DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS

Carlos López Pablo

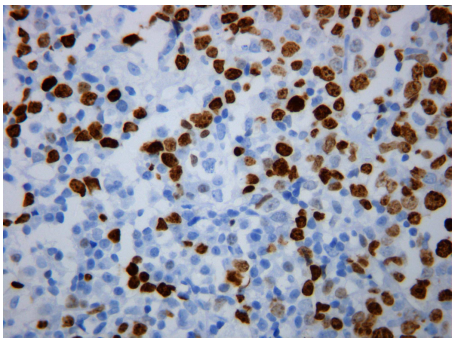
ISBN: 978-84-694-1253-4
Dipòsit Legal: T-315-2011

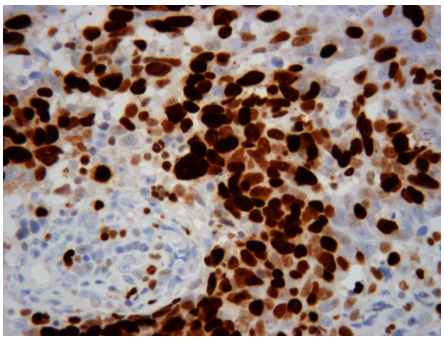
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Development of a new automated method for the quantification of nuclear immunohistochemical markers

Doctoral Thesis

Carlos López Pablo





Medicine and Surgery Department

UNIVERSITAT ROVIRA I VIRGILI

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Carlos López Pablo

DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS.

DOCTORAL THESIS

Supervised by Dr. Alfredo Bardají Ruíz

Medicine and Surgery Department

UNIVERSITAT ROVIRA I VIRGILI

Tarragona
2010

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT
ROVIRA I VIRGILI
Departament de Medicina i Cirurgia

Sant Llorenç, 21
43201 Reus
Tel.: 977759305
Fax: 977759322

Alfredo Bardají Ruíz, professor titular del Departament de Medicina i Cirurgia de la Universitat Rovira i Virgili.

CERTIFICO:

Que aquest treball, titulat "Development of a new automated method for the quantification of nuclear immunohistochemical markers", que presenta Carlos López Pablo per a l'obtenció del títol de Doctor, ha estat realitzat sota la meva direcció al departament de Medicina i Cirurgia d'aquesta universitat i que acompleix els requeriments per poder optar a Menció Europea.

Tarragona, 10 de Maig de 2010

Dr. Alfredo Bardají Ruiz

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

HTVC
Hospital de Tortosa
Verge de la Cinta
InstitutCatalàdelaSalut

Esplanetes, 14
43500 Tortosa
Telèfon 977 51 91 00
Telefax 977 51 91 11

Marylène Lejeune, coordinadora de l'Àrea de Biologia Molecular i Recerca de l'Hospital de Tortosa Verge de la Cinta,

CERTIFICA:
Que el present treball titulat: "Development of a new automated method for the quantification of nuclear immunohistochemical markers", que presenta Carlos López Pablo per a optar al grau de Doctor amb menció europea, ha estat realitzat sota la meva co-direcció a l'Àrea de Biologia Molecular i Recerca i al Servei d'Anatomia Patològica de l'Hospital de Tortosa Verge de la Cinta i que tots els resultats presentats i la seva anàlisi són fruit de la investigació realitzada per l'esmentat doctorand.

Tortosa, 28 de maig de 2010

Dra. Marylène Lejeune

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

AGRAÏMENTS.

Institucionals:

Agrair al *Instituto de Salud Carlos III* pel seu suport econòmic pels dos projectes que han permès realitzar aquesta tesi.

A *l'Hospital de Tortosa Verge de la Cinta* i al *Servei d'Anatomia Patològica* per haver permès realitzar aquest treball en les seves instal·lacions i desenvolupar la feina investigadora.

A la *Fundació Dr. Ferran* i a *l'Institut d'Investigació Sanitari Pere i Virgili* pel seu suport econòmic i institucional a l'hora de realitzar aquesta tesi.

A *l'Agència de Gestió d'Ajuts Universitaris i de Recerca* i a la *COST Action IC0604* sense el finançament de les quals no hagués estat possible realitzar les estances a l'estranger per obtenir la menció europea del doctorat.

Al *Nałęcz Institute of Biocybernetics and Biomedical Engineering Polish Academy of Sciences* per permètrem realitzar estances al seu centre d'investigació.

Personals:

Al Dr. Alfredo Bardají pel seu ajut, pel seu suport i per haver dirigit la meva tesi.

Al Dr. Marcial García-Rojo pel seu suport i pels seus savis consells en l'apassionant món de la Telepatologia i l'anàlisi digital d'imatges.

A la Dra. Anna Korzynska per la seva ajuda en la revisió d'aquesta tesi doctoral i per ajudar-me amb els seus coneixements en el món de l'anàlisi d'imatge.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Al Dr. Tomàs Alvaro per haver confiat en mi i donar-me l'oportunitat de començar a treballar a l'Hospital de Tortosa Verge de la Cinta.

Al Dr. Lluís Pons pel seu suport, pel seu ajut i per la seva insistència en aconseguir projectes d'investigació.

A Marc Iniesta, Vanesa Gestí, Bárbara Tomás, Ainhoa Montserrat, Maria del Mar Barberà, Ana Suñé, Verònica Echeverría, Montse Sebastià, Anna Carot i Rosa Cabrera perquè sense el seu ajut aquests treballs no haguessin estat possibles.

A Patrícia Escrivà pel seu recolzament i per compartir les penúries dels becaris.

A la Dra.Teresa Salvadó Usach pel seu suport, la seva confiança i per ajudar-me en totes les vicissituds estadístiques del món de la ciència.

Al Dr. Joaquín Jaén pel seu ajut, per haver confiat i apostat per mi, per donar-me l'oportunitat de desenvolupar-me com investigador i treballar en la seva línia d'investigació.

Al Dr. Ramon Bosch pel seu suport, per la seva ajuda i confiança, i per la seva contagiosa passió per la investigació i la ciència.

A la Dra. Marylène Lejeune per confiar en mi, pel seu suport, per ser la meva mentora i ensenyar-me tot el que cal per poder algun dia arribar a ser un científic.

A Sergi C., Nerea, Marcos, Cinta, Ismael, Lorena, Sergi P., Montse, Guillermo, Raquel, Gustavo, Conchi, Jesús, Nuria, René, Caterin, Juan Miguel, Rosa i Jorge per haver-me *ampeltat* de vatros.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

A J. Antonio Gil pel seu suport, confiança en mi i per la seva ajuda en la impressió d'aquesta tesi doctoral.

A Joan M. per ajudar-me a millorar com a persona i enriquir-me amb els seus consells.

A Joana pel seu recolzament i per haver donat al món les dues coses més meravelloses.

A Júlia pel seu suport, per aguantar les meves xapes, pel seu somriure i per fer que aquest món sigui un lloc millor on valgui la pena viure.

Als que estimo i ja no estan aquí.

Als meus pares i a la meva germana, pels seus esforços, pel seu suport i per haver-me acompanyat durant tot el trajecte de la meva vida.

A tu, Graciel·la, gràcies pel teu suport, per la teva ajuda, per donar-me valor, per estar sempre al meu costat i perquè sense tu no ho hauria aconseguit mai.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

INDEX.

Page/s

1. Endorsement letter by Prof. Anna Korzyńska.....5

2. Endorsement letter by Prof. Bernd Blobel.....7

3. Mentions.....9

4. List of publications 17

5. Abbreviations and Acronyms.....19

6. Abstract.....21

7. Resumen.....27

8. State of art in digital image analysis of immunohistochemical markers in pathology.....33

8.1. Technical variability.....33

8.2. Human variability.....34

8.3. Automated and semi-automated procedures in pathology.....35

8.4. Image format.....39

9. Objectives of the study.....41

10. Methods: Computer-assisted automated procedure.....45

10.1. Tissue preparation and immunohistochemistry.....45

10.2. Image acquisition.....46

10.3. Manual quantification.....47

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

10.4. Automatic nuclear quantification: segmentation and image analysis.48

10.4.1. Contrast enhancement……..48

10.4.2. Image segmentation……..49

10.4.3. Positive objects and drop data……50

10.5. Automatic nuclear quantification: data management……50

10.5.1. Algorithm 1 (image selection)……..51

10.5.2. Algorithm 2 (clusters with an area ≤25 μm^2)……51

10.5.3. Algorithm 3 (clusters with an area >25 μm^2)……51

10.5.4. Algorithm 4 (clusters with an area >25 μm^2)……52

10.6. Statistical analysis……52

11. Methods: Image format selection……55

11.1. Selection of histopathological material and immunohistochemistry……..55

11.2. Digitalisation procedures……..56

11.3. Image compression and quantification……56

11.4. Analysis of morphometric parameters altered by JPEG compression……..57

11.5. Methodology of comparison……58

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

12. Results……59

12.1. Automated quantification evaluation……59

12.1.1. Image classification……60

12.1.2. Automated methodology validation……60

12.2. Compression effects in the results of automated counting procedure…… 62

12.3. Compression effects in the study of morphological parameters……65

12.4. Correction of automated nuclei counting in compressed images..66

12.4.1. Determination of corrective factors……66

12.4.2 Quantification of nuclei with corrected roundness values in the Excel® macro……..……..……..……..68

13. Discussion……69

14. Conclusion……77

15. Reference list……79

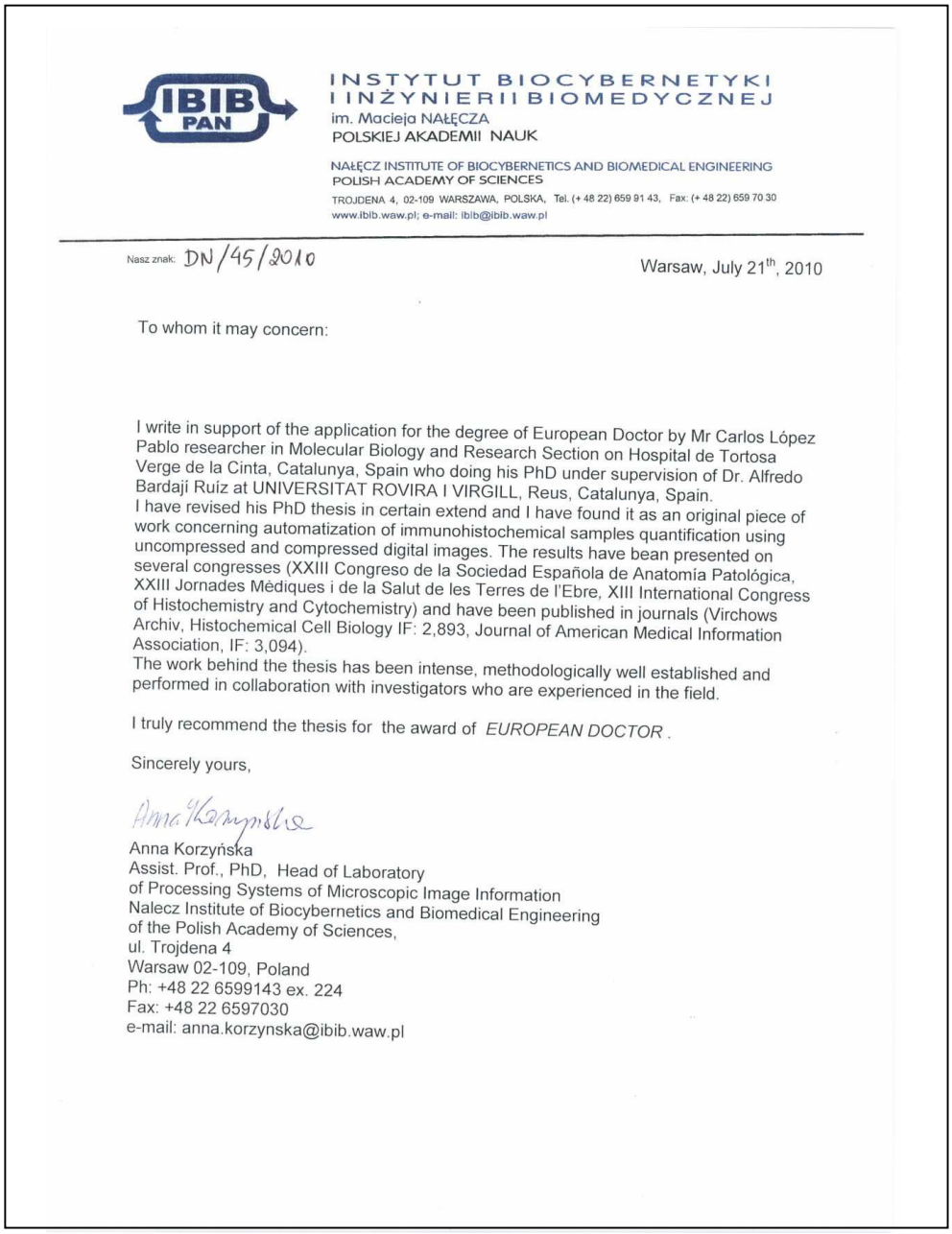
16. Tables (1 to 3)……..95

17. Figures (1 to 12)……..99

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

1. Endorsement letter by Prof. Anna Korzyńska.



INSTYTUT BIOCYBERNETYKI
I INŻYNIERII BIOMEDYCZNEJ
im. Macieja NAŁĘCZA
POLSKIEJ AKADEMII NAUK

NAŁĘCZ INSTITUTE OF BIOCYBERNETICS AND BIOMEDICAL ENGINEERING
POLISH ACADEMY OF SCIENCES
TROJDENA 4, 02-109 WARSZAWA, POLSKA, Tel. (+ 48 22) 659 91 43, Fax: (+ 48 22) 659 70 30
www.ibib.waw.pl; e-mail: ibib@ibib.waw.pl

Nasz znak: DN/45/2010

Warsaw, July 21th, 2010

To whom it may concern:

I write in support of the application for the degree of European Doctor by Mr Carlos López Pablo researcher in Molecular Biology and Research Section on Hospital de Tortosa Verge de la Cinta, Catalunya, Spain who doing his PhD under supervision of Dr. Alfredo Bardají Ruíz at UNIVERSITAT ROVIRA I VIRGILL, Reus, Catalunya, Spain.
I have revised his PhD thesis in certain extend and I have found it as an original piece of work concerning automatization of immunohistochemical samples quantification using uncompressed and compressed digital images. The results have bean presented on several congresses (XXIII Congreso de la Sociedad Española de Anatomía Patológica, XXIII Jornades Mèdiques i de la Salut de les Terres de l'Ebre, XIII International Congress of Histochemistry and Cytochemistry) and have been published in journals (Virchows Archiv, Histochemical Cell Biology IF: 2,893, Journal of American Medical Information Association, IF: 3,094).
The work behind the thesis has been intense, methodologically well established and performed in collaboration with investigators who are experienced in the field.

