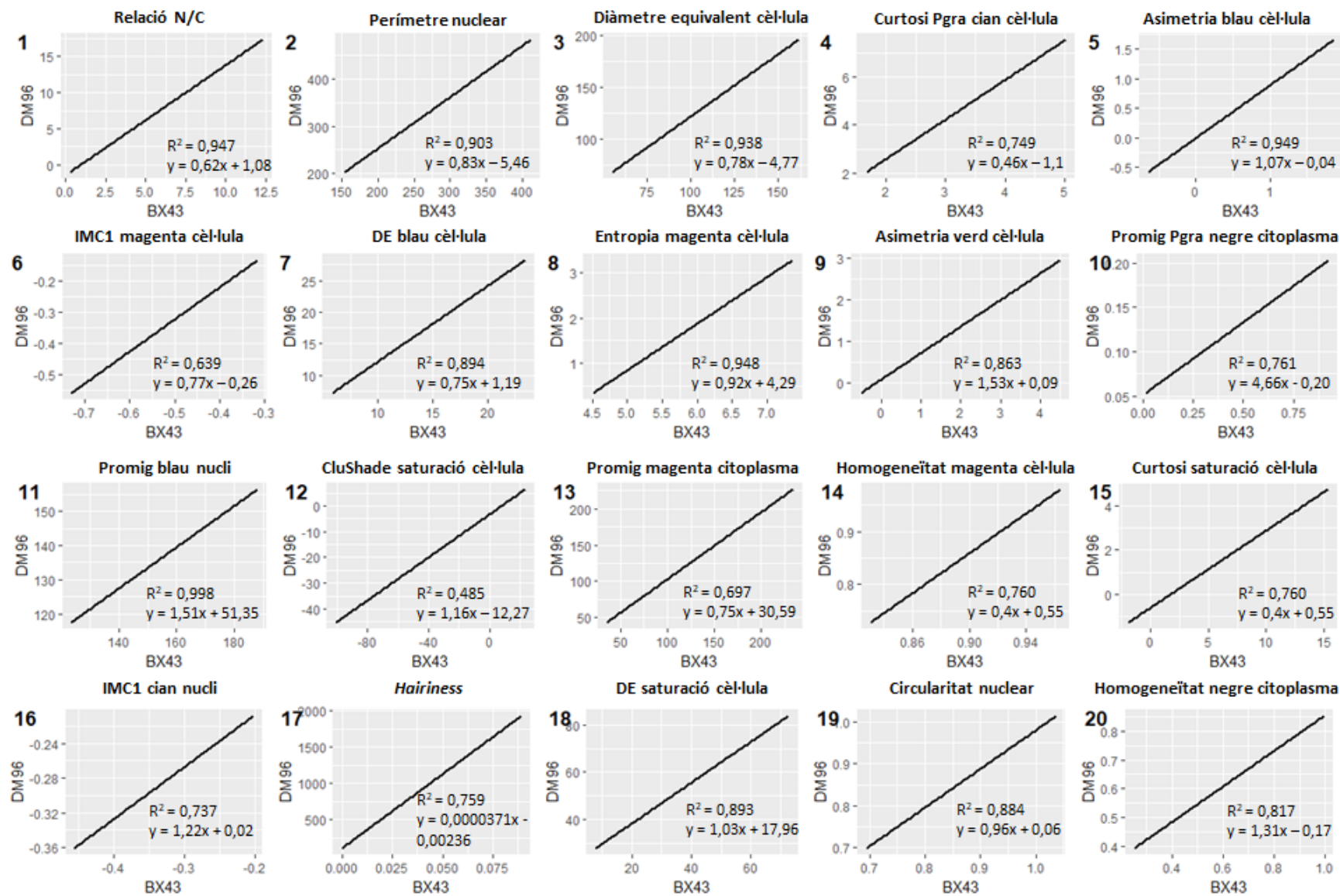


Figura 9.1: Correlacions de Pearson dels 20 descriptors entre el sistema DM96 i el BX43 per a tots els grups cel·lulars.

Taula 9.2: Coeficients de determinació (R²) dels 20 descriptors per a cada tipus cel·lular limfoide obtinguts entre el sistema DM96 i el BX43.

Descriptor	LI	LLC	HCL	LZME	LCM	LF	LPB	LPT	SS	LLGG	CLR	BL
1.Rati nucli/citoplasma	0,998	0,998	0,996	0,967	0,994	0,999	0,994	0,975	0,976	0,998	0,998	0,977
2.Perímetre nuclear	0,997	0,993	0,995	0,994	0,990	0,985	0,992	0,970	0,921	0,994	0,997	0,994
3.Diàmetre equivalent de la cèl·lula	0,997	0,991	0,994	0,993	0,995	0,984	0,983	0,979	0,969	0,996	0,996	0,994
4.Curtosi de la corba pseudo-granulomètrica del component cian de la cèl·lula	0,994	0,993	0,997	0,987	0,992	0,998	0,992	0,953	0,979	0,973	0,977	0,984
5.Asimetria de l'histograma del component blau de la cèl·lula	0,998	0,998	0,997	0,986	0,998	0,986	0,980	0,901	0,995	0,997	0,996	0,991
6. <i>Information measure of correlation 1</i> del component magenta de la cèl·lula	0,971	0,983	0,985	0,991	0,948	0,998	0,975	0,960	0,988	0,992	0,990	0,977
7.Desviació estàndard de l'histograma del component blau de la cèl·lula	0,998	0,996	0,978	0,988	0,974	0,960	0,989	0,958	0,983	0,996	0,990	0,997
8.Entropia de l'histograma del component magenta de la cèl·lula	0,990	0,994	0,989	0,998	0,988	0,987	0,992	0,977	0,989	0,997	0,996	0,985
9.Asimetria de l'histograma del component verd de la cèl·lula	0,997	0,998	0,989	0,986	0,984	0,992	0,982	0,945	0,994	0,997	0,997	0,997
10.Mitjana de la corba pseudo-granulomètrica del component negre del citoplasma	0,986	0,995	0,987	0,984	0,998	-	0,992	0,972	0,990	0,989	0,997	0,989
11.Mitjana de l'histograma del component blau del nucli	0,995	0,990	0,982	0,993	0,970	0,990	0,963	0,955	0,995	0,993	0,998	0,989
12. <i>Cluster shade</i> del component de la saturació de la cèl·lula	0,991	0,995	0,991	0,968	0,981	0,991	0,964	0,974	0,983	0,977	0,907	0,959
13.Mitjana de l'histograma del component magenta del citoplasma	0,994	0,997	0,985	0,995	0,993	0,986	0,982	0,877	0,996	0,984	0,995	0,996
14.Homogeneïtat del component magenta de la cèl·lula	0,988	0,992	0,992	0,996	0,997	0,989	0,978	0,957	0,997	0,994	0,996	0,995
15.Curtosi de l'histograma del component de la saturació de la cèl·lula	0,992	0,996	0,969	0,987	0,963	0,997	0,967	0,975	0,997	0,998	0,990	0,994
16. <i>Information measure of correlation 1</i> del component cian del nucli	0,997	0,986	0,993	0,995	0,997	0,992	0,928	0,978	0,995	0,956	0,999	0,985
17. <i>Hairiness</i>	0,974	-	0,982	0,945	0,941	-	0,983	0,974	0,981	0,971	0,869	0,944
18.Desviació estàndard de l'histograma del component de la saturació de la cèl·lula	0,989	0,998	0,989	0,990	0,885	0,998	0,985	0,947	0,985	0,980	0,990	0,990
19.Circularitat nuclear	0,999	0,998	0,994	0,997	0,993	0,989	0,997	0,961	0,997	0,996	0,998	0,992
20.Homogeneïtat del component negre del citoplasma	0,996	0,963	0,995	0,997	0,997	0,947	0,994	0,920	0,990	0,984	0,972	0,995

BL, blasts limfoides; CLR, cèl·lula limfoide reactiva; HCL, tricoleucèmia; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LI, limfòcits; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulars T; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

A continuació es detallen els resultats obtinguts de la correlació entre ambdós sistemes (DM96 i BX43) i de l'anàlisi estadística (descriptiva i comparació múltiple) per a cada un dels 20 descriptors.

9.1.1 *Relació nucli/citoplasma*

Pel que fa a la *relació nucli/citoplasma* (*N/C*), el coeficient de correlació r és de 0,973 ($p < 0,00001$, veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *relació N/C* obtinguts amb el sistema DM96 i el BX43 per als 12 grups de cèl·lules limfoides. En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,947 ($p < 0,00001$, veure Taula 9.2 i gràfic 1 Fig. 9.1), el qual implica que un 6,3 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust. Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular (veure Taula 9.2), tots ells són superiors a 0,983.

Pel que fa als valors obtinguts de la *relació N/C* amb el sistema BX43 (veure Fig. 9.2 i Taula 9.3), tal i com hem vist amb el sistema DM96 (Capítol 4), les dues cèl·lules limfoides velloses (HCL i LZME) i les CLR són els subconjunts amb les medianes més baixes de *relació N/C* (1,5, 1,5 i 1,4, respectivament), mentre que el LF és el grup amb la mediana més elevada (7,7) ($p < 0,0001$). En quant als resultats obtinguts de la comparació múltiple, totes les parelles de cèl·lules difícils de diferenciar, a excepció del binomi HCL-LZME, han mostrat diferències significatives en la *relació N/C* ($p < 0,0001$, Taula 6.5).