I truly recommend the thesis for the award of *EUROPEAN DOCTOR*.

Sincerely yours,

Anna Korzyńska

Anna Korzyńska
Assist. Prof., PhD, Head of Laboratory
of Processing Systems of Microscopic Image Information
Nalecz Institute of Biocybernetics and Biomedical Engineering
of the Polish Academy of Sciences,
ul. Trojdena 4
Warsaw 02-109, Poland
Ph: +48 22 6599143 ex. 224
Fax: +48 22 6597030
e-mail: anna.korzynska@ibib.waw.pl

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

2. Endorsement letter bv Prof. Bernd Blobel.

University of Regensburg Medical Center, eHealth Competence Center
Franz-Josef-Strauss-Allee 11, 93042 Regensburg, Germany

UNIVERSITY OF REGENSBURG MEDICAL CENTER
eHealth Competence Center
Head: Bernd Blobel, PhD, Associate Professor

Telefon: 0941/944-6769
Telefax: 0941/944-6766
E-mail: bernd.blobel@klinik.uni-regensburg.de
http://www.ehealth-cc.de

bl-la

Regensburg, 18. September 2010

To whom it concerns

I have carefully reviewed the Doctoral Thesis "Development of a new automated method for the quantification of nuclear immunohistochemical markers", authored by Mr. Carloz López Pablo, researcher in the field of Molecular Biology at the Research Section of the Hospital de Tortosa Verge de la Cinta, Catalunya, Spain, and submitted to the Universitat Rovira i Virgili, Reus, Catalunya, Spain.

As specialization of medical histology, the young discipline of immunohistochemistry serves the identification and classification of structures such as cancer cells. By that way, morphologically equally appearing cancers can be grouped according to their growing and distribution behavior as well as to their therapeutic response. In the context of targeted therapy, cell characteristics can be directly connected to the effectiveness of therapeutic molecules.. Therefore, the evaluation of immunohistochemical markers is an essential tool for diagnostic, predictive, therapeutic and research purposes.

The presented scientific work addresses the weakness of laborious and subjective manual quantification of immunohistochemically stained nuclear markers by automating this process, at the same time also tackling the problem of nuclear aggregates. As digital imaging produces an increasing number of image files of growing size, there is a need of – unfortunately lossy – compression of those images by preserving the contained information as good as possible.

The thesis investigates quality and reliability of automated quantification of nuclear immunohistochemical markers, the impact of image compression during the format transformation from TIFF to JPEG as well as the analysis of clusters, in comparison to the manual evaluation. The study has been performed scientifically and technically sound. This also applies to the statistical tests used to validate the results. The entire thesis in text, tables and figures has been presented in an understandable and consistent way. This does not exclude very few opportunities for improvements including some misspellings. The scientific work presented in the thesis has also been published in acknowledged international journals. Therefore, I recommend to accept the thesis for the award of a European Doctor.

Sincerely yours

Bernd Blobel

Bernd Blobel, PhD, Associate Professor

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

3. MENTIONS.

(1) This investigation was partly presented in the 21st European Congress of Pathology, Istanbul, Turkey, September 2007.

Joaquín Jaén, Carlos López, Marylène Lejeune, Patricia Escrivà, Maria Teresa Salvadó, Ramon Bosch, Lluís Pons, Jordi Baucells, Xavier Cugat, Jordi Roig, Tomás Álvaro. New approach for immunohistochemical nuclear markers quantification in digital images. Virchows Archiv 451(2), abstract PP4-295, pº585, 2007.

Carlos López, Joaquín Jaén, Marylène Lejeune, Patricia Escrivà, Maria Teresa Salvadó, Ramon Bosch, Lluís Pons, Jordi Baucells, Xavier Cugat, Jordi Roig, Tomás Álvaro. Effects of JPEG digital image compression on computer-assisted image quantification of nuclear immunohistochemical stained cells. Virchows Archiv 451(2), abstract PP4-296, pº585, 2007.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

PP4-294
COMPARATIVE STUDY BETWEEN SEMI-AUTOMATED AND MANUAL CELL QUANTIFICATION IN DIGITAL IMAGES IN PATHOLOGY

Marylène Lejeune[1], Joaquín Jaen[1], Carlos López[1], Patricia Escrivà[1], Maria Teresa Salvadó[1], Ramón Bosch[1], Lluís Pons[1], Jordi Baucells[2], Xavier Cugat[2], Jordi Roig[2], Tomás Alvaro[1]
[1] Verge de la Cinta Hospital, Pathology Department, Spain
[2] Verge de la Cinta Hospital, Informatics Department, Spain

Background: Manual quantification of immunohistochemical stained cells is the most frequent method in the current practice in pathology. To avoid the manual subjectivity, a number of automated and semi-automated processes have been previously developed and described. In the present study, we have developed different semi-automated processes in order to quantify different immunohistochemical markers. To obtain a gold-standard reference, we have performed manual quantifications and we have also evaluated inter and intra-observer variability. Method: 196 digital color images were obtained from sections of Hodgkin's lymphoma biopsies and stained with various nuclear, cytoplasmatic and membrane markers. Each image was manually quantified twice by 3 different observers. To analyse the images, we developed a specific macro for each marker with a commercial analysis software. The global dataset obtained with the software, were dropped to an Excel Datasheet, where the number of cells was obtained after we introduced different corrector factors. To evaluate the agreement degree between manual and semi-automated method and to quantify our intra and inter-observers variability, a comparative statistical analysis was performed with SPSS 11.0. Results: Globally, the variability was higher for intra-observers than for inter-observers counts, except in nuclear stains were both types of variations have similar differences. The comparison between semi-automated and manual methods shows that both methods have the same way of cell quantification. Both manual and automatic quantification showed a higher variation when the images contained more than 100 positive nuclei. Discussion: This study quantifies variation observed in manual quantification and shows that the different semi-automated procedures developed for each marker in our laboratory represent a valid alternative to manual quantification. (PI 04/1440, 04/1467, 05/1527).

PP4-295
NEW APPROACH FOR AUTOMATIC QUANTIFICATION OF IMMUNOHISTOCHEMICAL NUCLEAR MARKERS IN DIGITAL IMAGES

Joaquín Jaén[1], Carlos López[1], Marylène Lejeune[1], Patricia Escrivà[1], Maria Teresa Salvadó[1], Ramón Bosch[1], Lluís Pons[1], Jordi Baucells[2], Xavier Cugat[2], Jordi Roig[2], Tomás Alvaro[1]
[1] Verge de la Cinta Hospital, Pathology Department, Spain
[2] Verge de la Cinta Hospital, Informatics Department, Spain

Background: Several procedures of digital image analysis have been developed for quantitative evaluation of nuclear immunohistochemical markers with prognostic, diagnostic and therapeutic significance in current medical practice. To our knowledge, do not exist effective algorithms that permit to obtain a precise cell quantification in digital images with a high grade of complexity. The aim of this study is to develop a new methodology capable to analyze correctly, images with high variability in their morphology, cellular density, stain intensity and cell distribution. Method: 118 digital images from 4 different immunohistochemical nuclear markers were captured with different grade of cell concentration and clusters composition. Two coordinated macros were elaborated to perform the automatic count of positive nuclei. The first macro was developed with a commercial software that allows the modification and segmentation of the images. All extracted information was dropped to an Excel datasheet, where we have developed a macro with Visual Basic and introduced different algorithms that manage the dataset obtaining a final number of positive nuclei on each image. All statistical analysis was performed with SPSS 11.0. Results: t-Student test, Spearman correlation and ICC, showed no significant differences between the manual and the automatic count, whatever the image complexity. Kaplan-Meier and Bland Altman graphic representations indicate that cluster composition and a high nuclei density, increase variability on both types of count, and globally in more than 90% of images, automatic count was similar to the manual. Conclusion: This study, describes a new methodology for the automatic count of digital images for different immunohistochemical nuclear markers. This method improves the quantification of images with a high complexity in their nuclear composition. Variability observed between human and automatic count are the same as inter or intra-observer variability, which is accepted in the current clinical practice. (PI 04/1440, 04/1467, 05/1527).

PP4-296
EFFECTS OF DIGITAL IMAGE COMPRESSION ON COMPUTER-ASSISTED IMAGE QUANTIFICATION OF IMMUNOHISTOCHEMICAL STAINED CELL NUCLEI

Carlos López[1], Joaquín Jaén[1], Marylene Lejeune[1], Patricia Escrivà[1], Maria Teresa Salvadó[1], Ramón Bosch[1], Lluís Pons[1], Jordi Baucells[2], Xavier Cugat[2], Jordi Roig[2], Tomás Alvaro[1]
[1] Verge de la Cinta Hospital, Pathology Department, Spain
[2] Verge de la Cinta Hospital, Informatics Department, Spain

Background: The analysis of digital images and standard compression algorithms in current clinical practice has been used since some years ago. Image compression may reduce the amount of computer memory required for store images. The consequences of image compression have been previously evaluated on different diagnostic techniques. The aim of this study is to analyse the consequences of computer-assisted quantification, between uncompressed Tiff format and different levels of image compression in Jpeg format of immunohistochemical stained cells. Method: Digital images were captured with the software Leica IM50 4.0, in Tiff format from tissue samples stained with immunohistochemical markers, Ki67 (n=24) and FOXP3 (n=24). In a second step Tiff images were converted with the software ACDSee 6.0 to Jpeg files with a compression factor of 0, 50 and 100%. All the captured and converted images (n=196), were analysed with two coordinated macros developed with an image analysis software and an Excel datasheet. Quantified parameters were the total stained positive area and the number of positive nuclei. Tiff images results were compared with the different compression factor in Jpeg format, with Kaplan-Meier and Bland Altman plots using SPSS 11.5 Statistical Software. Results: Globally, variations in the quantification of positive stained areas between Tiff and the different compressed files are similar. On the other hand, cell count differences are lower with a compression factor of 0% and higher with a 50 or 100% of compression. Nevertheless, all compressed formats have a similar variability when images have less than 100 nuclei, and in this case, this variability is lower than globally. Conclusion: This descriptive study shows that image compression is a source of variation in automatic count of digital images. In these conditions, image compression should be omitted for immunohistochemical digital image quantification, overall in images with more than 100 nuclei. (PI 04/1440, 04/1467, 05/1527)

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

(2) This investigation was partly presented in the XXIII Congreso de la Sociedad Española de Anatomía Patológica, Tarragona, Spain, May 2007. C López, J Jaén, M Lejeune, MT Salvadó, P Escrivà, Ll Pons, T Álvaro. Nuevas estrategias de cuantificación inmunohistoquímica digital automatizada.

Sociedad Española de Anatomía Patológica
División española de la International Academy of Pathology

CERTIFICADO DE PRESENTACIÓN
XXIII CONGRESO DE LA SEAP

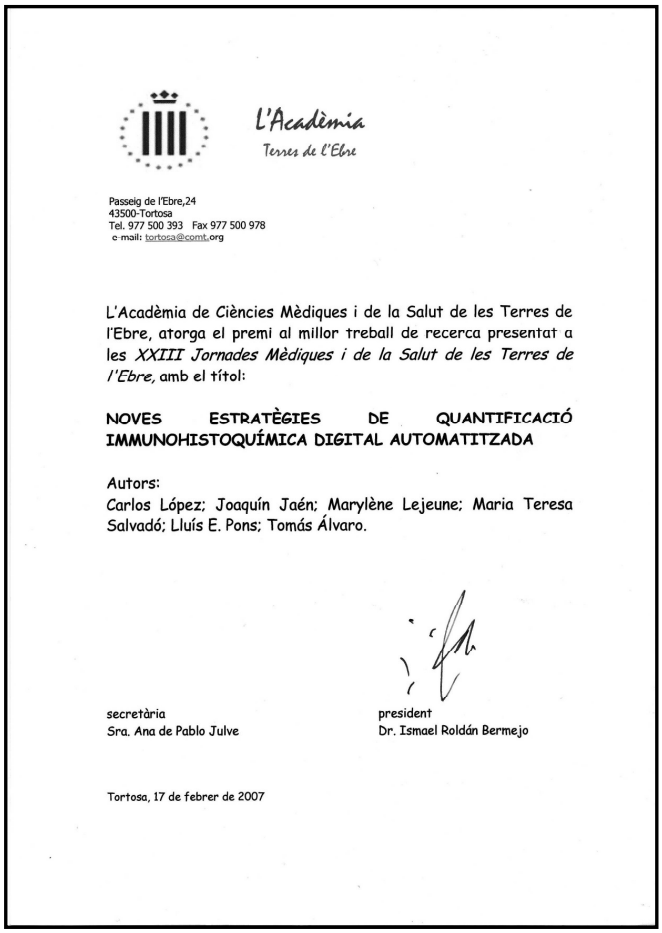
FERNANDO PINEDO MORALEDA, SECRETARIO DE LA SOCIEDAD ESPAÑOLA DE ANATOMÍA PATOLÓGICA Y DIVISIÓN ESPAÑOLA DE LA ACADEMIA INTERNACIONAL DE PATOLOGÍA.

CERTIFICA:

Que el Dr.; ***CARLOS LOPEZ PABLO*** , han presentado en el XXIII Congreso Nacional de la SOCIEDAD ESPAÑOLA DE ANATOMÍA PATOLÓGICA Y DIVISIÓN ESPAÑOLA DE LA ACADEMIA INTERNACIONAL DE PATOLOGÍA., celebrado en Tarragona, entre los días 16 al 19 de Mayo de 2007, la comunicación oral titulada: ***"NUEVAS ESTRATEGIAS DE CUANTIFICACIÓN IMNUNOHISTOQUÍMICA DIGITAL AUTOMATIZADA"***.

Para que conste donde proceda firmo el siguiente certificado en Tarragona a diecinueve de Mayo de dos mil siete.

(3) This investigation was partly presented in the XXIII Jornades Mèdiques i de la Salut de les Terres de l'Ebre, Tortosa, Spain, February, 2007. Lopez C, Jaén J, Lejeune M, Salvadó MT, Pons Ll, Álvaro T. Noves estratègies de quantificació immunohistoquímica digital automatitzada. **This present research was further awarded in that congress with the "X award of the Acadèmia de les Ciències i de la Salut de Catalunya i les Balears" as the best work in research presented.**



L'Acadèmia
Terres de l'Ebre

Passeig de l'Ebre,24
43500-Tortosa
Tel. 977 500 393 Fax 977 500 978
e-mail: tortosa@comit.org

L'Acadèmia de Ciències Mèdiques i de la Salut de les Terres de l'Ebre, atorga el premi al millor treball de recerca presentat a les *XXIII Jornades Mèdiques i de la Salut de les Terres de l'Ebre,* amb el títol:

NOVES ESTRATÈGIES DE QUANTIFICACIÓ IMMUNOHISTOQUÍMICA DIGITAL AUTOMATITZADA

Autors:
Carlos López; Joaquín Jaén; Marylène Lejeune; Maria Teresa Salvadó; Lluís E. Pons; Tomás Álvaro.

secretària
Sra. Ana de Pablo Julve

president
Dr. Ismael Roldán Bermejo

Tortosa, 17 de febrer de 2007

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

(4) This investigation was partly presented in the XIII International Congress of Histochemistry and Cytochemistry ICH2008 in Gdansk, Poland, August 2008.

Carlos López, Marylène Lejeune, Patricia Escrivà, Ramon Bosch, Salvado Teresa, Lluis Pons, Alvaro Tomas, Joaquin Jaen. Effects of image compression on area and roundness parameters in digital image analysis. Folia Histochemica et Cytobiologica, 2008:46(S2):S80 P2.63.

P2.63

Effects of image compression on area and roundness parameters in digital image analysis

Carlos Lopez, Marylene Lejeune, Patricia Escriva, Ramon Bosch, Salvado Teresa, Lluis Pons, Alvaro Tomas, Joaquin Jaen

Department of Pathology, Hospital de Tortosa Verge de la Cinta, Tortosa, Spain

Contact: clopezp.htvc.ics@gencat.cat

Introduction: Image compression and digital image analysis has been used in several fields of medicine. Two of the most used image formats are TIFF(uncompressed) and JPEG(compressed). In previous studies, we have observed that image compression could affect negatively the quantification of immunohistochemical stained nuclei. Automated nuclei counts on TIFF images differ from those made on the same images compressed in different JPEG formats. This variability increases with the level of compression and was particularly elevated in images with a high number of nuclei and with high area clusters. This study evaluates the variations in the area and roundness of positive stained nuclei due to image compression and assesses their effect on automatic nuclei quantification. In addition, a mathematical algorithm is implemented in the macros to decrease these differences between TIFF and JPEG. Material and methods: The present study was performed with 47 digital images of Ki67 and FOXP3 immunohistochemical markers captured and saved in the uncompressed TIFF format. Each one of these images was converted to JPEG files with 3x,23x and 46x compression using the ACDSee 9.0 program. Sixty-five positive objects with different shape and size were selected from TIFF images, and their area and roundness were compared with the same objects in the different compression levels. All images and positive objects were analyzed with macros previously elaborated, tested and validated with Image-Pro® Plus 5.0 program and Excel® datasheet complemented or not with a novel linear regression model. All statistical analyses were carried out using SPSS 11.0. Results: Boxplot graphics of 65 objects analyzed in TIFF and JPEG formats show practically similar area distribution whereas the roundness values decrease as the level of compression increases. Roundness variations could be corrected with a linear regression for each different level of compression calculated from these objects. When the linear regression is applied to all the positive objects analyzed, the original nuclei count differences between TIFF and the compressed images, were considerably reduced (Kaplan-Meier curves). In these conditions, the automatic nuclei counts of compressed images are more similar to the TIFF format, mainly in images with more than 100 nuclei or high area clusters. Discussion: This work shows how digital image compression could produce variations in distinct parameters (area and roundness) used to determine the positive object during an automated process of image analysis. Our new mathematical algorithms correct roundness variability measured in JPEG images, improving the accuracy of the macros developed to the automatic quantification of immunohistochemical nuclear markers. These results indicate that JPEG digital images could be an useful format for compact storage and quantitative immunohistochemical digital analysis.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

(5) This present research was awarded with the grant "Ajut predoctoral per la realització de la tesi doctoral curs 2009-10" supported by the Fundació Privada Reddis and the Universitat Rovira i Virgili.

Fundació Privada Reddis

AJUT PREDOCTORAL
per
LA REALITZACIÓ DE LA TESI DOCTORAL
CURS 2009-10
a
Carlos López Pablo

Reus, 28 de setembre de 2010

La Degana

Dra. Montserrat Giralt

UNIVERSITAT ROVIRA I VIRGILI
Facultat de Medicina
i Ciències de la Salut

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

(6) The present research was supported by:

5.1 Beca del Instituto de Salud Carlos III PI04/1440: "Desarrollo de patrones de análisis de imágenes digitales asistido por ordenador: Evaluación cuantitativa semiautomática de técnicas inmunohistoquímicas".

5.2 Beca del Instituto de Salud Carlos III PI08/0796: "Desarrollo de algoritmos informáticos para cuantificación inmunohistoquímica de dianas moleculares y marcadores pronósticos en cáncer de mama".

5.3 Beca de la COST-Action IC0604 EURO-TELEPATH "Telepathology Network in Europe" para estancia en el Institute of Biocybernetics and Biomedical Engineering en Varsovia (Polonia) del 5 de mayo al 13 de Junio de 2009

5.4 Beca de la Agència de Gestió d'Ajuts Universitaris i de Recerca de la Generalitat de Catalunya (Modalitat A: Beques i ajuts per a estades per activitats de recerca relacionades amb estudis de doctorat), para estancia en el Institute of Biocybernetics and Biomedical Engineering en Varsovia (Polonia) del 1 de Marzo al 30 de Abril del 2010.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

5. LIST OF PUBLICATIONS.

The present research was published in the following indexed and peer-reviewed journals:

- López C, Lejeune M, Salvadó MT, Escrivà P, Bosch R, Pons LE, Alvaro T, Roig J, Cugat X, Baucells J, Jaén J. Automated quantification of nuclear immunohistochemical markers with different complexity. Histochem Cell Biol. 2008 Mar;129(3):379-87. IF: 2,893. PMID: 18172664. Quartile: 1

- López C, Lejeune M, Escrivà P, Bosch R, Salvadó MT, Pons LE, Baucells J, Cugat X, Alvaro T, Jaén J. Effects of image compression on automatic count of immunohistochemically stained nuclei in digital images. J Am Med Inform Assoc. 2008 Nov-Dec;15(6):794-8. IF: 3,094. PMID: 18172664. Quartile: 1

- López C, Jaén Martinez J, Lejeune M, Escrivà P, Salvadó MT, Pons LE, Alvaro T, Baucells J, García-Rojo M, Cugat X, Bosch R. Roundness variation in JPEG images affects the automated process of nuclear immunohistochemical quantification: correction with a linear regression model Histochem Cell Biol. 2009 Oct;132(4):469-77. IF: 2,320. PMID: 19652993. Quartile: 2

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

4. ABREVIATIONS AND ACRONYMS.

DCT = Discrete Cosine Transform

DI = Digital Image

ER = Estrogen Receptor

FOXP3 = Forkhead Box P3

HSI = Hue Saturation Intensity

IC = Image Compression

ICC = Intraclass Correlation Coefficient

IHC = Immunohistochemical

JPEG = Joint Photographic Experts Group

PR = Progesterone Receptor

RGB = Red Green Blue

TIFF = Tagged Image File Format

WHO = World Health Organitzation

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

6. ABSTRACT.

Background: In the current practice of pathology, the evaluation of immunohistochemical (IHC) markers represents an essential tool for diagnostic, predictive, therapeutic and research purposes. Manual quantification of immunohistochemically stained nuclear markers is still laborious and subjective and the use of computerized systems for digital image (DI) analysis have not yet resolved the problems of nuclear aggregates (clusters). Furthermore, the volume of DI storage continues to be an important problem in computer-assisted pathology. In order to solve the storage problems, Joint Photographic Experts Group (JPEG) DI compression enables to reduce files sizes in lossy compression but some details may be lost.

Objectives: In this study, a new automatic procedure for quantifying immunohistochemical nuclear markers is proposed. Furthermore, the effects of DI compression on IHC nuclear markers counts have been observed. These effects have been studied using the proposed automatic computerized procedure. Moreover, this study attempts to show, with respect to immunohistochemically stained nuclei, which morphometric

shape parameters may be altered by the different levels of JPEG compression, and the implications of these alterations for automated nuclear counts. Finally, a method for correcting this discrepancy in the nuclear count is proposed.

Methods: The automated quantification procedure developed during the study consisted of two combined macros. The first one, based on a commercial software, analyzes DIs using a masking process, color segmentation and morphological descriptors of segmented objects. All information extracted in this first step is automatically exported to an Excel datasheet, where a second macro, composed of four different algorithms, analyzes all the information and calculates the total estimated number of positive nuclei on each image. 118 images of lymphomas and breast cancer with various levels of complexity were analyzed and compared with the manual quantification obtained by a trained observer. These samples were stained for the estrogens and progesterone receptors, Ki67 and FOXP3 IHC markers.

Furthermore, a group of 47 original images captured in Tagged Image File Format (TIFF) of lymphomas and breast cancer samples were used as

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

reference to study how image compression (IC) influences on the results of automatic quantification. The original images were converted to JPEG format with 3x, 23x and 46x compression rate. These original images were compared with these three groups of compressed images. Moreover, sixty-five positive objects were selected from these images and six morphological parameters (area, roundness, perimeter, diameter, length and width) were measured and compared in uncompressed and compressed images.

Results: Reliability (Intraclass correlation coefficient (ICC)>0.950) and differences between the manual quantification and the automated method (no statistically significant) were validated using statistical analysis. Bland-Altman plot and Kaplan-Meier curves indicated that there was an agreement with the results of both methods in almost 90% of the analyzed images.

To summarize, automated count differences between TIFF and JPEG images increased as the percentage of compression increased too. Compressed images with ≤100 cells/field, without clusters or with small-area clusters had small differences in the number of nuclei compared with

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

TIFF format (differences ≤5 nuclei/image in 95-100% of cases). By contrast, JPEG images with >100 cells/field or with large-area clusters showed substantial differences (< 35-50 nuclei/image in 95-100% of cases) at all levels of compression. Moreover, the roundness appears to be the only morphological parameter that was significantly affected by IC. Linear regression models were used to correct the variations in roundness produced by JPEG compression. These models were calculated for each compression rate and were incorporated in the automated macros. These correction factors reduced the nuclear counting differences arising from all IC levels, eliminating the statistically significant differences between compressed and uncompressed equivalent images.