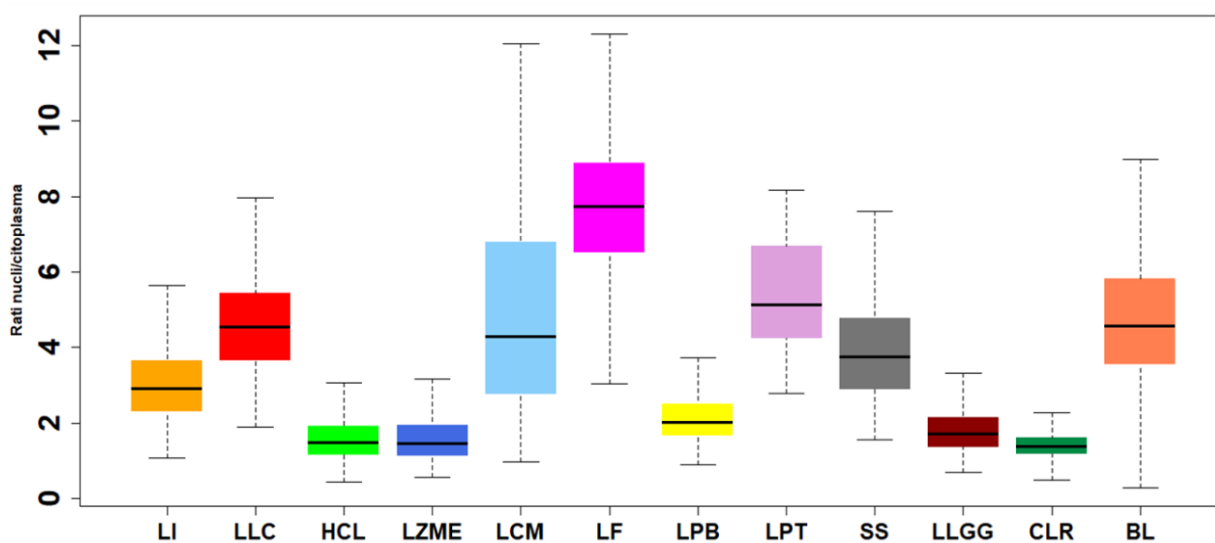


Figura 9.2: Diagrama de caixes corresponents a la *relació nucli/citoplasma* per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.

A la Taula 9.3 es mostren els valors quantitatius dels cinc descriptors geomètrics més rellevants per a diferenciar entre els 12 grups de cèl·lules limfoides obtinguts a partir de les imatges adquirides amb el sistema BX43, els quals són: *relació N/C*, *el perímetre nuclear*, *el diàmetre equivalent cel·lular*, *el hairiness* i *la circularitat nuclear*.

Taula 9.3: Valors quantitius dels descriptors geomètrics més rellevants obtinguts amb el sistema BX43. Els valors que es mostren són les medianes (rang interquartílic).

Posició	1	2	3	17	19
Descriptor	Relació nucli/citoplasma	Perímetre nuclear (píxels)	Diàmetre equivalent (píxels)	<i>Hairiness</i> ($\times 10^{-3}$)	Circularitat nuclear
LI	2,9 (1,4)	226 (21)	82 (11)	82 (11)	0,99 (0,03)
LLC	4,5 (1,8)	223 (25)	79 (9)	79 (9)	1,00 (0,01)
HCL	1,5 (0,8)	257 (28)	105 (16)	105 (16)	1,00 (0,02)
LZME	1,5 (0,8)	260 (30)	107 (21)	107 (21)	1,00 (0,02)
LCM	4,3 (4,0)	261 (50)	92 (22)	92 (22)	0,99 (0,03)
LF	7,7 (2,4)	231 (36)	77 (10)	77 (10)	0,99 (0,04)
LPB	2,0 (0,8)	281 (29)	107 (16)	107 (16)	0,97 (0,04)
LPT	5,1 (2,4)	232 (33)	80 (11)	80 (11)	0,95 (0,07)
SS	3,8 (1,9)	287 (75)	100 (24)	100 (24)	0,98 (0,04)
LLGG	1,7 (0,8)	252 (23)	98 (16)	98 (16)	0,96 (0,06)
CLR	1,4 (0,5)	310 (51)	122 (20)	122 (20)	0,90 (0,10)
BL	4,6 (2,3)	294 (39)	100 (14)	100 (14)	0,94 (0,07)

BL, cèl·lula blàstica limfoide; CLR, cèl·lula limfoide reactiva; HCL, tricoleucèmia; LCM, leucèmia cèl·lules del mantell; LF, limfoma fol·licular; LI, limfòcit normal; LLC, leucèmia limfàtica crònica; LLGG, leucèmia de limfòcits grans granulars; LPB, leucèmia prolimfocítica B; LPT, leucèmia prolimfocítica T; LZME, limfoma de la zona marginal esplènica; SS, síndrome de Sézary.

9.1.2 *Perímetre nuclear*

Pel que fa al *perímetre nuclear* (2n descriptor més rellevant), mesurat en píxels, el coeficient de correlació r obtingut és de 0,950 (veure Taula 9.1), per tant, existeix una relació directa entre els valors del *perímetre nuclear* obtinguts amb els dos sistemes d'adquisició d'imatges. En relació a la intensitat de la correlació, només un 9,7 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust ($R^2 = 0,903$, $p < 0,00001$, veure Taula 9.2 i gràfic 2 Fig. 9.1). En quant als R^2 calculats per cada grup cel·lular, tots ells són superiors a 0,99 (veure Taula 9.2), a excepció del LF (0,985), la LPT (0,970) i la SS (0,921).

Pel que fa als valors obtinguts del *perímetre nuclear* amb el sistema BX43, com ja succeïa amb les imatges adquirides amb el DM96 (Capítol 4), la SS, el LCM i les CLR són els subconjunts cel·lulars que mostren més variabilitat en els seus valors: el rang interquartílic és de 75, 50 i 51 píxels, respectivament (veure Taula 9.3 i Fig. 9.3).

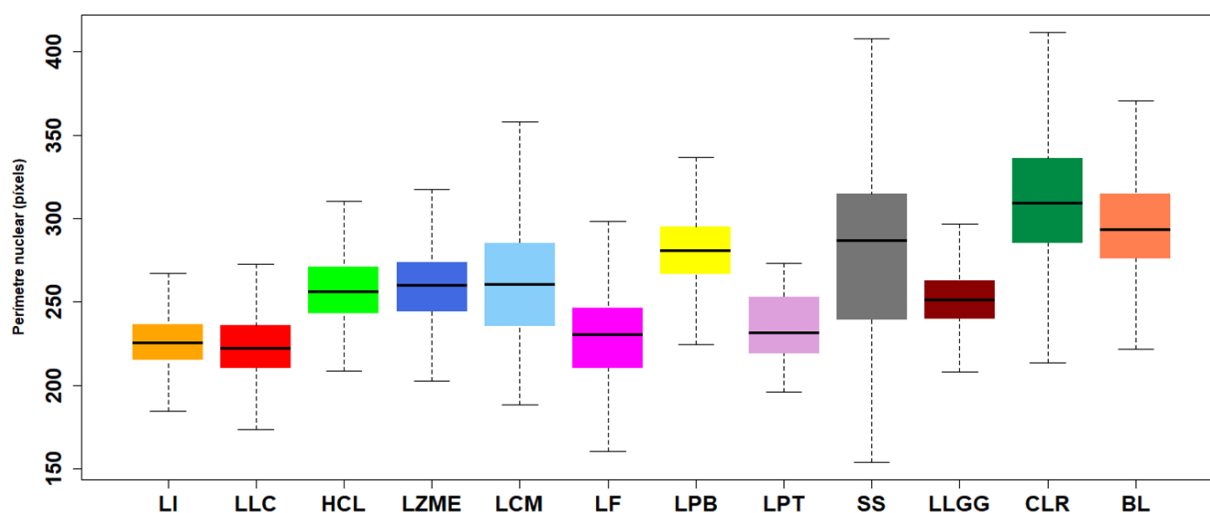


Figura 9.3: Diagrama de caixes corresponents al *perímetre nuclear* per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.

9.1.3 Diàmetre equivalent de la cèl·lula

Com s'ha vist al Capítol 4, el *diàmetre equivalent de la cèl·lula* (3r més rellevant) ressaltava les diferències de mida entre alguns subconjunts de cèl·lules limfoides anormals. El coeficient de correlació obtingut entre ambdós sistemes (DM96 i BX43) és de 0,969 ($p < 0,00001$, veure Taula 9.1). Això es tradueix en que existeix una relació directa entre els valors obtinguts amb ambdós sistemes per als 12 grups de cèl·lules limfoides. En quant a la intensitat de la correlació, només un 6,2 % de la variació del sistema BX43 no s'explica pel sistema DM96 a través de la recta d'ajust ($R^2 = 0,938$, $p < 0,00001$, veure Taula 9.2 i gràfic 3 Fig. 9.1). Si ens fixem en els R^2 de cada grup cel·lular, tots ells són superiors a 0,969 (veure Taula 9.2).

En relació als valors obtinguts del *diàmetre equivalent* per cada tipus de cèl·lula limfoide (veure Fig. 9.4 i Taula 9.3), el grup de les CLR presenta les medianes més altes, 122 píxels, mentre que el LF les més baixes, 77 píxels ($p < 0,0001$). La LPB presenta valors més alts en comparació amb la LPT, 107 i 80 píxels, respectivament ($p < 0,0001$). Els resultats obtinguts per al *diàmetre equivalent de la cèl·lula* en els diferents subconjunts cel·lulars limfoides concorden amb la seva morfologia i amb els resultats obtinguts amb el sistema DM96.

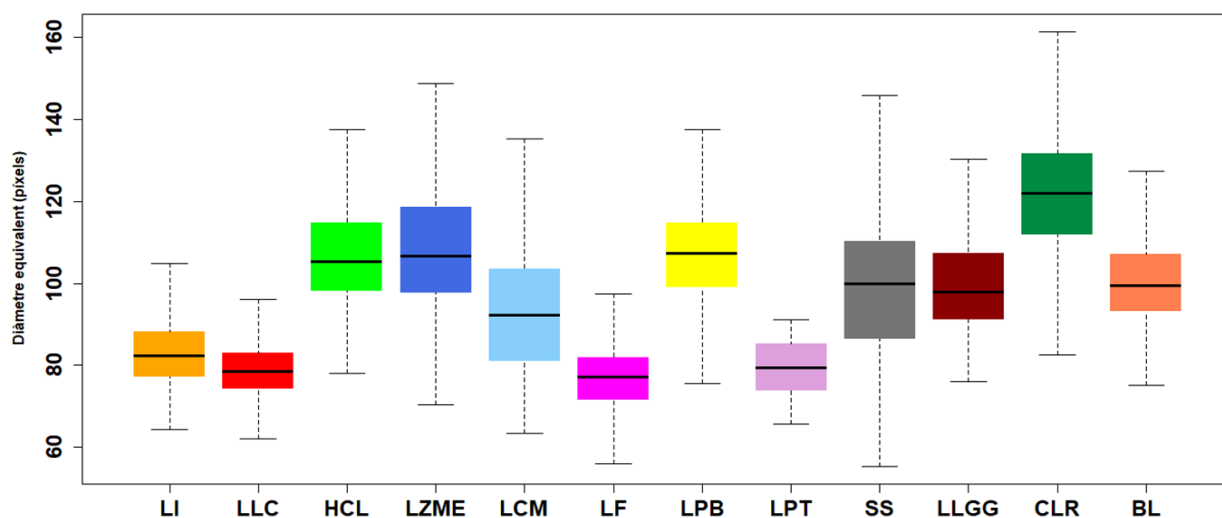


Figura 9.4: Diagrama de caixes corresponents al *diàmetre equivalent* per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.

9.1.4 Curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula

Aquest descriptor és el quart més rellevant obtingut amb el sistema DM96 per a la diferenciació de les 12 cèl·lules limfoides objecte d'estudi del Capítol 4 i és el primer descriptor de textura més rellevant. Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,865 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,749 ($p < 0,00001$), el qual implica que un 24,1 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 4 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,973 a excepció de la LPT (0,953, veure Taula 9.2).

La Figura 9.5 mostra els diagrames de caixes de la *curtosi de la corba pseudo-granulomètrica del cian de la cèl·lula* per als 12 grups de cèl·lules limfoides: la mediana més elevada es troba als BL (3,5), mentre que les medianes més baixes es troben a l'HCL (2,6) i la SS (2,6) ($p < 0,0001$).

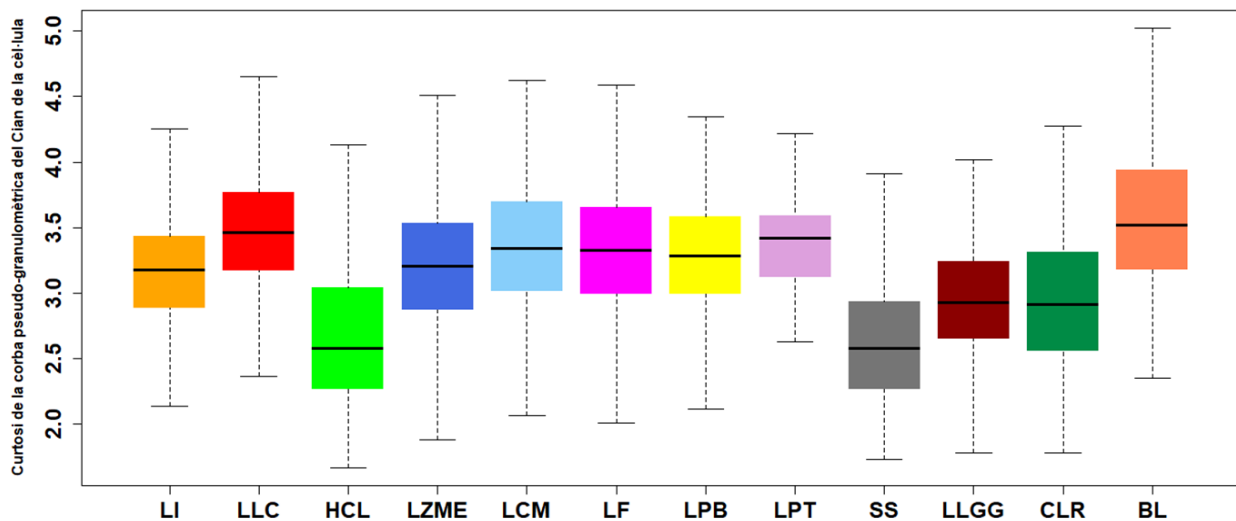


Figura 9.5: Diagrama de caixes corresponents a la *curtosi de la corba pseudogranulomètrica del component cian de la cèl·lula* per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.

9.1.5 Asimetria de l'histograma del blau de la cèl·lula

El cinquè millor descriptor per a la diferenciació dels 12 grups de cèl·lules limfoides, com s'ha presentat al Capítol 4, va resultar ser *l'asimetria de l'histograma del component blau de la cèl·lula*. Es tracta d'un descriptor de color i textura estadístic de primer ordre basat en l'histograma del component de color blau i a la Figura 4.11 se n'ha mostrat un exemple per tal de comprendre millor la seva relació amb les característiques discriminants morfològiques.

Pel que fa al coeficient de correlació r obtingut és de 0,974 (veure Taula 9.1), per tant, existeix una relació directa entre els valors de *l'asimetria de l'histograma del component blau de la cèl·lula* obtinguts amb els dos sistemes d'adquisició d'imatges. En relació a la intensitat de la correlació, només un 5,1 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust ($R^2 = 0,949$, $p < 0,00001$, veure Taula 9.2 i gràfic 5 Fig. 9.1). En quant als R^2 calculats per cada grup cel·lular, tots ells són superiors a 0,980 (veure Taula 9.2), a excepció de la LPT (0,901).

La Figura 9.6 conté els diagrames de caixes corresponents a aquest descriptor. Les medianes més elevades s'observen als limfòcits anormals de la SS (1,04), LPT (1,03) i LLC (1,00), mentre que les més baixes es troben al LZME (0,37) i les CLR (0,21) ($p < 0,0001$).

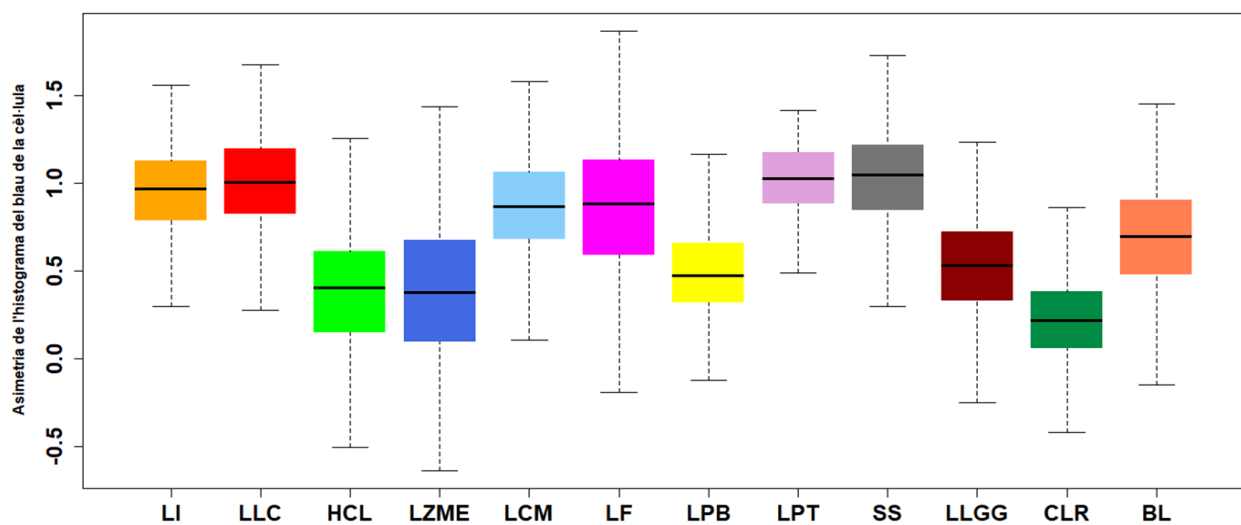


Figura 9.6: Diagrames de caixes corresponents a *l'asimetria de l'histograma del component blau de la cèl·lula* per als 12 grups de cèl·lules limfoides adquirides amb el sistema BX43.

9.1.6 Information measure of correlation 1 del magenta de la cèl·lula

El sisè millor descriptor per a la diferenciació dels 12 grups de cèl·lules limfoides, com s'ha presentat al Capítol 4, va resultar ser la *information measure of correlation 1 (IMC1) del magenta de la cèl·lula*.

Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,799 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *IMC1 del magenta de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,639 ($p < 0,00001$), el qual implica que un 36,1 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 6 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,971 a excepció del LCM (0,948) i la LPT (0,96, Taula 9.2).

En relació als valors obtinguts de la *IMC1 del magenta de la cèl·lula* per cada tipus de cèl·lula limfoide (veure Fig. 9.7), el LF presenta la mediana més elevada (-0,40), mentre que les CLR la més baixa (-0,61) ($p < 0,0001$). La LPB presenta valors més baixos en comparació amb la LPT, -0,55 i -0,40, respectivament ($p < 0,0001$).

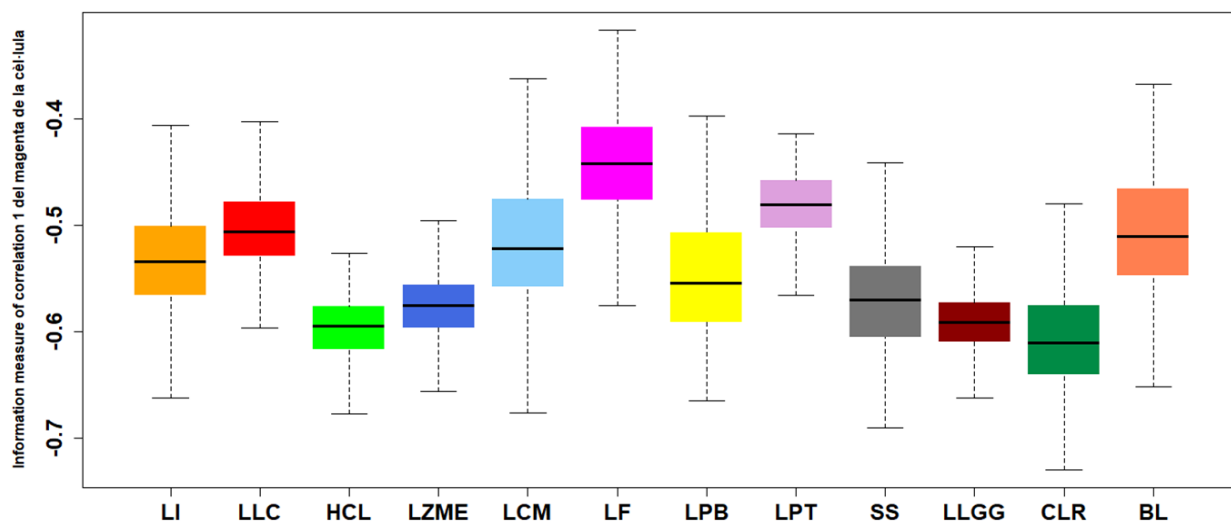


Figura 9.7: Diagrames de caixes corresponents a la *Information measure of correlation 1 del component magenta de la cèl·lula* per als 12 grups de cèl·lules limfoides amb el sistema BX43.