Conclusions: The proposed automated procedure is an objective, faster than manual counting and reproducible method that has 90% of similarity with manual count, even with DIs with >100 cells/field or with large-area clusters. JPEG compression of DIs with ≤100 cells/field, without clusters or with small-area clusters does not compromise the accuracy of automatic quantification of IHC nuclear marker and could be an efficient method for storing these DIs, even when the correction factors are not used. Moreover, the results in this study demonstrate that with the roundness

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

correction factors, it is possible to carry out unbiased automated quantifications on IHC nuclear markers in compressed DIs with >100 cells/field or with large-area clusters with the proposed automated methodology. These correction factors could be easily incorporated in different systems of DIs analysis.

Keywords: Quantification, nuclei, image, analysis, immunohistochemistry, segmentation, algorithm, compression.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

7. RESUMEN.

Antecedentes: En la práctica habitual de la patología, la evaluación de marcadores inmunohistoquímicos representa una herramienta indispensable con fines diagnósticos, predictivos, terapéuticos y en investigación. La cuantificación manual de marcadores nucleares teñidos inmunohistoquímicamente todavía es laboriosa y subjetiva y la utilización de sistemas informatizados para el análisis de imágenes digitales, no ha resuelto todavía el problema de los agregados nucleares (*clusters*). Por otro lado, el volumen de las imágenes digitales almacenadas continúa siendo un problema importante en la patología asistida por ordenador. La compresión JPEG (Joint Photographic Experts Group) de imágenes digitales permite reducir el tamaño de los archivos pero con la desventaja que se pueden perder algunos detalles.

Objetivos: En el presente estudio, hemos diseñado un nuevo procedimiento automático con la finalidad de cuantificar varios marcadores nucleares inmunohistoquímicos con (*clusters)* de diferentes grados de complejidad.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Por otro lado, los efectos de la compresión de imágenes digitales sobre los recuentos de marcadores nucleares inmunohistoquímicos fueron estudiados según los resultados obtenidos con el procedimiento automático.

Al mismo tiempo, este estudio intenta demostrar qué parámetros morfológicos pueden alterarse debido a los diferentes niveles de la compresión JPEG; las implicaciones que estas alteraciones tienen en la cuantificación nuclear, así como desarrollar un método para corregir estas alteraciones en el recuento nuclear.

Métodos: El procedimiento para la cuantificación automatizada desarrollado durante el estudio, consiste en dos macros combinadas. La primera, desarrollada con un programa comercial, permite el análisis de imágenes digitales usando una segmentación por color y morfológica que incluye una máscara. Toda la información extraída con esta primera macro es automáticamente volcada a una hoja de Excel, dónde una segunda macro compuesta de cuatro algoritmos diferentes, analiza toda la información y calcula el número definitivo de núcleos positivos según el nivel de complejidad de cada imagen. 118 imágenes de linfomas y cáncer

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

de mama con *clusters* de diferentes niveles de complejidad fueron analizadas y comparadas mediante la cuantificación manual realizada por un observador experimentado. Las muestras utilizadas fueron teñidas para los marcadores inmunohistoquímicos de receptores de estrógenos, receptores de progesterona, Ki67 y FOXP3.

Por otro lado, un grupo de 47 imágenes capturadas en formato TIFF (Tagged Image File Format) en muestras de linfomas y cáncer de mama teñidas con Ki67 y FOXP3 fueron utilizadas como referencia y comparadas con tres grupos que contenían las mismas imágenes pero convertidas a formato JPEG con un nivel de compresión 3x, 23x y 46x, con el objetivo de estudiar la compresión de imágenes en la metodología automatizada. Además, se seleccionaron 65 objetos de estas imágenes, se midieron seis parámetros morfológicos y se compararon entre el formato TIFF y los diferentes grados de compresión.

Resultados: En relación con la validación de la cuantificación de los núcleos, el análisis estadístico demostró una gran concordancia entre la cuantificación manual y el método automatizado (ICC>0.950) y sin diferencias estadísticamente significativas. Las gráficas de Bland-Altman y

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

las curvas de Kaplan-Meier mostraron que los resultados obtenidos con ambas metodologías concuerdan en un 90% de las imágenes analizadas. Globalmente, en relación con los efectos de la compresión de imágenes, las diferencias en los recuentos automatizados entre las imágenes TIFF y JPEG aumentaron con el porcentaje de compresión. Las imágenes comprimidas con baja complejidad (≤100 células / campo, sin *clusters* o con *clusters* de área pequeña) tuvieron pocas diferencias comparadas con el formato TIFF (≤5 núcleos / imagen en el 95-100% de los casos). Por el contrario, los recuentos automatizados en las imágenes JPEG con elevada complejidad (>100 células o *clusters* de área grande) mostraron diferencias importantes en todos los niveles de compresión comparados con los realizados en el formato TIFF (diferencias <35-50 núcleos / imagen en el 95-100% de los casos). Además, la redondez fue el único parámetro morfológico que estuvo significativamente afectado por la compresión. Los factores para corregir las discrepancias en la estimación de la redondez se obtuvieron de modelos de regresión lineal para cada nivel de compresión, eliminándose así las diferencias estadísticamente significativas entre las medidas de la redondez en las imágenes equivalentes. Estos factores correctores fueron incorporados en las macros automatizadas, los cuales

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

redujeron las diferencias en los recuentos nucleares producidos por la compresión de imágenes.

Conclusiones: Este nuevo procedimiento automatizado es un método objetivo, más rápido y reproducible que tiene un excelente nivel de precisión, incluso con imágenes digitales de elevada complejidad.

La compresión JPEG en imágenes digitales de baja complejidad no compromete la eficiencia de la cuantificación automatizada de los marcadores nucleares inmunohistoquímicos y puede ser un método eficiente para almacenar imágenes digitales, incluso cuando no se utilizan los factores correctores desarrollados. Además, nuestros resultados demuestran que utilizando nuestra metodología automatizada conjuntamente con los factores correctores de la redondez permiten realizar recuentos automatizados sin desviación en imágenes comprimidas de elevada complejidad. Estos factores correctores pueden ser fácilmente incorporados en diferentes sistemas de análisis de imagen.

Palabras clave: Cuantificación, núcleos, imagen, análisis, inmunohistoquímica, segmentación, algoritmo, compresión.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

8. STATE OF ART IN DIGITAL IMAGE ANALYSIS OF IMMUNOHISTOCHEMICAL MARKERS IN PATHOLOGY.

8.1. Technical variability.

In the current practice of pathology, the identification of cell markers and the evaluation and quantification of positively immunostained cells represent an indispensable routine for diagnostic, predictive, therapeutic and research purposes. At present, a substantial number of these markers are detected in situ, in tissue sections, with IHC techniques and they are evaluated directly with the microscope (manual evaluation). The principal disadvantages associated with these classical manual methods of sample evaluation are the labor-intensiveness of quantification (time consuming) and the variability obtained from the cell counts when several evaluations are performed in the same sample. This variability arises from the heterogeneity of stain intensities, variation in cell size, shape, and distribution and the variability in appearance of the immunoreactive objects. This is basically due to a number of variables that influence antigen staining in paraffin-embedded tissues. The principal extrinsic factors of variability are related to the specimen and include, principally, the clone, the dilution of the antibody, the detection system and chromogen, the

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

antigen retrieval method, and the external/internal reference standards.[1] All these extrinsic factors can be modified and controlled as part of the technique, unlike the intrinsic factors related to the tissue sample, which are almost difficult to standardize between laboratories, or even within the same laboratory.[2] These intrinsic factors include the type of fixative, fixation time, tissue processing and the level of antigen expression and preservation. Changes in these variables can produce variability in the results of manual counting done by one observer (intra-observer variability) and between different observers (inter-observer variability).

8.2. Human variability.

Evaluation and quantification methods range from simple counting of stained cells to more complex grading of their intensity on an arbitrary scale performed by the pathologist using a conventional microscope. Under these conditions, the specific and precise quantification of stained cells is subject to evaluation by the human eye, which has difficulties in distinguishing subtle differences in staining intensity, especially at the extremes of the continuous color scale.

Nevertheless, the necessary time to quantify large samples is very high and the interpretation of immunostains is difficult because of the technical variability and the human subjectivity in samples evaluations. So, the reproducibility among observers is not always achived due to the interpretation of immunostains.[2-5] Sometimes, subjectivity makes it difficult to establish criteria for cell positivity, and has given rise to different scoring definitions.[1,4-8] Additionally, variations in manual counting corresponded to intra- or inter-observer variability. While some authors have demonstrated that inter-observer differences are greater than intra-observer differences,[9-11] others have published results indicating the opposite relationship or have observed no significant differences.[12,13] Previous studies carried out in our laboratory found no significant differences between inter- and intra-observer results for nuclear marker quantifications.[14]

8.3. Automated and semi-automated procedures in pathology.

The lack of reproducibility of some of the human evaluations has promoted the use of automated image analysis technologies as a means of avoiding subjectivity. Computer-assisted image analysis is not only easy to use and economical, but it can also be an effective method of quantifying and scoring immunohistochemically stained markers. Several automated image

analysis programs (Image-pro® Plus, ACIS®, Metamorph®, Visilog®) suitable for use in histopathology have recently been commercialized. Used in conjunction with a microscope and a digital camera, these programs can detect, quantify, and classify immunostained cells in DIs based on color, size, and shape of these cells.[15] These systems of processing images are based on a method of segmentation that divides images into two classes of areas: the first ones are the areas of interest and the second ones are the background. This classification is done in order to determine which pixel of the image belongs to each class.

Various semi-automated or automated processes have been developed to identify individual stained cells in different tissue samples and to classify them using cellular shape descriptors. These methods could improve levels of sensitivity, precision, reproducibility and standardization of these kinds of measurements, since sometimes semi-automatic methods require the user interaction for object boundaries modification. However, the selection of multiple parameters that can yield a correct score is not an easy task. In addressing this problem, several authors have developed a number of analytical systems used to perform semi-automated or automated image analysis of DIs of cells.[6,16-25] Many of these studies have focused their

attention on nucleus quantification in simple images. Nevertheless, they have not solved the problems of analyzing images of high nuclear density and clusters. Although the development of specific algorithms for quantifying images has yielded satisfactory results, whereby 90-95% of nuclei may be correctly quantified, these studies present results on images clearly separated or simple clusters with adjacent and non-overlapping nuclei.[26-30] A few studies have applied segmentation techniques in images of more complex clusters but they obtained a relatively low rate of successful quantification.[31-33] However, markers used in anatomical pathologic diagnosis are not limited to the nuclear markers alone but they can also stain membrane or be disseminated throughout the cytoplasm in granules. At present, there are not reliable fully automated methods which had already been described and that could be applicable to these subcellular locations of proteins.[2,14,34]

There are several studies which have reported the great value for the diagnosis and treatment when using automated or semi-automated methods to detect and quantify cells or nuclei in pathological tissues in lymphomas,[14,35-38] breast cancer,[17,39-42] prostate[6,19] and other types of neoplasm.[29,43-45] Other studies have revealed the importance of measuring

the morphological parameters of cells or nuclei using computer-assisted image analysis. Various morphological characteristics, like area, roundness, length, width and optical density, amongst others, have been studied in order to describe their biological, histological and clinical significance in malignant neoplasms of the bladder,[46] prostate,[47] skin[48] and face.[49] These types of studies analyze the differences between malignant and benign cells with parameters of Euclidian geometry. Otherwise, some aspects of cells morphology are better described in terms of fractal geometry.[50-52] These previously mentioned studies use the fractal dimension and the lacunarity as descriptor factors of differences between malignant and benign cells. In other studies, a reliable classification of immunopositive patterns has also been found when the feature extraction methods were based only on fractal dimension analysis.[53] Fractal geometry has been described to be specially sensitive in measuring dimensions of complex irregular objects compared to the Euclidean geometry.[54,55] Moreover, several studies in pathology have revealed that fractals could help to the characterization of malignant cells.[50-52] However, depending on the type of measurements and segmentation procedures, fractal or Eculidean geometry should be chosen to perform the best image analysis procedure.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

8.4. Image format.

The first step needed to count automatically the number of immunopositive nuclei is the acquisition of optimal digital-microscopy images. The quality of DI depends on various technical items, mostly from microscope optics, digital cameras, computer monitors and more. Digital cameras allow to capture images in different formats. TIFF and JPEG are two of the most frequently used image formats in medicine such as radiology, microscopics and others.[19,37,56-61] TIFF is an example of a lossless format, which has the advantage that no data are lost, but with the disadvantage that files sizes are quite large. JPEG is an example of lossy format based on combination of a discrete cosine transform (DCT) and Huffman encoding, which has the advantage that it can reduce file sizes but with the disadvantage that some detail may be lost in the compression.