9.1.7 Desviació estàndard de l'histograma del component blau de la cèl·lula

La desviació estàndard (DE) de l'histograma del component blau de la cèl·lula és el setè descriptor més rellevant i es calcula a partir del mateix histograma que el cinquè descriptor, però en aquest cas el paràmetre estadístic és la desviació estàndard de la distribució. Després de realitzar el test de correlació de Pearson, el coeficient de correlació r és de 0,945 ($p < 0,00001$, veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la DE de l'histograma del component blau de la cèl·lula obtinguts amb el sistema DM96 i el BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,894 ($p < 0,00001$), el qual implica que un 10,6 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 7 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,974 (veure Taula 6.4) a excepció del LF (0,96) i la LPT (0,958).

La Figura 9.8 conté els diagrames de caixes corresponents a aquest descriptor. Les medianes més baixes d'aquest descriptor s'observen al LF (9,6), a la LPT (9,4) i als BL (10,0), mentre que els limfòcits vellosos de l'HCL (17,1) i el LZME (15,0) i la LLGG (16,0) mostren les més altes ($p < 0,0001$).

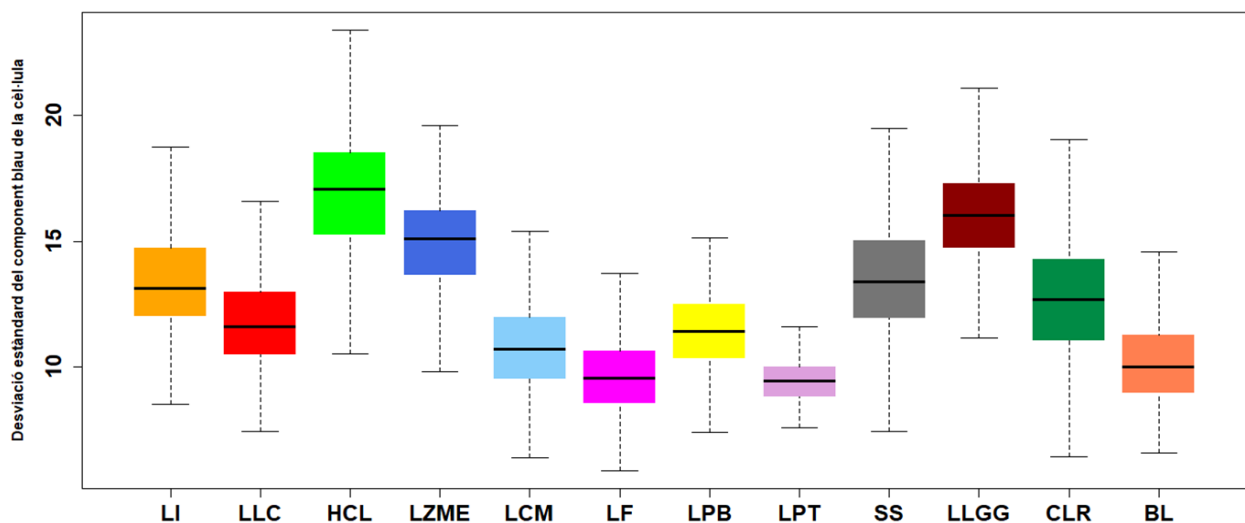


Figura 9.8: Diagrama de caixes corresponents a la desviació estàndard de l'histograma del component blau de la cèl·lula per als 12 grups de cèl·lules limfoides amb el sistema BX43.

9.1.8 Entropia de l'histograma del component magenta de la regió de la cèl·lula

El vuitè descriptor més rellevant correspon a l'entropia de l'histograma del component magenta de la regió de la cèl·lula. Com hem vist al Capítol 4, l'entropia és una mesura d'incertesa o d'aleatorietat. És a dir, com més alta sigui l'entropia, més informació s'inclourà.

Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,974 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de l'entropia de l'histograma del component magenta de la regió de la cèl·lula obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,948 ($p < 0,00001$), el qual implica que només un 5,2 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 8 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,987 a excepció de la LPT (0,977, veure Taula 9.2).

La Figura 9.9 conté els diagrames de caixes corresponents a aquest descriptor i es pot veure que les CLR, els limfòcits anormals de l'HCL, LZME, LLGG i LPB són els grups de cèl·lules limfoides que presenten valors més alts d'entropia (6,8, 6,8, 6,8, 6,7 i 6,5, respectivament). Per contra, els limfòcits anormals del LF i els prolimfòcits T mostren els valors més baixos d'aquest descriptor (5,5 i 5,7, respectivament, $p < 0,0001$).

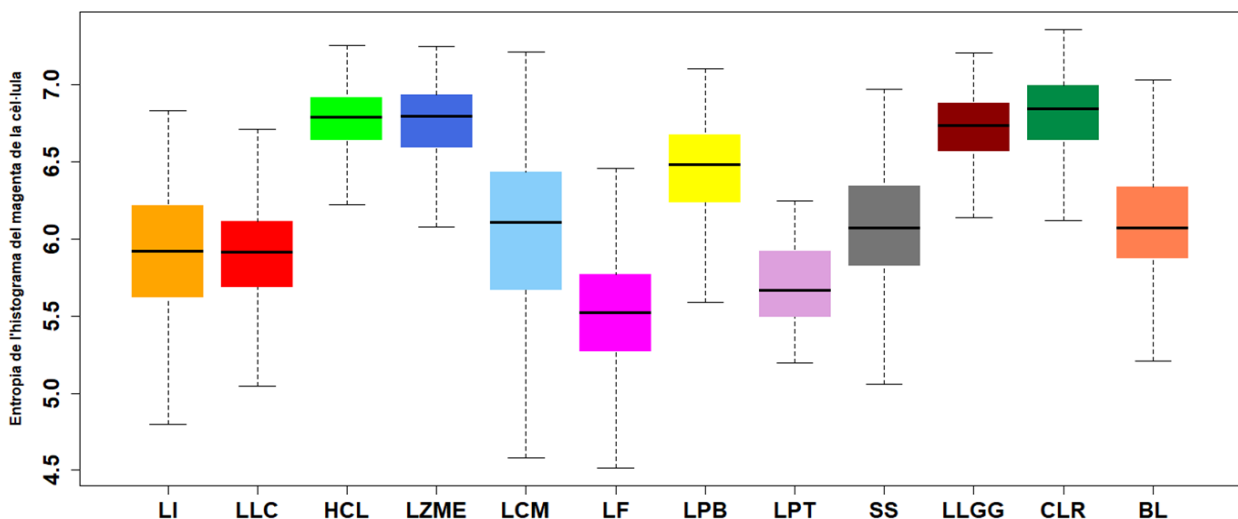


Figura 9.9: Diagrames de caixes obtinguts per al descriptor de l'entropia de l'histograma del magenta de la cèl·lula per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.9 Asimetria de l'histograma del component verd de la cèl·lula

El novè descriptor obtingut és *l'asimetria de l'histograma del component verd de la cèl·lula*, el qual és un descriptor de primer ordre i està relacionat amb la intensitat de verd de tota la cèl·lula (veure Capítol 4 secció 4.3.1 Fig.4.19).

Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,929 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de *l'asimetria de l'histograma del component verd de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,863 ($p < 0,00001$), el qual implica que només un 14,7 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 9 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,982 a excepció de la LPT (0,945, veure Taula 9.2).

Al Capítol 4 hem vist que les cèl·lules limfoides que tenen una baixa relació N/C presenten valors d'aquest descriptor més baixos a causa de la quantitat similar de píxels verds brillants i foscos que hi ha a les dues regions cel·lulars (nucli i citoplasma). Com podem veure al diagrama de caixes de la Figura 9.10, això també es compleix amb les imatges adquirides amb el BX43. Les CLR i els limfòcits vellosos (LZME i HCL) són els subconjunts de cèl·lules limfoides que mostren les medianes més baixes (0,54, 0,61 i 0,65, respectivament) mentre que les més altes s'observen al LF i als prolimfòcits T (2,45 i 2,15) ($p < 0,0001$).

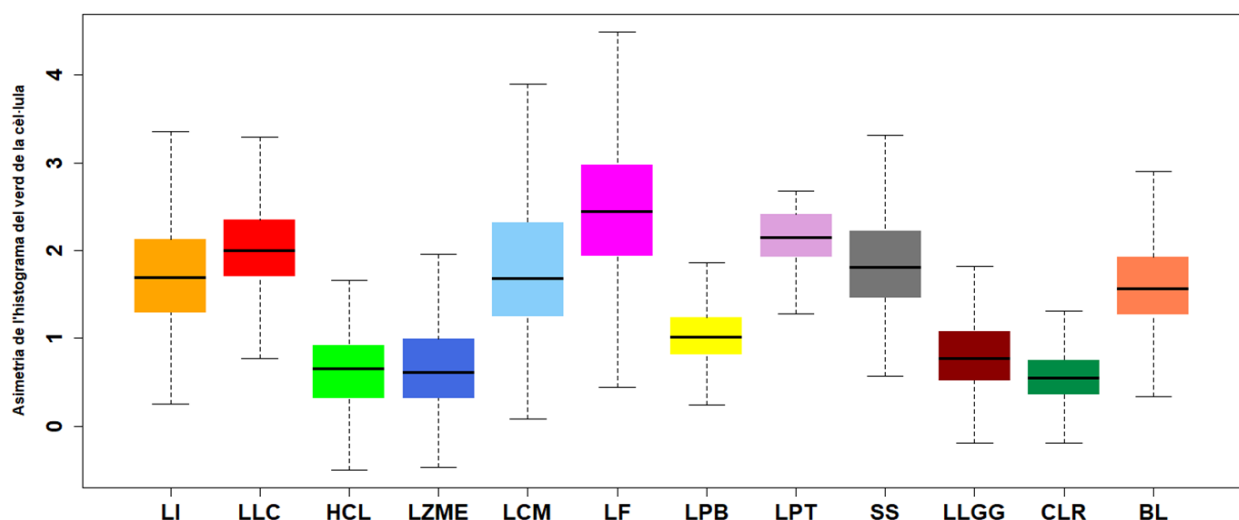


Figura 9.10: Diagrames de caixes obtinguts per al descriptor de *l'asimetria de l'histograma del component verd de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.10 Mitjana de la corba pseudo-granulomètrica del negre del citoplasma

El descriptor més rellevant obtingut del citoplasma és el desè descriptor: *la mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma*. Com més contribueix el component negre a la regió del citoplasma, major valor s'obté d'aquest descriptor.

En quant al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,872 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides. El LF, com no conté gairebé citoplasma habitualment, aquest descriptor pren valors de zero o propers a zero, fet que no ha permès correlacionar els valors amb el sistema DM96.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,761 ($p < 0,00001$), el qual implica que un 23,9 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 10 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,972 (veure Taula 9.2).

A la Figura 9.11 es mostren els diagrames de caixes per aquest descriptor. Com ja hem vist amb el sistema DM96, els prolimfòcits B i els limfòcits normals presenten les medianes més altes (0,292 i 0,295, respectivament), mentre que els limfòcits anormals del LF les més baixes (0,0) ($p < 0,0001$).

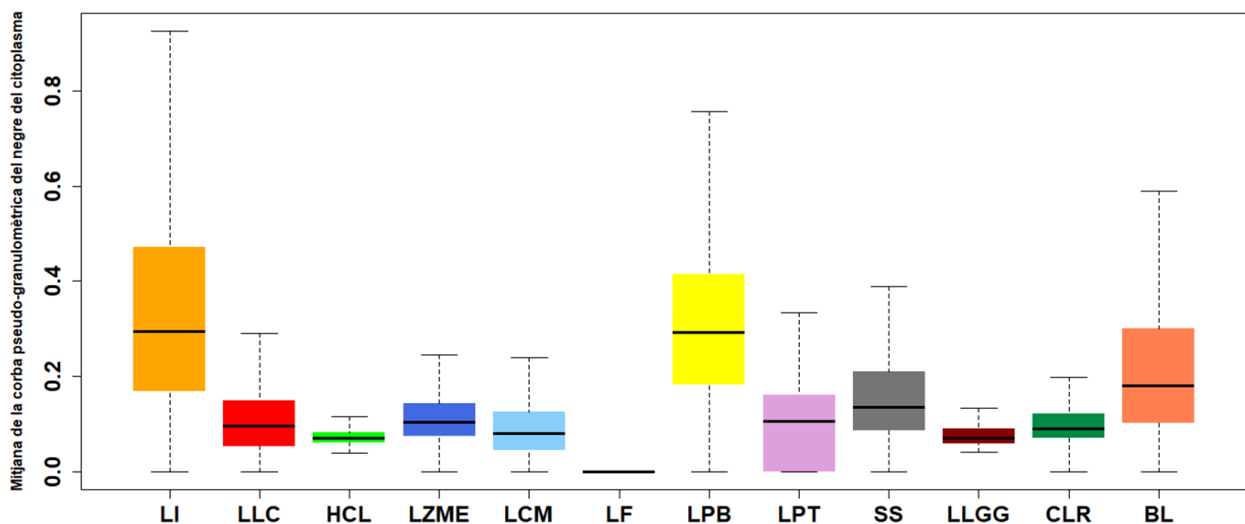


Figura 9.11: Diagrames de caixes obtinguts per al descriptor de la *mitjana de la corba pseudo-granulomètrica del component negre de la regió del citoplasma* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43

9.1.11 Mitjana de l'histograma del component blau del nucli

La mitjana de l'histograma del component blau del nucli és el descriptor més important obtingut del nucli i ocupa l'onzena posició. En relació al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,955 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la mitjana de l'histograma del component blau del nucli obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,998 ($p < 0,00001$), el qual implica que només un 0,2 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 11 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,970 a excepció dels dos tipus de prolimfòcits (veure Taula 9.2).

La Figura 9.12 conté els corresponents diagrames de caixes. La mediana més alta es troba a les CLR (168), mentre que la més baixa als prolimfòcits T (144) ($p < 0,0001$). Aquests resultats coincideixen amb els obtinguts amb el sistema DM96.

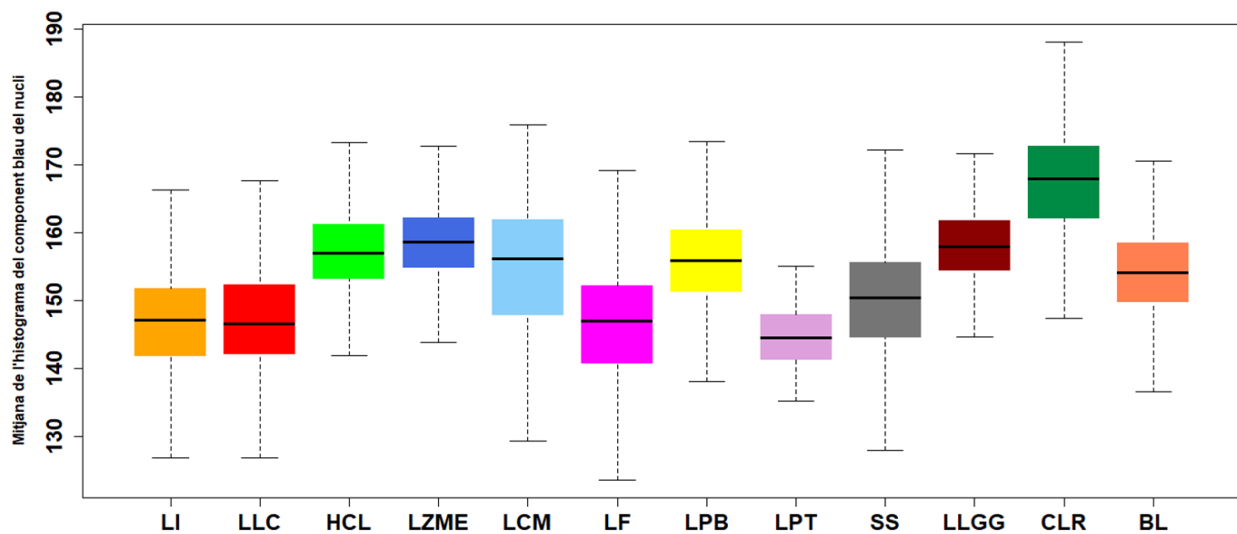


Figura 9.12: Diagrames de caixes obtinguts per al descriptor de la mitjana de l'histograma del component blau del nucli per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.12 Cluster shade de la saturació de la cèl·lula

Aquest descriptor va resultar ser el 12è més rellevant en la diferenciació dels 12 tipus de cèl·lules limfoides i es tracta d'un descriptor de color i textura estadístic de segon ordre. A l'anàlisi de correlació de Pearson s'ha obtingut un coeficient de correlació r entre ambdós sistemes de 0,696 (veure Taula 9.1), el qual indica que no existeix una relació directa entre els valors del *cluster shade de la saturació de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,485 ($p < 0,00001$), el qual implica que un 51,5 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 12 Fig. 9.1). Per tant, aquest descriptor no es correlaciona entre ambdós sistemes d'adquisició d'imatges.

A la Figura 9.13 es mostren els diagrames de caixes per aquest descriptor. Els limfòcits anormals del LF presenten els valors més elevats (-7,2) mentre que els limfòcits grans granulars i els tricoleucòcits mostren els valors més baixos (-56,4 i -51,0 respectivament) ($p < 0,0001$).

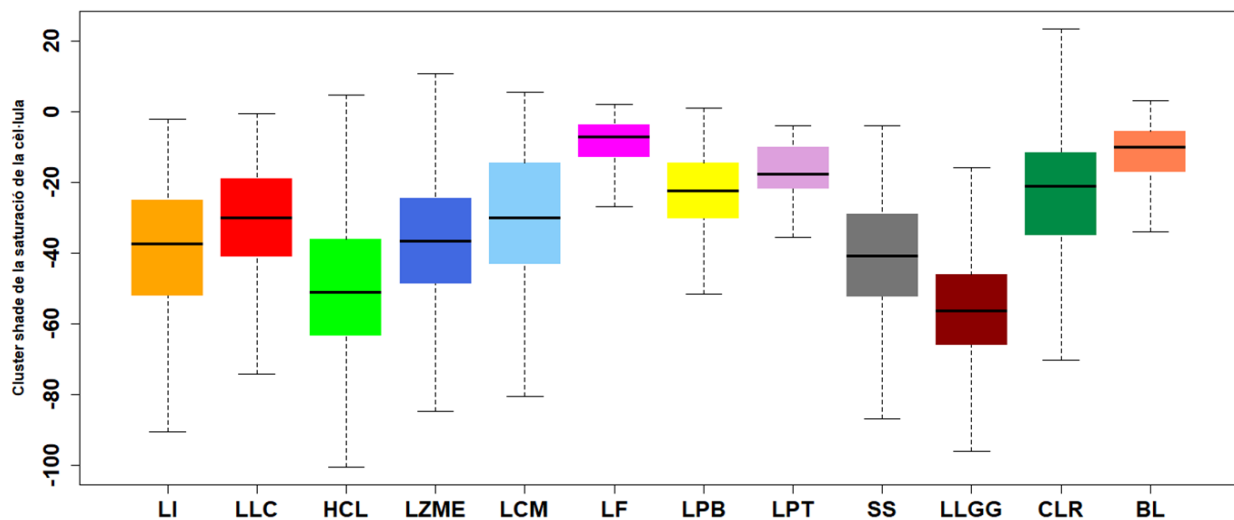


Figura 9.13: Diagrames de caixes obtinguts per al descriptor del *cluster shade de la saturació de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.13 Mitjana de l'histograma del magenta del citoplasma

Pel que fa a la mitjana de l'histograma del component magenta de la regió del citoplasma, el coeficient de correlació r obtingut entre ambdós sistemes ha estat de 0,835 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la mitjana de l'histograma del component magenta del citoplasma obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,697 ($p < 0,00001$), el qual implica que un 30,3 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 13 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,982 a excepció de la LPT (0,877, veure Taula 9.2).

La Figura 9.14 mostra els corresponents diagrames de caixes. La mediana més alta es troba als prolimfòcits B (169), mentre que la més baixa als limfòcits anormals del LF (0,0) ($p < 0,0001$).