Acquired images in TIFF format could be transformed in a JPEG format with different degree of compression. At low and moderate compression levels, it is very difficult for the human eye to discern any difference from the original, even at extreme magnification. However, JPEG inevitably introduces digital artefacts. Changes produced during JPEG IC principally affect the values of pixels in each of the three Red, Green and Blue (RGB)

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

channels in the 24-bit images.[57] This could alter results in some positive pixels being altered so that they fall outside the range of positive selected colour values, and in some negative pixels appearing to fall within the positive range of values.

Nevertheless, in most published IHC studies, the images have been captured with the maximum quality in an uncompressed TIFF format. However, very large DI databases require a huge digital storage space, so storing DIs remains one of the biggest problems.[2] At present, the effects of compression in Feulgen-stained samples have already been studied[57] but we could not find other works evaluating the consequence of this kind of compression in the automatic quantification of DIs with ICH markers stained with DAB(3,3'Diaminobenzidine).

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

9. OBJECTIVES OF THE STUDY.

In recent years, the evaluation of samples in pathology has become more precise due to the tumoral classifications developed for the World Health Organization (WHO) because different types of treatment are administrated depending of the pathologist's evaluation of the samples.

These protocols allow specialists to reduce inter- and intra-observer variability but not to eliminate them. A previous study done by our group, but not included in this thesis, has estimated inter and intra-observer human variability in cytoplasmatic, membrane and nuclear IHC markers.[14] The methodologies developed in this study were semi-automated and were applied to samples stained with nuclear markers that had a low number of positive nuclei. The aim described in the present thesis was to develop a totally automated computer-assisted methodology to quantify nuclear markers in images with several degrees in clustering complexity, in order to avoid human subjectivity. The accuracy and reliability of this process was compared with those of the manual method. The new proposed method was dedicated to segment and quantify immunohistochemically stained

nuclei from DIs involving four different nuclear markers from various tumoral tissue types, including images with nuclear clustering.

To achieve a more efficient way to store IHC DIs, this study also attempted to determine the influence of different JPEG IC rates on the automatic quantification. Moreover, the influence of image complexity in JPEG compression was also studied. Finally, the rates of JPEG compression that did not compromise the accuracy of the automatic nucleus count of IHC markers for diagnostic or research purposes were determined.

The image analysis procedure developed in the present study is based on color, morphological features of nuclei and thresholding. These types of segmentation have been used in several studies with various types of softwares.[19,62-64] The most important features in this type of IHC nuclear analysis are intensity level of positive stain (DAB) according to the values of each pixel in the RGB channel, the size (area, diameter, perimeter, length, width) and shape (roundness) of the positive objects. The effect of the JPEG DI compression on these parameters has not been evaluated in the particular case of the quantitative analysis of IHC nuclear markers. This study identifies which morphometric parameters may be altered by the

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

different levels of JPEG compression, to determine the implications of these alterations for the total counting of nuclei. The present study also attempts to develop corrective factors that might yield very similar nuclei counts to those of the original uncompressed images.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

10. METHODS: COMPUTER-ASSISTED AUTOMATED PROCEDURE

López C, Lejeune M, Salvadó MT, Escrivà P, Bosch R, Pons LE, Alvaro T, Roig J, Cugat X, Baucells J, Jaén J. Automated quantification of nuclear immunohistochemical markers with different complexity. Histochem Cell Biol. 2008 Mar;129(3):379-87.

10.1. Tissue preparation and immunohistochemistry.

The present study has been carried out on samples of different kinds of tumoral tissues collected from the archives of the Department of Pathology of the Hospital de Tortosa Verge de la Cinta. Tissues were fixed, processed and sections of paraffin-embedded tissues were prepared as previously described.[65] Briefly, 3-µm-thick sections were dried, deparaffinized in xylene, rehydrated in graded ethanol, and washed in water and PBS. Antigen retrieval was achieved by heat treatment in a pressure cooker or pronase digestion, as necessary. Samples were incubated with the appropriate dilution of the primary antibodies: Ki67 (clone MIB-1, Dako, Carpinteria, CA), FOXP3 (clone FOXP3-236A/E7, CNIO), Estrogen Receptors (ER) (clone NCL-ER-6F11, Novocastra, Newcastle upon Tyne, UK), and Progesterone Receptors (PR) (NCL-PGR-312, Novocastra). Automatic immunostaining was performed with a Horizon

TechMate (Dako). The secondary antibody was conjugated to the reagents in the DAKO EnVision™ system (Dako EnVision+™, Dako Corporation), which include peroxidase block, labeled polymer, and buffered substrate/DAB$^+$ chromogen. Upon oxidation, DAB forms a brown end-product at the site of the target antigen. Finally, tissues were counterstained with hematoxylin, dehydrated, and mounted. The entire process has been standardized according to the manufacturer's instructions to ensure high reproducibility and stain homogeneity, which are very important requirements for image analysis.[4,8,66,67]

10.2. Image acquisition.

Tissue sections were viewed using brightfield illumination under a Leica DM LB2 upright light microscope (Leica Microsystems Wetzlar GmbH, Wetzlar, Germany) with a 40x plane-apochromatic objective with a numerical aperture of 0.63. Representative areas of the different sections with different degrees of cellular complexity were captured on a Leica DFC320 digital camera (Leica Microsystems Digital Imaging, Cambridge, UK) controlled by Leica DFC Twain software 6.3.0 (Leica Microsystems Digital Imaging) running on a Compaq Professional Workstation computer (3.6Ghz Pentium IV CPU, 1.00GB RAM). These images have a resolution

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

of 2088x1550 pixels with RGB 24 true color format, and were saved in uncompressed TIFF format. Homogeneous light intensity was manually adjusted to allow maximum reproducibility since variations in the illumination can produce significant differences in the results of image quantification.

10.3. Manual quantification.

Each image was manually evaluated by a trained observer to obtain a reference value (Gold Standard), which was compared with automatic results. The observer quantified the different DIs with the help of the Manual Tag tool of the Image-Pro® Plus 5.0 program (Media Cybernetic, Silver Spring, USA), which allows to put a mark on each nucleus evaluated as positive. The tally of positive nuclei was increased automatically in the Manual Tag view menu in the program's object counter. The images were quantified twice by the same observer with one month's interval between counts. The mean value of these two counts was used as a reference for evaluating the accuracy of the automated nuclear quantification.

10.4. Automatic nuclear quantification: segmentation and image analysis.

The automatic nuclear quantification procedure presented here in this study has two separate parts: the analysis of the DIs and the post-processing of the information obtained in the analysis (Figure 1). In the first part of the procedure, images were analyzed with the Image-Pro® Plus 5.0 program which features a range of image processing tools. This software supports a macro language for specific procedures that allows the sequential reproduction and automation of all steps. The macro was made up of a set of "actions" or algorithms that allows the non-stop analysis of up to 50 consecutive images taken from the different IHC markers used in this study (Ki67, FOXP3, ER, PR).

10.4.1. Contrast enhancement.

The image was loaded onto the screen of the image analysis program and spatially calibrated using the calibration tool of Image-Pro® Plus 5.0, which includes the corresponding calculated distance for the 40x objective. Then, the Contrast Enhancement tool is applied using predetermined values that can be found in the text macro presented in the article that describes the macro explained in this thesis.[35] The contrast enhancement made it easy to

choose the colour of the positive objects and minimizes variations in staining intensity, increasing the efficiency of the detection algorithms.

10.4.2. Image segmentation.

The second step of this first macro allows simultaneous color segmentation and objects selection according to morphological features.[68] Several authors use the Hue, Saturation and Intensity (HSI) color model for color segmentation, since it is intended to correspond more closely to the way humans perceive color.[63,69,70] The proposed method uses the RGB color model because there was observed less overlapping of the pixel values of negative nuclei. Some overlapping values of colors were eliminated using a special mask,[35] which modifies the pixel color values, displacing color of background and negative nuclei outside the segmentation color range of positive nuclei.

Positive objects obtained by the segmentation process corresponded to nuclear fragments, individual nuclei, and nuclear clusters. All segmented positive objects with an area <0.8 μm^2 were discarded because they were considered to be background or detritus. All positive objects ≥0.8 μm^2 were examined according to their size and shape. Eight different ranges were

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

established empirically to extract the maximum information. For each of these ranges, special algorithms of number of nuclei estimation were used in the second macro.

10.4.3. Positive objects and drop data.

The last step included in this first macro consisted of exporting all the extracted information about selected positive objects to an Excel datasheet, in order to use it in the macro to determine the exact number of positive nuclei (individual and those included in clusters).

10.5. Automatic nuclear quantification: data management.

The macro applied in the second part of the quantification procedure uses all the information extracted in the first macro to estimate the number of nuclei. The second macro, written in Visual Basic, was created in an Excel datasheet as an automated procedure. The four algorithms incorporated in this macro, which will be later described, took into account the nuclear architecture and the morphological features as shown in Figure 1.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

10.5.1. Algorithm 1 (image selection).

This algorithm selects images into two categories according to the presence or the absence of clusters. In the first category, there are images without nuclear clusters or with clusters of small area (≤25 μm^2). The second category includes images with clusters >25 μm^2.

10.5.2. Algorithm 2 (clusters with an area ≤25 μm^2).

This algorithm was applied to the images of the categories with no clusters or low-area clusters (≤25 μm^2). This algorithm incorporated a mean area for each image calculated from their own positive nuclei. This mean area was applied to both the individual nuclei and to the clusters of small area, which permits us to obtain the final number of nuclei in each image. This last step allows us to extrapolate the number of nuclei inside the clusters. This extrapolation is based in the assumption that all the objects inside the clusters have the same area, although some of them can have an area higher or lower than the mean area.

10.5.3. Algorithm 3 (clusters with an area >25 μm^2).

This algorithm was applied to the images of the categories with >100 nuclei or high-area clusters. It uses the standard deviation of nuclei areas and the

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

sum of areas of all positive objects to redirect each image towards the different variant of algorithm 4.

10.5.4. Algorithm 4 (clusters with an area >25 μm^2).

This final algorithm was based on the area and roundness of the positive objects, and had seven different ways of data analysis. Each way is chosen based on the mean area of the positive objects and its standard deviation in each image. The mean area of the nuclei in each case is calculated using different ranges values of the area and the roundness. The mean area calculated is used to estimate the number of nuclei inside of the clusters. This estimation is based in an assumption similar to the one used in the second algorithm.

10.6. Statistical analyses.

All statistical analyses were done using SPSS v. 11.0. The extent of agreement between the manual and automatic results was evaluated with Student's t-test, Spearman's correlation coefficient, the ICC, and the Bland-Altman and Kaplan-Meier analyses with their corresponding graphic representation. Student's t-test was applied to evaluate differences

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

between the results of manual quantification and the automated proposed counting method. Spearman's correlation coefficients were computed to check the associations between the manual and automated quantifications. Bland-Altman graphs illustrate the difference between results obtained with the manual and automated methods against the mean of both counts. Kaplan-Meier curves portray the conditional probability of observing differences between results obtained with the manual and automated methods.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

11. METHODS: IMAGE FORMAT SELECTION.

López C, Lejeune M, Escrivà P, Bosch R, Salvadó MT, Pons LE, Baucells J, Cugat X, Alvaro T, Jaén J. Effects of image compression on automatic count of immunohistochemically stained nuclei in digital images. J Am Med Inform Assoc. 2008 Nov-Dec;15(6):794-8.

López C, Jaén Martinez J, Lejeune M, Escrivà P, Salvadó MT, Pons LE, Alvaro T, Baucells J, García-Rojo M, Cugat X, Bosch R. Roundness variation in JPEG images affects the automated process of nuclear immunohistochemical quantification: correction with a linear regression model Histochem Cell Biol. 2009 Oct;132(4):469-77.

11.1. Selection of histopathological material and immunohistochemistry.

Representative slides immunostained with monoclonal antibodies directed against the nuclear protein Ki67 (clone MIB-1, Dako, Carpinteria, CA, USA) and FOXP3 (clone FOXP3-236A/E7, CNIO, Spain) were selected from the archives of the Department of Pathology of the Hospital de Tortosa Verge de la Cinta. All slides were prepared, processed and immunohistochemically stained in the Pathology Department using previously described standard protocols.[35,65]

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

11.2. Digitalisation procedures.

Samples were viewed with a Leica DM LB2 microscope (Leica Microsystems Wetzlar GmbH, Wetzlar, Germany) with x40 magnification with a 0.65 numeric aperture with similar illumination. Forty-seven DIs of Ki67 and FOXP3 IHC markers were captured and saved in the uncompressed TIFF format with a Leica DFC320 digital camera at a resolution of 3.3 Mpixels with Bayer Array RGB filter. To avoid differences in the illumination that might produce significant differences in the quantification of the image, adjusted values of illumination for the images were used to ensure maximum reproducibility. The camera was connected to a Compaq Professional Workstation computer (2 GHz Pentium IV CPU, 750 MB RAM) and controlled with the Leica IM50 v4.0 program.

11.3. Image compression and quantification.

ACDSee software allows us to compress images with different ratios of JPEG compression. Three compression ratios were chosen to compress the original TIFF images (around 9.27 Mbytes). The first compression ratio with the lowest compression and the highest quality that the software allows is 3x compression or a ratio compression 1:3 (around 3.25 Mbytes). The second compression ratio with intermediate quality images is 23x

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

compression or a ratio compression 1:23 (through 360-460 400 Kbytes). Finally, images with maximum compression and poor quality are 46x compression or a ratio compression 1:46 (through 180-246 Kbytes).