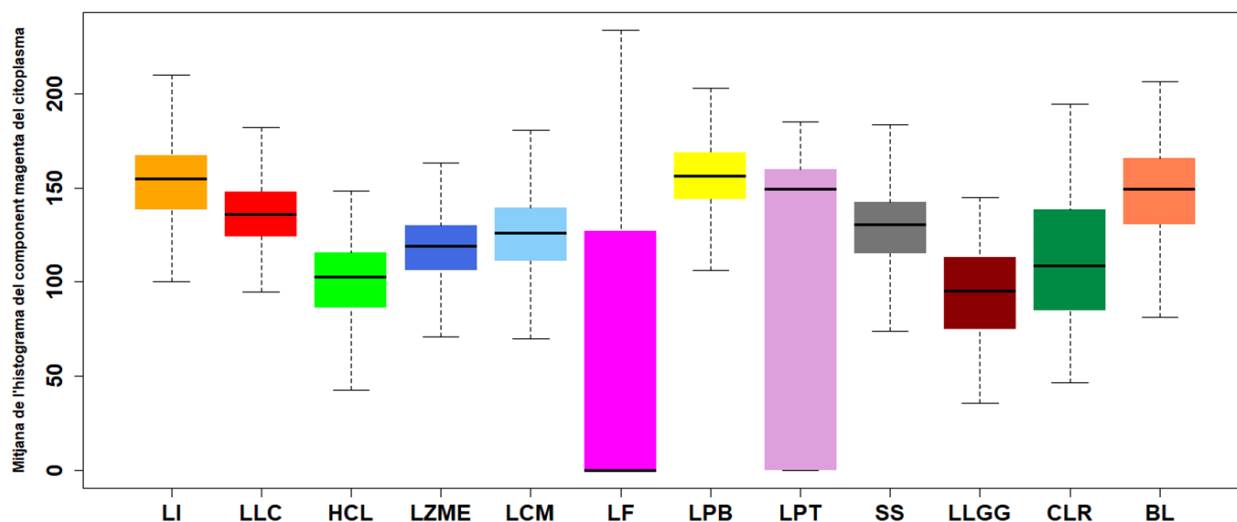


Figura 9.14: Diagrames de caixes obtinguts per al descriptor de la mitjana de l'histograma del component magenta de la regió del citoplasma per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.14 Homogeneïtat del component magenta de la cèl·lula

El tretzè descriptor més rellevant correspon a l'*homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula*. Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,872 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de l'*homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,76 ($p < 0,00001$), el qual implica que un 24 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 14 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,978 a excepció de la LPT (0,957, veure Taula 9.2).

Com es pot veure al diagrama de caixes de la Figura 9.15, els limfòcits vellosos del LZME presenten les medianes més baixes (0,87) mentre que les més altes s'observen a les cèl·lules de Sézary (0,90) ($p < 0,0001$).

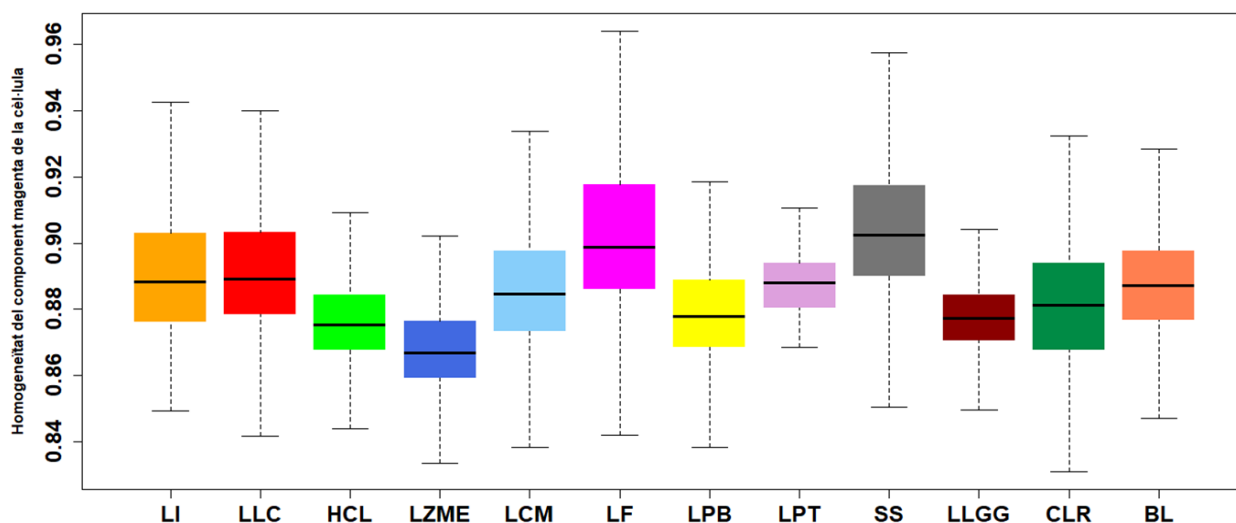


Figura 9.15: Diagrames de caixes obtinguts per al descriptor de *la homogeneïtat de l'histograma del component magenta de la regió de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.15 Curtosi de l'histograma de la saturació de la cèl·lula

Aquest descriptor és el 15è més rellevant obtingut amb el sistema DM96 per a la diferenciació de les 12 cèl·lules limfoides objecte de la present *Tesi Doctoral*. Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,893 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *curtosi de l'histograma del component de la saturació de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,797 ($p < 0,00001$), el qual implica que un 20,3 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 15 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,975 a excepció del LCM i de la LPB (veure Taula 9.2).

La Figura 9.16 conté els corresponents diagrames de caixes. Els valors més elevats es troben en les cèl·lules amb una relació N/C elevada com és el cas del LF (5,7), la LPT (4,1) o la LLC (3,2) mentre que les medianes més baixes es troben en els tipus de cèl·lules limfoides que tenen una relació N/C baixa com l'HCL (-1,4), el LZME (-1,3), la LLGG (-1,1) i les CLR (-1,1) ($p < 0,0001$).

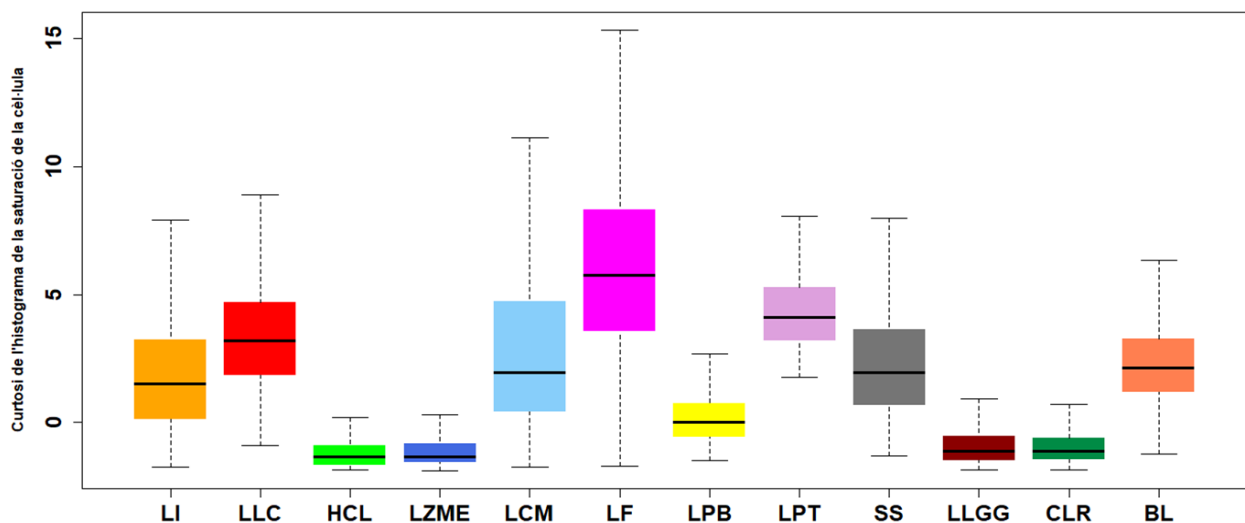


Figura 9.16: Diagrames de caixes obtinguts per al descriptor de la curtosi de l'histograma de la saturació de la regió de la cèl·lula per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.16 *Information measure of correlation 1 del cian del nucli*

Aquest és el segon descriptor més rellevant obtingut del nucli i ocupa la setzena posició. Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,859 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *IMC1 del component cian del nucli* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,737 ($p < 0,00001$), el qual implica que un 26,3 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 16 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,978 a excepció de la LPB i la LLGG (veure Taula 9.2).

La Figura 9.17 conté els corresponents diagrames de caixes. La mediana més alta es troba a les cèl·lules blàstiques limfoides (-0,30), mentre que la més baixa als limfòcits de la LLC (-0,36) ($p < 0,0001$). Aquests resultats coincideixen amb els obtinguts amb el sistema DM96.

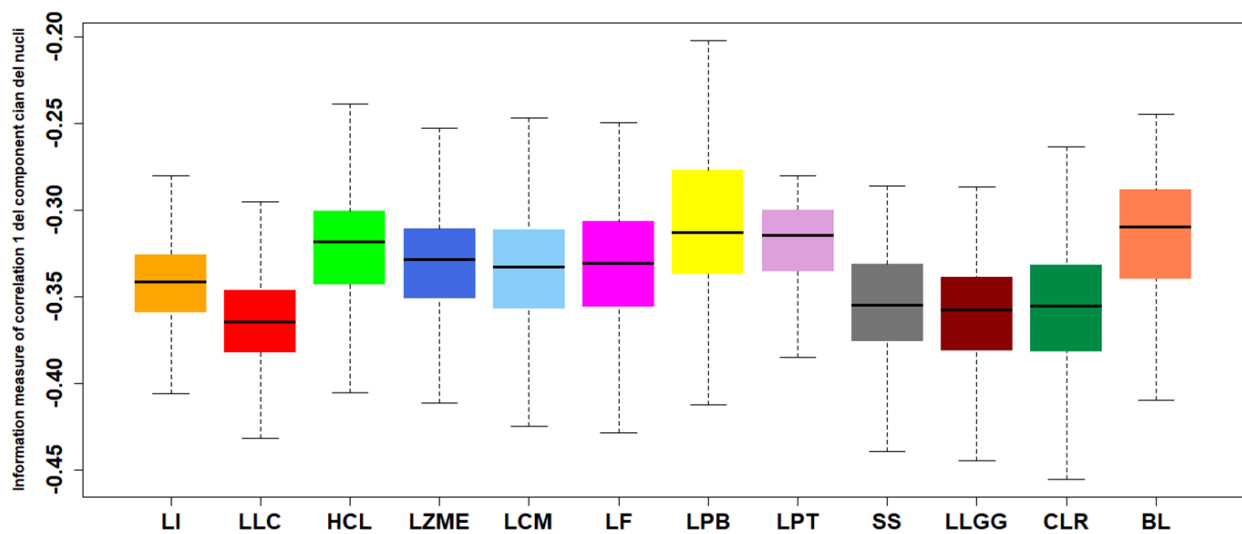


Figura 9.17: Diagrames de caixes obtinguts per al descriptor de la *Information measure of correlation 1 del component cian de la regió del nucli* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.17 *Hairiness*

El descriptor del perfil citoplasmàtic o *hairiness* amb la formulació ajustada per les imatges del sistema BX43 ha presentat la mateixa especificitat pels limfòcits vellosos que la demostrada amb el sistema DM96. El coeficient de correlació r obtingut entre ambdós sistemes ha estat de 0,871 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de *hairiness* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,76 ($p < 0,00001$), el qual implica que un 24 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 17 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,941 a excepció de les cèl·lules limfoides reactives (0,869, veure Taula 9.2).

A la Taula 9.3 es mostren els valors de *hairiness* per als 12 grups de cèl·lules limfoides i la Figura 9.18 presenta els corresponents diagrames de caixes. Es confirma que aquest descriptor, calculat sobre les imatges adquirides amb el sistema BX43, és específic de les cèl·lules limfoides velloses, permetent discriminar els limfòcits vellosos d'altres limfòcits anormals.

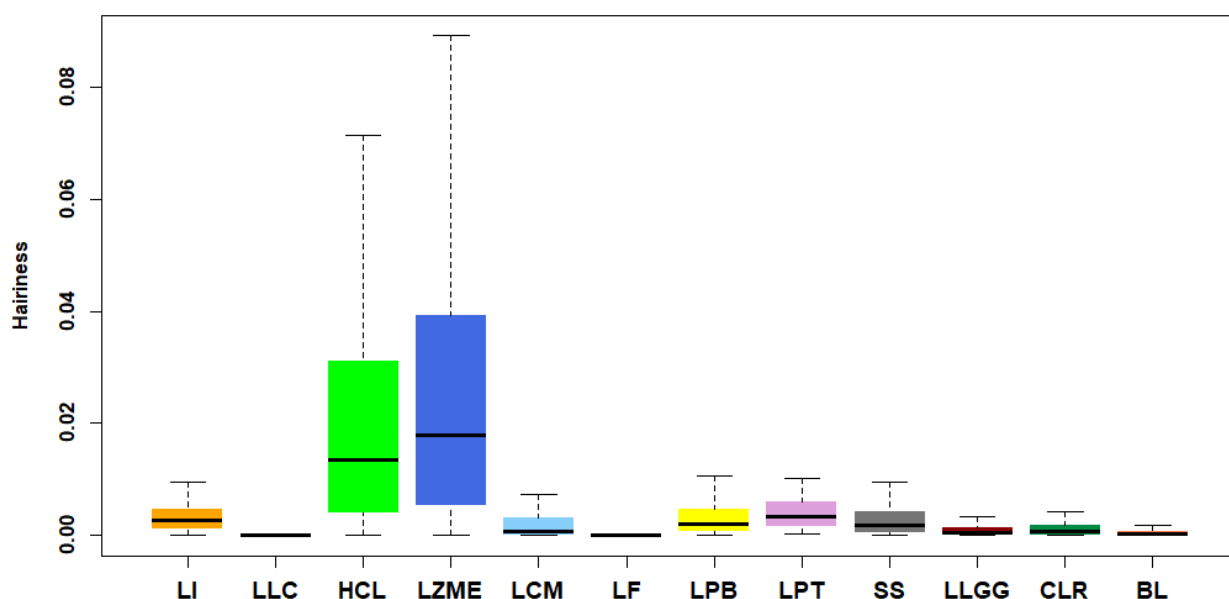


Figura 9.18: Diagrames de caixes obtinguts per al descriptor *hairiness* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.18 *Desviació estàndard de l'histograma de la saturació de la cèl·lula*

El divuitè descriptor més rellevant correspon a la *desviació estàndard de l'histograma de la saturació de la cèl·lula*. En quant al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,945 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *desviació estàndard de l'histograma de la saturació de la cèl·lula* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,893 ($p < 0,00001$), el qual implica que un 10,7 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 18 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,980 a excepció de la LPT (0,947, veure Taula 9.2).

La Figura 9.19 conté els corresponents diagrames de caixes. Els valors més elevats es troben a l'HCL (51,5) i a la LLGG (48,5) mentre que les medianes més baixes es troben al LF (18,5) i a la LPT (20,5) ($p < 0,0001$). Aquests resultats coincideixen amb els obtinguts amb el sistema DM96.

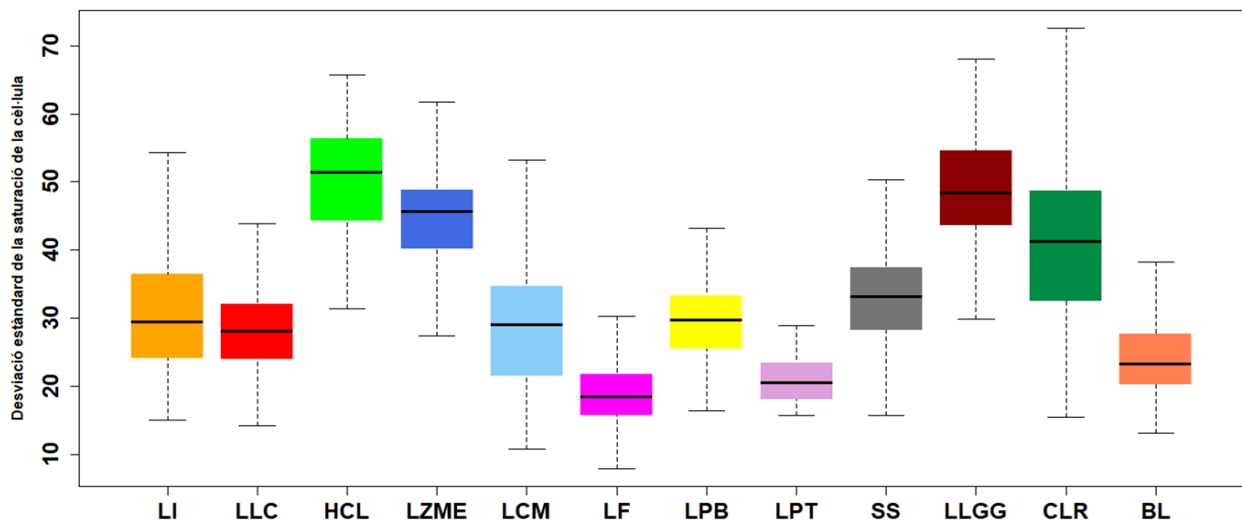


Figura 9.19: Diagrames de caixes obtinguts per al descriptor de la *desviació estàndard de l'histograma de la saturació de la cèl·lula* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.19 *Circularitat nuclear*

La *circularitat nuclear* (19a posició), el valor de la qual va des de zero (baixa circularitat nuclear) fins a un (nucli rodó), presenta un coeficient de correlació r entre ambdós sistemes de 0,940 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de la *circularitat nuclear* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,893 ($p < 0,00001$), el qual implica que un 10,7 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 19 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,989 a excepció de la LPT (0,961, veure Taula 9.2).

A la Taula 9.3 es mostren els valors de *circularitat nuclear* per als 12 grups de cèl·lules limfoides i la Figura 9.20 presenta els corresponents diagrames de caixes. Els limfòcits anormals de la LLC i HCL mostren els valors més alts (1,00 i 1,00, respectivament), mentre que les CLR, els BL i la LPT presenten els valors més baixos (0,90, 0,93 i 0,94, respectivament, $p < 0,0001$).

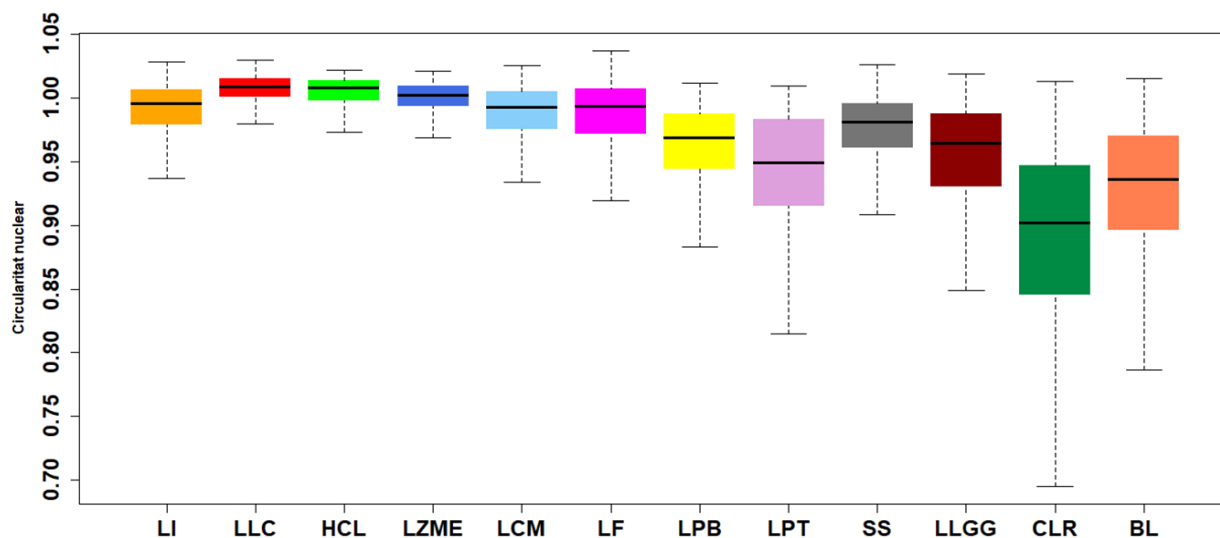


Figura 9.20: Diagrames de caixes obtinguts per al descriptor de la *circularitat nuclear* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.1.20 Homogeneïtat del negre del citoplasma

El vintè descriptor obtingut és *l'homogeneïtat del component negre de la regió del citoplasma*, la qual és un descriptor de primer ordre i està relacionat amb la intensitat de negre de tota la cèl·lula. Pel que fa al coeficient de correlació r obtingut entre ambdós sistemes, aquest ha estat de 0,904 (veure Taula 9.1), el qual indica que existeix una relació directa entre els valors de *l'homogeneïtat del component negre de la regió del citoplasma* obtinguts del sistema DM96 i del BX43 per als 12 grups de cèl·lules limfoides.