The set of 47 original uncompressed TIFF DIs and their JPEG versions (47 with 3x, 47 with 23x and 47 with 46x JPEG compression) were analysed using the automatic computer-assisted procedure presented in this thesis.

11.4. Analysis of morphometric parameters altered by JPEG compression.

The positive objects from TIFF and JPEG images were detected and quantified with the previously described method. Then, 65 positive objects (individuals and clusters) with different shape and size were chosen from TIFF images. The morphometric parameters of shape of these selected objects for this study were area (μm^2); classical perimeter (perimeter1: length of the object's outline in µm), diameter (average of diameters measured at 2 degree intervals and passing through object's centroid in µm), length (Feret diameter along the major axis of the object in µm), width (Feret diameter along the minor axis of the object in µm) and the roundness which include the perimeter2 (the chain code length of the outline:

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

[perimeter2^2] / [4 x Π x area]). These parameter values were calculated and compared with those coming from various versions of JPEG compression.

11.5. Methodology of comparison.

Computer-assisted counts and morphological parameters of objects shapes of TIFF DIs were the reference in statistical comparisons. Differences in the nuclear count obtained between paired results (TIFF versus each group of compressed JPEG images with 3x, 23x and 46x compression) were evaluated by the paired t-test, the ICC as well as the Bland-Altman and Kaplan-Meier analysis with their corresponding graphical illustration. Student's t-test was used to determine the significance of the differences between the mean values of each morphological feature of corresponding objects in TIFF format and each compressed JPEG formats. All statistical analyses were carried out using SPSS version 11.0.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

12. RESULTS.

López C, Lejeune M, Salvadó MT, Escrivà P, Bosch R, Pons LE, Alvaro T, Roig J, Cugat X, Baucells J, Jaén J. Automated quantification of nuclear immunohistochemical markers with different complexity. Histochem Cell Biol. 2008 Mar;129(3):379-87.

López C, Lejeune M, Escrivà P, Bosch R, Salvadó MT, Pons LE, Baucells J, Cugat X, Alvaro T, Jaén J. Effects of image compression on automatic count of immunohistochemically stained nuclei in digital images. J Am Med Inform Assoc. 2008 Nov-Dec;15(6):794-8.

López C, Jaén Martinez J, Lejeune M, Escrivà P, Salvadó MT, Pons LE, Alvaro T, Baucells J, García-Rojo M, Cugat X, Bosch R. Roundness variation in JPEG images affects the automated process of nuclear immunohistochemical quantification: correction with a linear regression model Histochem Cell Biol. 2009 Oct;132(4):469-77.

12.1. Automated quantification evaluation.

Time needed to count the number of immunostained nuclei in a single image (taken with the 40x objective and with an area of $3.2x10^6$ μm^2) with the automated procedure takes approximately 20 seconds. The time needed for manual counting for the same images is longer and depends on

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

the image complexity and on the number of nuclei. According to the experiments, the mean number of stained nuclei is 160 nuclei per image. The manual counting has a time cost that ranges from 180 to 200 seconds. So, using the automated version of the nuclei counting there is a higher saving time.

12.1.1. Image classification.

For statistical analysis, images were classified into four different groups. Images were grouped on the basis of size and complexity or on the basis of the number of positive nuclei. The first group includes 51 images with individual nuclei or very low area clusters (≤25 μm^2). The second group includes 67 images with a cluster area >25 μm^2. The third group includes 51 images with ≤100 nuclei and the fourth group of 67 images >100 nuclei. Figure 2 shows images of different nuclear classification in which we can observe high area clusters.

12.1.2. Automated methodology validation.

Statistics of the comparison between the manual and automated methods are shown in Table 1. The mean and standard deviation of the two manual counts were taken as references for comparison with the results of the

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

automatic process. Table 1 shows that there were no significant differences between the mean of the automatic and the manual counts (t-test), p>0.050). Spearman's correlation coefficient was generally >0.950, and the ICC, derived from the analysis of variance, agreed very closely, with values generally >0.950. The large standard deviations of the mean of the manual and automated quantifications are due to the wide range of the number of positive nuclei in several image counts. The high correlations between the results obtained with the two methods for the various image categories are shown in Figure 3.

However, plotting the differences between the paired measurements against the mean of the two values (Bland-Altman's graphical representation) shows that these differences are not homogeneous (Figure 4A). A clear difference can be observed when images contained >100 positive nuclei (Figure 4C) or were composed of clusters with a great area (Figure 4E). Nevertheless, the difference in manual counting versus the automatic counting is in the range of intra-observer variability in manual counting. The conditional probabilities of observing differences between paired quantifications obtained with manual and automatic procedures were estimated by the Kaplan-Meier method (Figure 5). Globally, the range

of differences obtained from intra-observer differences and manual versus automatic probabilities of differences were similar in more than 90% of the images (Figure 5A). The range of differences observed in images with a few positive nuclei (Figure 5B) or with small clusters (Figure 5D) were almost the same in 100% of the images. In images with >100 nuclei (Figure 5C) or with large clusters (Figure 5E), the range of differences was almost equal between the two curves for 86% of images.

12.2. Compression effects in the results of automated counting procedure.

The statistical analysis of the counting objects differences is presented in the Figure 6. The agreement between the counts in TIFF image and in the different levels of JPEG compressed images demonstrated that the differences were not homogeneous. The automatic quantifications performed in JPEG images, in all compression levels, yielded a lower cell count in most cases whereby most of the points represented in the Bland-Altman figure are above zero on the abscissa.

Moreover, the Kaplan Meier curves indicate a greater probability of observing substantial differences in the cell counts in the pair-wise

comparisons involving JPEG images compressed at 23x and 46x (Figure 6B).

As demonstrated in the previous part of the study,[35] the complexity of the images (number of nuclei/field and/or the presence of clusters) may affect the results of the automated quantification. According to the image classification previously described, the different degree of JPEG compression did not appear to affect the automatic quantification of images with ≤100 cells/field, without clusters or with small-area clusters. Effectively, the differences between the TIFF and the JPEG compressed images appeared in close proximity to the zero on the abscissa (Figure 7A and 7B). The probability of observing differences of ≤5 nuclei/image is around of 95-100% in images with a low density of positive cells (Figure 7C) as well as in images with or without small clusters (Figure 7D).

On the other hand, when images had >100 cells/images (Figure 8A) or large clusters (Figure 8B), there was greater dispersion of the difference in the Bland-Altman representations. The number of nuclei counted in TIFF format images was also higher than that in the corresponding JPEG format images. The probability of observing differences in JPEG images with a

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

high density of positive cells (Figure 8C) and also with large clusters (Figure 8D) was considerably higher. Count differences were <15 nuclei in more than 90% of cases for JPEG images with 3x compression. Otherwise, in images with 23x or 46x compression count differences were <40 nuclei in more than 90% of cases.

Although the Bland-Altman and Kaplan-Meier methods show very similar or small variations between counting results for uncompressed and compressed images, the t-test did not agree with these results. Globally, the paired t-test demonstrated significant differences in counting results between the TIFF and the JPEG images of all three compression levels (Table 2; $p<0.001$ for TIFF vs. 3x, 23x and 46x compressed JPEG images). However, these differences appear to be only significant for images with >100 cells/field or with high area cluster at all compression rates ($p<0.001$) and for TIFF vs. JPEG 3x in images with no clusters or small-area clusters ($p=0.002$). TIFF vs. JPEG 23x images tended to present significant differences with values near 0.05 in images with no clusters or small area clusters ($p=0.059$) and in images with ≤100 nuclei ($p=0.052$). These findings must be interpreted with caution since the significance was mostly due to the fact that variance of differences is very small (count differences

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

very high but equally distributed on both sides of zero on the abscissa in the Bland-Altman plot). On the other hand, the ICC coefficients showed excellent agreement but the condition of equality of variances was no satisfied.

12.3. Compression effects in the study of morphological parameters.

As showed in the Figure 9A, the different ratios of compression modify the color values of the pixels and more specifically the borderline pixels. These changes affect the positive objects shapes and size that modifies the results of the automated procedure as showed in Figure 9B. Variations in the size parameters were observed but the more pronounced variations were observed for the shape parameter (roundness) in single objects as well as in clusters. It was found that the roundness values decreased as the compression level increased.

The statistical analysis was applied to the mean of each parameter measured for all the selected objects. As illustrated in Table 3, values obtained for the different size parameters (area, perimeter1, diameter, length and width) did not show significant differences. The statistical differences between the mean values of the perimeter2 alone give only

statistical differences in 1:23 compression ratio. Roundness, the shape parameter that includes the perimeter2, was the only morphological feature of the objects that showed significant differences when comparing TIFF vs. JPEG 1:23 (Table 3A, p = 0.003) and TIFF vs. JPEG 1:46 (Table 3B, p = 0.002).

12.4. Correction of automated nuclei counting in compressed images.

12.4.1. Determination of corrective factors.

In this study, roundness was the only parameter considered that yielded statistically significant differences as a result of compression. Despite of there is only statistical differences in 1:23 and 1:46 compression ratios, a linear regression model was fitted to design the correction factors for each of the three compression levels. For this purpose, the roundness value of the object in TIFF format was taken as the dependent variable and that of the same object in the compressed JPEG images was the independent variable. The final linear regression models, where x and y are the original and corrected values, respectively, were:

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

JPEG 1:3 compression ratio:	y = 0.190 + 1.218 x	r =0.998
JPEG 1:23 compression ratio:	y = 0.048 + 1.885 x	r =0.987
JPEG 1:46 compression ratio:	y = 1.163 + 1.668 x	r =0.993

The three sets of regression parameters were incorporated in the automated procedure to achieve the correction of roundness in the compressed images. Student's t-test confirmed the accurate correction obtained with the linear regression coefficients. As a consequence, the previous significant differences in the roundness of positive objects in the TIFF compared with the JPEG compressed images (TIFF vs. JPEG 1:23, TIFF vs. JPEG 1:46) became non-significant by the corrections. Despite there are no significant differences in JPEG 1:3 compression, the roundness of each object is corrected. The real benefits of roundness correction in all three compression rates will be explained in the next section in the nuclei count correction. The roundness correction can be easily observed with Bland-Altman graphs (Figure 10) in which the differences between the uncorrected and corrected roundness values for TIFF vs. JPEG1:3, TIFF vs. JPEG1:23 and TIFF vs. JPEG1:46 were plotted against the mean of the roundness values for each pair of comparisons. Figure 10A1 shows the roundness differences between TIFF

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

and JPEG 1:3 without correction in roundness values of the compressed images. The Figure 10A2 shows the roundness differences between TIFF and JPEG 1:3 with correction in the roundness values of the compressed images. The same differences corrected and uncorrected roundness differences are represented for the other compression levels. In all the cases could be observed how corrected roundness values in compressed images are more similar to the TIFF roundness values due to differences are smaller in this comparisons. Despite of the outliers values in the graph, the majority of roundness values are between 0 and 10, and for this reason the mean roundness values are in the range from 4 to 7 (Table 3).

12.4.2. Quantification of nuclei with corrected roundness values in the Excel® macro.

Results obtained with the automated counting procedure with correction factors were compared with the counts made without roundness correction. In Figure 11, the Kaplan-Meier curves illustrate the conditional probability of observing differences between the TIFF and JPEG images for each comparison. The graphs indicate that there is a great reduction in count differences because of the use of correction factors. The best correction results were achieved in images with >100 nuclei or large area clusters.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

13. DISCUSSION.

It was developed a new single automatic procedure that allows the analysis of compressed and uncompressed DIs of various nuclear markers (ER, PR, Ki67, and FOXP3) with various degrees of complexity. These markers are used as prognostic markers in breast and lymphomas cancers.[36,37,65,72,73]

The reasons for the development of this new automatic procedure come from the fact that the great majority of the previously published studies used relatively simple images and did not estimate images with high area clusters of positive nuclei.[26-30] These procedures are unable to discriminate these types of images correctly. There are several approaches to segment nuclei in IHC images. Among them, various types of threshold, edge-based methods and mathematical morphology approaches. Edge-based methods are generically sensitive to image noise and artifacts. Algorithms for detecting the nuclear boundary are prone to error. The identification of noisy edges and discontinuous boundaries requires complex post-processing.[74,75] Other image analysis techniques, which only use the knowledge coming from the test of morphological features of the objects and multiple thresholding, do not provide good results, especially with high

number of nuclei and when there is an important variation in the size, shape and staining intensity of the cells.[30] Segmentation techniques that only use region-based algorithms are less sensitive to image noise, usually computationally more expensive and only useful for low area clusters.[22] Other studies use spatial filters such as dilation or erosion which can produce modifications and variations of the images.[24,27,76] Moreover their use is extremely complicated and it is needed to study in each case how to apply it to the different types of objects in the images.

The developed procedure presented in this thesis uses color segmentation, contrast enhancement, color masking and morphological classification of the objects according to features such as area and roundness to reduce artefacts in noisy images. The proposed method avoids obstacles described in the bibliography. Among them, the problems of staining intensity variations that are basically due to the technical procedures of IHC staining, conditioned by extrinsic and intrinsic variables related to specimen preparation and antigen staining.[2,7,8] The real problem in the quantification of this type of IHC images is the quantification number of nuclei in images with high area cluster. The developed procedure in this study makes an estimation of the number of nuclei in clusters that allows us to obtain an

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

accurate and definitive number of nuclei in the images. In high area clusters, other methods as local segmentation or watershed can lead to inaccurate results or oversegmented clusters.[19,21,28] The manual counting method was chosen as a reference because the human visual system is always which finally decides wether the quantification is good or not. The proposed method results agree in 90% of cases with those of the manual counting. The principal count differences appear for images with a high density of positive nuclei or with large clusters of these nuclei.

As previously mentioned, the biggest problem in automated IHC quantification[35] is the variation of the density of positively stained cells in each image and the presence of large clusters in images. These both features affect image compression, whereby as the compression level increases the differences in counts between uncompressed and compressed images increase.[71] However, the common problem of storage capacity of TIFF format images has led many investigators to use images in JPEG compressed formats without having tested the real impact of the compression on the final quantitative results. The investigations made in this study demonstrate that in these cases the efficiency of the automated procedure decreases with the degree of DI compression. Changes

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

produced during JPEG IC affect the colour values of each pixel[57] so that some positive pixels altered by the compression fall outside our selected range of positive colour and some negative pixels appearing to fall within the positive range of values. This variability could reduce the efficiency of the segmentation procedure and lead to inaccuracy the final nucleus count.

The alterations of the color of borderlines pixel in JPEG compressed images affect clearly the boundary (limits) of the objects selected during the segmentation process. Since the values of the different parameters (area, perimeter, diameter, length, width and roundness) are calculated from these limits. Morphometric analysis of the nucleus is known to be of great importance as a predictive, diagnostic factor and as an indicator of aggressive behavior.[47-49,77] It was found that from all measured parameters, roundness was the only one which showed significant differences between the TIFF format and the JPEG 1:23 and JPEG 1:46 compressed images, which initially prevented this compression formats from being used.

Furthermore, it has been shown how to improve computer-assisted nuclear quantification in compressed DIs. The results showed that IC affects roundness parameter since part of the automated methodology developed

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

is based on this parameter. As a result, a linear regression model was developed to correct the alterations in this parameter. The correction is good in more than 90% of cases, as illustrated by the very high correlation of “r” coefficients between the roundness values obtained from equivalent TIFF images and JPEG compressed files. These linear regressions were included in the automated count algorithms and produced corrected counts for the compressed formats. The count differences between TIFF images and compressed images with the correction were not statistically significantly. The roundness correction allowed us to use IC for nuclear quantification with the automated procedure presented. The residual differences from the definitive number of nuclear markers (≤5 cells per image) may be considered acceptable in images with low or big number of stained nuclei, since such small numbers would not be of clinicobiological relevance.

The small uncorrected nuclei count differences after roundness correction may be explained by the presence of very small positive (borderline) objects detected in TIFF format images that are not detected in those of the various compressed JPEG format images. Under these conditions, the compression eliminates some pixels in the positive range in such a way

that they cannot be detected according to the size criteria. In some cases, a borderline object disappears, while in others, large clusters may be split into two separate objects due to the presence of these borderline pixels (Figure 12).

The more accurate analysis of all types of complex images, improving the range of images that can be compressed for DI analysis, thereby allowing a greater reduction in the requirement for electronic media storage. The fast advance of technology for digital pathology (telepathology) and the high-capacity data of the actual computers are now a reality and permits to scan complete slides each one around 20-40GB of size in TIFF format. However, the current IHC studies with research and clinical practice purposes implicate the evaluation of a great number of slides containing entire tissues or extracts of tissues that occupy a lot of space.[71] In general, these studies required more that one image for each case which increases considerably the space necessary for the storage of these files. The great amount of information generated, makes nowadays necessary, the files size reduction and investigate which are the effects of IC in order to make an efficient and accurate image analysis.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

The results presented in this thesis showed that the measurement of the roundness parameter in compressed images could be corrected with a simply regression linear model for each compression level. This correction could be used in all computerized procedures that want to measure roundness in compressed images. As previously mentioned, the measurement of morphological parameters can have clinical implications in the study of several cancers, so this corrections allow us to correlate roundness with clinical parameter in compressed images without problems. Furthermore, any automated quantification procedure that uses roundness parameter will be able to obtain nuclei or cell counting without bias due to this parameter.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

14. CONCLUSION.

The presented study proposes a new automatic procedure used in our Pathology Department, which enables the quantification of IHC nuclear markers in DIs with varying degrees of complexity. This automation provides accurate results in Ki67, ER, PR and FOXP3 nuclear markers in breast and lymphoma tissue specimens. The procedure overcomes most of the difficulties currently described in the literature and offers independence of intra- and inter-observer variability. The results of quantification available with this new procedure allow us to confront the great diagnostic and therapeutic challenges in pathology.

This study demonstrated that a high degree of compression can be used, without any correction, in computer-assisted quantification of DIs with ≤100 nuclei or with low area clusters. The same levels of accuracy and reproducibility are obtained as with the original TIFF format images. The use of compression in these images allows reducing up to 46x the original size of the file, the overall disks space, the economical costs, and the time spent making security copies.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Furthermore, it is possible to use compressed DIs in order to quantify IHC markers in images with >100 nuclei or with high area clusters when a standard and simple linear regression model that permits roundness correction is incorporated to the procedure. This roundness correction permits to obtain similar nuclei counting results for compressed and uncompressed images. So, the presented system appears to be useful to clinical pathologists and can be one of the solutions for developing this type of systems.

The presented methodology can be incorporated in any commercial image analysis systems since the segmentation criteria used here can be introduced in any of these programs.

In the future, the potential uses of IC will be investigated in other automated quantification procedures in order to generalize counting of other nuclear, cytoplasmatic and membrane IHC markers.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

15. REFERENCE LIST

1. Taylor CR: Quantifiable internal reference standards for immunohistochemistry: the measurement of quantity by weight. Appl Immunohistochem Mol Morphol 14:253-9, 2006

2. Leong FJ, Leong AS: Digital imaging in pathology: theoretical and practical considerations, and applications. Pathology 36:234-41, 2004

3. Rhodes A, Jasani B, Anderson E, et al: Evaluation of HER-2/neu immunohistochemical assay sensitivity and scoring on formalin-fixed and paraffin-processed cell lines and breast tumors: a comparative study involving results from laboratories in 21 countries. Am J Clin Pathol 118:408-17, 2002

4. Seidal T, Balaton AJ, Battifora H: Interpretation and quantification of immunostains. Am J Surg Pathol 25:1204-7, 2001

5. Taylor CR: The total test approach to standardization of immunohistochemistry. Arch Pathol Lab Med 124:945-51, 2000

6. Rubin MA, Zerkowski MP, Camp RL, et al: Quantitative determination of expression of the prostate cancer protein alpha-methylacyl-CoA racemase using automated quantitative analysis (AQUA): a novel paradigm for automated and continuous biomarker measurements. Am J Pathol 164:831-40, 2004

7. Yaziji H, Barry T: Diagnostic Immunohistochemistry: what can go wrong? Adv Anat Pathol 13:238-46, 2006

8. de Jong D, Rosenwald A, Chhanabhai M, et al: Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol 25:805-12, 2007

9. McCarty KS, Jr., Miller LS, Cox EB, et al: Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. Arch Pathol Lab Med 109:716-21, 1985

10. Mosedale DE, Metcalfe JC, Grainger DJ: Optimization of immunofluorescence methods by quantitative image analysis. J Histochem Cytochem 44:1043-50, 1996

11. Johansson AC, Visse E, Widegren B, et al: Computerized image analysis as a tool to quantify infiltrating leukocytes: a comparison between high- and low-magnification images. J Histochem Cytochem 49:1073-79, 2001

12. Jagoe R, Steel JH, Vucicevic V, et al: Observer variation in quantification of immunocytochemistry by image analysis. Histochem J 23:541-7, 1991

13. Johnsson A, Olsson C, Anderson H, et al: Evaluation of a method for quantitative immunohistochemical analysis of cisplatin-DNA adducts in tissues from nude mice. Cytometry 17:142-50, 1994

14. Lejeune M, Jaen J, Pons L, et al: Quantification of diverse subcellular immunohistochemical markers with clinicobiological relevancies: validation

of a new computer-assisted image analysis procedure. J Anat 212:868-78, 2008

15. Gil J, Wu HS: Applications of image analysis to anatomic pathology: realities and promises. Cancer Invest 21:950-9, 2003

16. Veltri RW, Partin AW, Miller MC: Quantitative nuclear grade (QNG): a new image analysis-based biomarker of clinically relevant nuclear structure alterations. J Cell Biochem Suppl Suppl 35:151-7, 2000

17. Mofidi R, Walsh R, Ridgway PF, et al: Objective measurement of breast cancer oestrogen receptor status through digital image analysis. Eur J Surg Oncol 29:20-4, 2003

18. Nabi G, Seth A, Dinda AK, et al: Computer based receptogram approach: an objective way of assessing immunohistochemistry of androgen receptor staining and its correlation with hormonal response in metastatic carcinoma of prostate. J Clin Pathol 57:146-50, 2004

19. Singh SS, Kim D, Mohler JL: Java Web Start based software for automated quantitative nuclear analysis of prostate cancer and benign prostate hyperplasia. Biomed Eng Online 4:31, 2005

20. Anderson NH, Hamilton PW, Bartels PH, et al: Computerized scene segmentation for the discrimination of architectural features in ductal proliferative lesions of the breast. J Pathol 181:374-80, 1997

21. Ranefall P, Egevad L, Nordin B, et al: A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei. Anal Cell Pathol 15:145-56, 1997

22. Wu HS, Barba J, Gil J: A parametric fitting algorithm for segmentation of cell images. IEEE Trans Biomed Eng 45:400-7, 1998

23. Bauer KD, de la Torre-Bueno J, Diel IJ, et al: Reliable and sensitive analysis of occult bone marrow metastases using automated cellular imaging. Clin Cancer Res 6:3552-9, 2000

24. Loukas CG, Wilson GD, Vojnovic B, et al: An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections. Cytometry A 55:30-42, 2003

25. Matkowskyj KA, Cox R, Jensen RT, et al: Quantitative immunohistochemistry by measuring cumulative signal strength accurately measures receptor number. J Histochem Cytochem 51:205-14, 2003

26. Poon SS, Ward RK, Palcic B: Automated image detection and segmentation in blood smears. Cytometry 13:766-74, 1992

27. Lockett SJ, Herman B: Automatic detection of clustered, fluorescent-stained nuclei by digital image-based cytometry. Cytometry 17:1-12, 1994

28. Malpica N, de Solorzano CO, Vaquero JJ, et al: Applying watershed algorithms to the segmentation of clustered nuclei. Cytometry 28:289-97, 1997

29. Keenan SJ, Diamond J, McCluggage WG, et al: An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN). J Pathol 192:351-62, 2000

30. Lin G, Adiga U, Olson K, et al: A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. Cytometry A 56:23-36, 2003

31. Bartels PH, Bibbo M, Graham A, et al: Image understanding system for histopathology. Anal Cell Pathol 1:195-214, 1989

32. Bibbo M, Kim DH, di Loreto C, et al: Tissue architectural features for the grading of prostatic carcinoma. Anal Quant Cytol Histol 12:229-36, 1990

33. Garrido A. PdlBN: Applying deformable templates for cell image segmentation. Pattern Recognition 33:821-832, 2000

34. Camp RL, Chung GG, Rimm DL: Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 8:1323-7, 2002

35. Lopez C, Lejeune M, Salvado MT, et al: Automated quantification of nuclear immunohistochemical markers with different complexity. Histochem Cell Biol 129:379-87, 2008

36. Alvaro T, Lejeune M, Salvado MT, et al: Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. Clin Cancer Res 11:1467-73, 2005

37. Alvaro T, Lejeune M, Salvado MT, et al: Immunohistochemical patterns of reactive microenvironment are associated with clinicobiologic behavior in follicular lymphoma patients. J Clin Oncol 24:5350-7, 2006

38. Zhang K, Prichard JW, Yoder S, et al: Utility of SKP2 and MIB-1 in grading follicular lymphoma using quantitative imaging analysis. Hum Pathol 38:878-82, 2007

39. Trere D, Montanaro L, Ceccarelli C, et al: Prognostic relevance of a novel semiquantitative classification of Bcl2 immunohistochemical expression in human infiltrating ductal carcinomas of the breast. Ann Oncol 18:1004-14, 2007

40. Davis DW, Buchholz TA, Hess KR, et al: Automated quantification of apoptosis after neoadjuvant chemotherapy for breast cancer: early assessment predicts clinical response. Clin Cancer Res 9:955-60, 2003

41. Gil J, Wu H, Wang BY: Image analysis and morphometry in the diagnosis of breast cancer. Microsc Res Tech 59:109-18, 2002

42. Wang S, Saboorian MH, Frenkel EP, et al: Assessment of HER-2/neu status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence in situ hybridization assay as the standard. Am J Clin Pathol 116:495-503, 2001