En quant a la intensitat de la correlació, el coeficient de determinació (R^2) és igual a 0,817 ($p < 0,00001$), el qual implica que un 18,3 % de la variació del sistema BX43 no és explicat pel sistema DM96 a través de la recta d'ajust (veure Taula 9.2 i gràfic 20 Fig. 9.1). Si ens fixem en els coeficients de determinació calculats per cada grup cel·lular, tots ells són superiors a 0,963 a excepció de la LPT (0,92, veure Taula 9.2).

La Figura 9.21 conté els corresponents diagrames de caixes. El valor més elevat es troba a les cèl·lules limfoides anormals del LF (1,00) mentre que la mediana més baixa es troba a la LLC (0,69) ($p < 0,0001$).

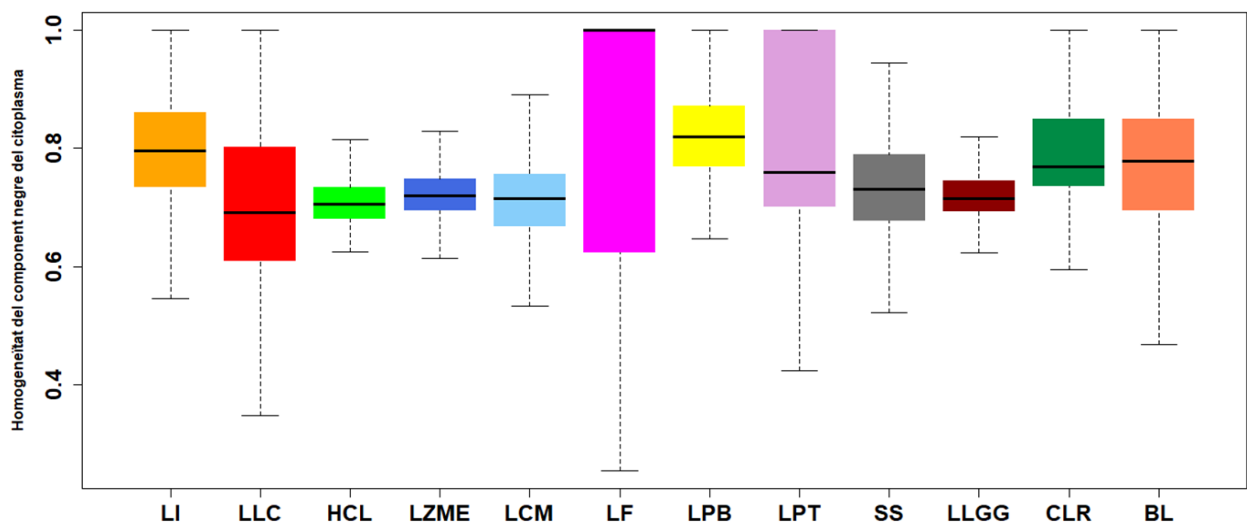


Figura 9.21: Diagrames de caixes obtinguts per al descriptor de la *homogeneïtat del component negre del citoplasma* per als diferents grups de cèl·lules limfoides incloses a l'estudi amb el sistema BX43.

9.2 Annex B

En aquest annex B es mostren els resultats de la correlació de tres dels quatre descriptors analitzats al Capítol 5, amb utilitat en la quantificació de característiques citològiques específiques, obtinguts mitjançant els dos sistemes d'adquisició d'imatges (DM96 i BX43), ja que els resultats de correlació del descriptor *hairiness* s'han detallat a l'Annex A secció 9.1.17.

A la Figura 9.22 i a la Taula 9.4 es mostren els resultats de la correlació de Pearson de la *correlació del cian del nucli*, la *desviació estàndard (DE) de la corba granulomètrica del cian del nucli* i l'*asimetria de l'histograma del component u del citoplasma*.

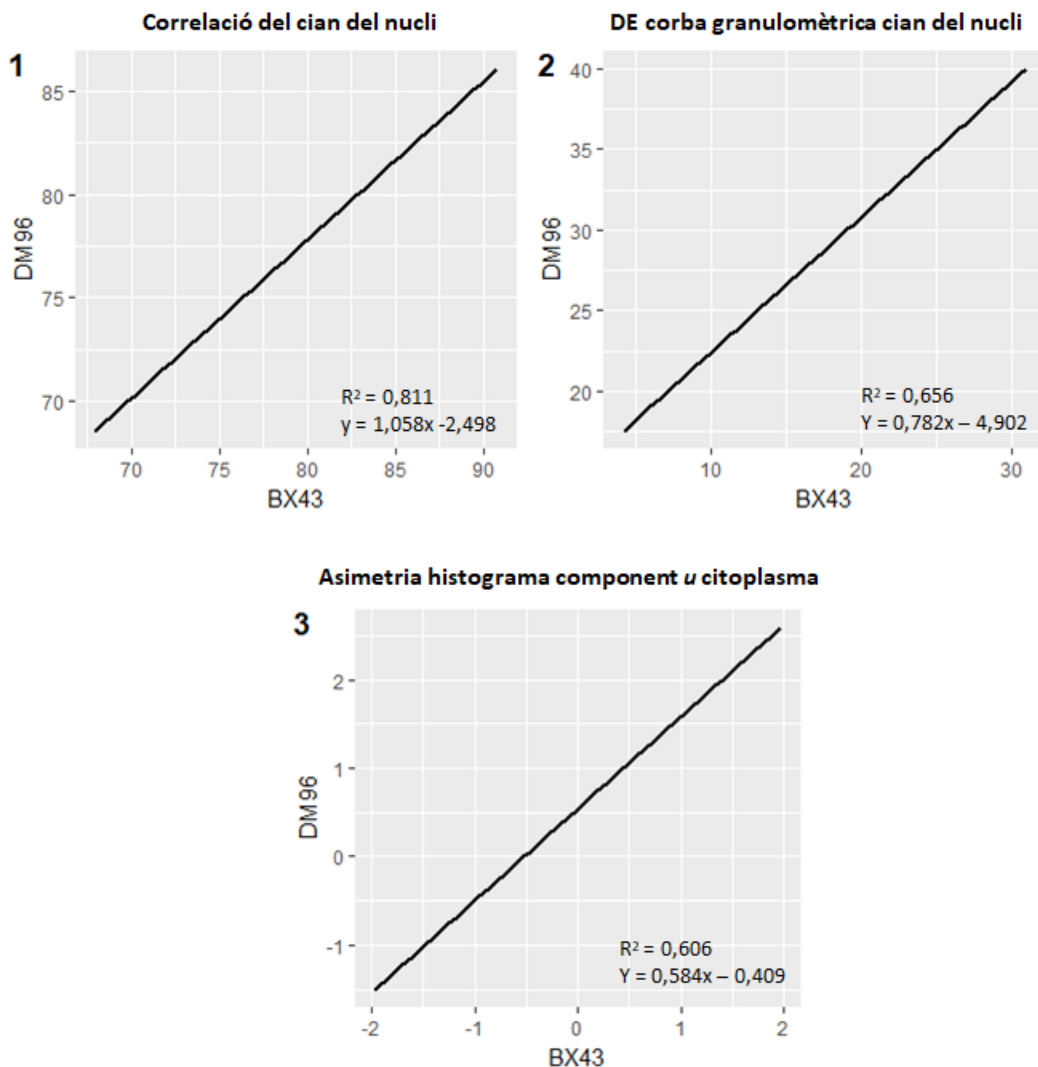


Figura 9.22: Correlacions de Pearson dels tres descriptors entre el sistema DM96 i el BX43 per a tots els grups cel·lulars

Taula 9.4: Coeficient R de Pearson, recta d'ajust, interval de confiança del 95 % i coeficients de determinació (R²) dels tres descriptors per a cada tipus cel·lular limfoide obtinguts entre el sistema DM96 i el BX43.

Descriptor	<i>Correlació del cian del nucli (x 10⁻²)</i>	<i>DE de la corba granulomètrica del cian del nucli (x 10⁻³)</i>	<i>Asimetria de l'histograma del component u del citoplasma</i>
Coeficient R de Pearson	0,901	0,810	0,779
Recta d'ajust entre sistema BX43 i DM96	y = 1,058x - 2,498	y = 0,782x - 4,902	y = 0,584x - 0,409
Interval de confiança del 95 %	0,896 - 0,905	0,802 - 0,818	0,770 - 0,787
Coeficients R²	0,811	0,656	0,606
Coeficients R² per grup cel·lular			
Limfòcit	0,991	0,947	0,975
Limfòcit anormal de la leucèmia limfàtica crònica	0,999	0,968	0,978
Limfòcit vellós de la tricoleucèmia	0,995	0,992	0,998
Limfòcit vellós del limfoma de la zona marginal esplènica	0,996	0,994	0,975
Limfòcit anormal del limfoma fol·licular	0,983	0,928	0,877
Cèl·lula del limfoma del mantell morfologia blàstica	0,982	0,882	0,917
Cèl·lula del limfoma del mantell morfologia clàssica	0,997	0,980	0,975
Cèl·lula plasmàtica	0,983	0,986	0,995
Prolimfòcit B	0,981	0,994	0,993
Prolimfòcit T	0,978	0,974	0,962
Cèl·lula de Sézary variant clàssica	0,996	0,980	0,994
Cèl·lula de Sézary variant Lutzner o petita	0,994	0,898	0,99
Limfòcit gran granular	0,989	0,954	0,984
Cèl·lula limfoide reactiva	0,998	0,990	0,986
Blast limfoide	0,981	0,985	0,997