43. Elhafey AS, Papadimitriou JC, El-Hakim MS, et al: Computerized image analysis of p53 and proliferating cell nuclear antigen expression in benign,

hyperplastic, and malignant endometrium. Arch Pathol Lab Med 125:872-9, 2001

44. Hannen EJ, van der Laak JA, Kerstens HM, et al: Quantification of tumour vascularity in squamous cell carcinoma of the tongue using CARD amplification, a systematic sampling technique, and true colour image analysis. Anal Cell Pathol 22:183-92, 2001

45. Williams SG, Buscarini M, Stein JP: Molecular markers for diagnosis, staging, and prognosis of bladder cancer. Oncology (Huntingt) 15:1461-70, 1473-4, 1476; discussion 1476-84, 2001

46. Milord RA, Lecksell K, Epstein JI: An objective morphologic parameter to aid in the diagnosis of flat urothelial carcinoma in situ. Hum Pathol 32:997-1002, 2001

47. Makarov DV, Marlow C, Epstein JI, et al: Using nuclear morphometry to predict the need for treatment among men with low grade, low stage prostate cancer enrolled in a program of expectant management with curative intent. Prostate 68:183-9, 2008

48. Cheretis C, Angelidou E, Dietrich F, et al: Prognostic value of computer-assisted morphological and morphometrical analysis for detecting the recurrence tendency of basal cell carcinoma. Med Sci Monit 14:MT13-19, 2008

49. Appel T, Bierhoff E, Appel K, et al: Predictive variables for the biological behaviour of basal cell carcinoma of the face: relevance of morphometry of the nuclei. Br J Oral Maxillofac Surg 41:147-50, 2003

50. Einstein AJ, Wu HS, Sanchez M, et al: Fractal characterization of chromatin appearance for diagnosis in breast cytology. J Pathol 185:366-81, 1998

51. Tambasco M, Magliocco AM: Relationship between tumor grade and computed architectural complexity in breast cancer specimens. Hum Pathol 39:740-6, 2008

52. Franzen LE, Hahn-Stromberg V, Edvardsson H, et al: Characterization of colon carcinoma growth pattern by computerized morphometry: definition of a complexity index. Int J Mol Med 22:465-72, 2008

53. Gerger A, Bergthaler P, Smolle J: An automated method for the quantification and fractal analysis of immunostaining. Cell Oncol 26:125-34, 2004

54. Cross SS: The application of fractal geometric analysis to microscopic images]. Micron 25:101-13, 1994

55. Cross SS: Fractals in pathology. J Pathol 182:1-8, 1997

56. Slone RM, Foos DH, Whiting BR, et al: Assessment of visually lossless irreversible image compression: comparison of three methods by using an image-comparison workstation. Radiology 215:543-53, 2000

57. Atalag K, Sincan M, Celasun B, et al: Effects of lossy image compression on quantitative image analysis of cell nuclei. Anal Quant Cytol Histol 26:22-7, 2004

58. Ohgiya Y, Gokan T, Nobusawa H, et al: Acute cerebral infarction: effect of JPEG compression on detection at CT. Radiology 227:124-7, 2003

59. Belair ML, Fansi AK, Descovich D, et al: The effect of compression on clinical diagnosis of glaucoma based on non-analyzed confocal scanning laser ophthalmoscopy images. Ophthalmic Surg Lasers Imaging 36:323-6, 2005

60. Gurdal P, Hildebolt CF, Akdeniz BG: The effects of different image file formats and image-analysis software programs on dental radiometric digital evaluations. Dentomaxillofac Radiol 30:50-5, 2001

61. Li F, Sone S, Takashima S, et al: Effects of JPEG and wavelet compression of spiral low-dose ct images on detection of small lung cancers. Acta Radiol 42:156-60, 2001

62. Oishi T, Sasaki A, Hamada N, et al: Proliferation and cell death of human glioblastoma cells after carbon-ion beam exposure: morphologic and morphometric analyses. Neuropathology 28:408-16, 2008

63. Brey EM, Lalani Z, Johnston C, et al: Automated selection of DAB-labeled tissue for immunohistochemical quantification. J Histochem Cytochem 51:575-84, 2003

64. Weaver DL, Krag DN, Manna EA, et al: Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. Mod Pathol 16:1159-63, 2003

65. Alvaro-Naranjo T, Lejeune M, Salvado-Usach MT, et al: Tumor-infiltrating cells as a prognostic factor in Hodgkin's lymphoma: a quantitative tissue microarray study in a large retrospective cohort of 267 patients. Leukemia Lymphoma 46:1581-91, 2005

66. Karlsson MG, Davidsson A, Hellquist HB: Quantitative computerized image analysis of immunostained lymphocytes. A methodological approach. Pathol Res Pract 190:799-807, 1994

67. Mesker WE, vd Burg JM, Oud PS, et al: Detection of immunocytochemically stained rare events using image analysis. Cytometry 17:209-15, 1994

68. Levine MD: Vision in Man and Machines. New York, McGraw-Hill, 1985

69. Ma W, Lozanoff S: A full color system for quantitative assessment of histochemical and immunohistochemical staining patterns. Biotech Histochem 74:1-9, 1999

70. van Der Laak JA, Pahlplatz MM, Hanselaar AG, et al: Hue-saturation-density (HSD) model for stain recognition in digital images from transmitted light microscopy. Cytometry 39:275-84, 2000

71. Lopez C, Lejeune M, Escriva P, et al: Effects of image compression on automatic count of immunohistochemically stained nuclei in digital images. J Am Med Inform Assoc 15:794-8, 2008

72. Alvaro T, Lejeune M, Camacho FI, et al: The presence of STAT1-positive tumor-associated macrophages and their relation to outcome in patients with follicular lymphoma. Haematologica 91:1605-12, 2006

73. Bosch Princep R, Lejeune M, Salvado Usach MT, et al: Decreased number of granzyme B+ activated CD8+ cytotoxic T lymphocytes in the inflammatory background of HIV-associated Hodgkin's lymphoma. Ann Hematol 84:661-6, 2005

74. Garbay C, Chassery JM, Brugal G: An iterative region-growing process for cell image segmentation based on local color similarity and global shape criteria. Anal Quant Cytol Histol 8:25-34, 1986

75. Ancin H, Roysam B, Dufresne TE, et al: Advances in automated 3-D image analyses of cell populations imaged by confocal microscopy. Cytometry 25:221-34, 1996

76. Underwood RA, Gibran NS, Muffley LA, et al: Color subtractive-computer-assisted image analysis for quantification of cutaneous nerves in a diabetic mouse model. J Histochem Cytochem 49:1285-91, 2001

77. Babik TM: Morphometric characteristics of epitheliocytes in the choroid plexus of the cerebral ventricles in humans during aging. Neurosci Behav Physiol 37:107-9, 2007

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

16. TABLES.

Table 1. Comparative analysis of manual and automated quantifications of 118 digital images.

Classification	Manual* [Mean ± SD]	Automatic [Mean ± SD]	Range	Student's t test	Intra-class correlation coefficient [IC 95%]	Spearman coefficient, Rho
Global	155.81 ± 125.95	158.94 ± 127.72	0-633	0.102	0.987 [0.981-0.991]	0.992 [p<0.0001]
Fewer than 100 nuclei	47.47 ± 27.64	48.92 ± 30.63	0-98	0.109	0.975 [0.956-0.985]	0.982 [p<0.0001]
More than 100 nuclei	244.14 ± 103.61	248.66 ± 104.22	102-633	0.187	0.965 [0.943-0.979]	0.963 [p<0.0001]
No nuclei or low area clusters	70.01 ± 62.81	69.09 ± 61.47	0-262	0.199	0.997 [0.994-0.999]	0.993 [p<0.0001]
High area cluster	221.11 ± 122.94	227.34 ± 122.76	18-633	0.062	0.975 [0.960-0.985]	0.975 [p<0.0001]

* mean of two manual quantifications by the same observer.

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Table 2. Probability values from Student's t-Tests comparing mean counts in TIFF and JPEG images of different compression rates.

Differences	TIFF vs. 3x JPEG	TIFF vs. 23x JPEG	TIFF vs. 46x JPEG
Globally	<0.001	<0.001	<0.001
No cluster/Low area cluster	0.002	0.059	0.125
Fewer than 100 nuclei	0.125	0.052	0.175
Large-area cluster	<0.001	<0.001	<0.001
More than100 nuclei	<0.001	<0.001	<0.001

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Table 3. Mean and standard deviation (SD) of morphological parameter estimates of global positive objects selected from the TIFF and different JPEG format images. a and b indicate statistically significant differences in mean values in comparisons (Student's t-test) with TIFF format images.

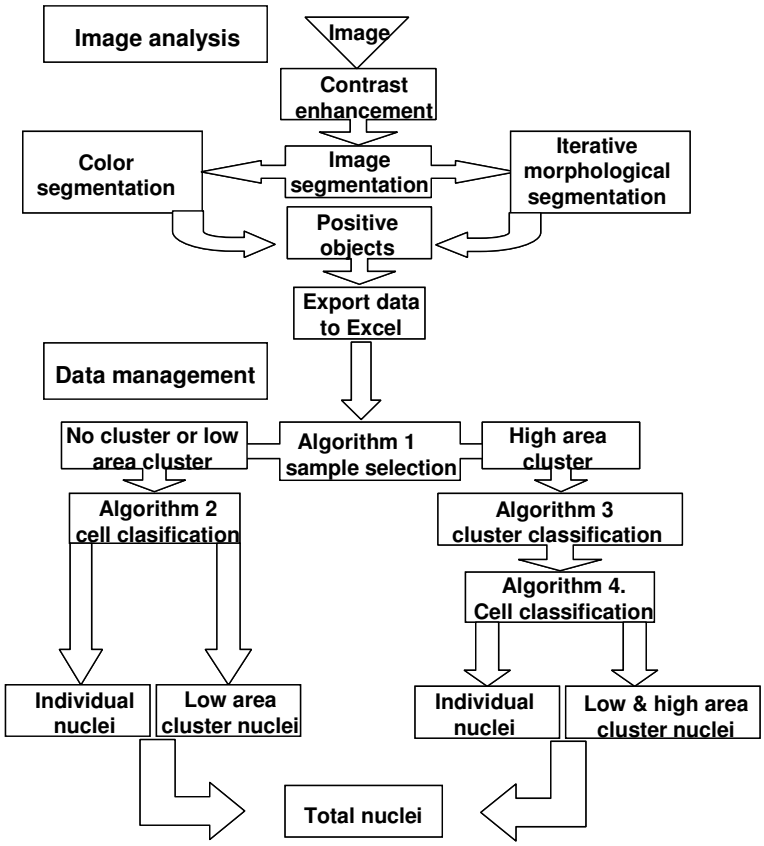
Mean of morphological measures	Area	Perimeter	Diameter	Length	Width	Roundness
 TIFF	67.71±101.98	36.20±41.37	7.04±5.69	10.11±9.31	6.70±6.33	7.50±7.67
JPEG 1:3	67.36±101.44	35.65±41.67	6.89±5.37	10.04±9.31	6.64±6.43	6.21±6.32
JPEG 1:23	67.43±101.48	35.80±42.01	6.98±5.66	10.03±9.26	6.62±6.35	3.94±4.02 [a]
JPEG 1:46	67.33±101.37	34.92±40.33	7.03±5.73	10.03±9.36	6.62±6.41	3.73±3.92 [b]

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

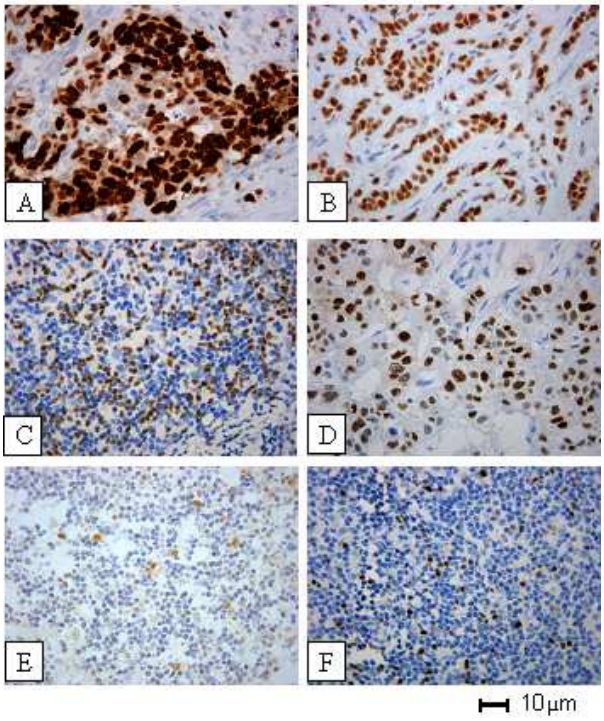
17. FIGURES.

Figure 1. Schematic representation of automatic quantification procedure in two steps that allows image analysis to determine the number of individual nuclei. Image analysis was carried out with Image-Pro® Plus 5.0 and data management with the algorithms introduced in the macro developed in the Excel datasheet.



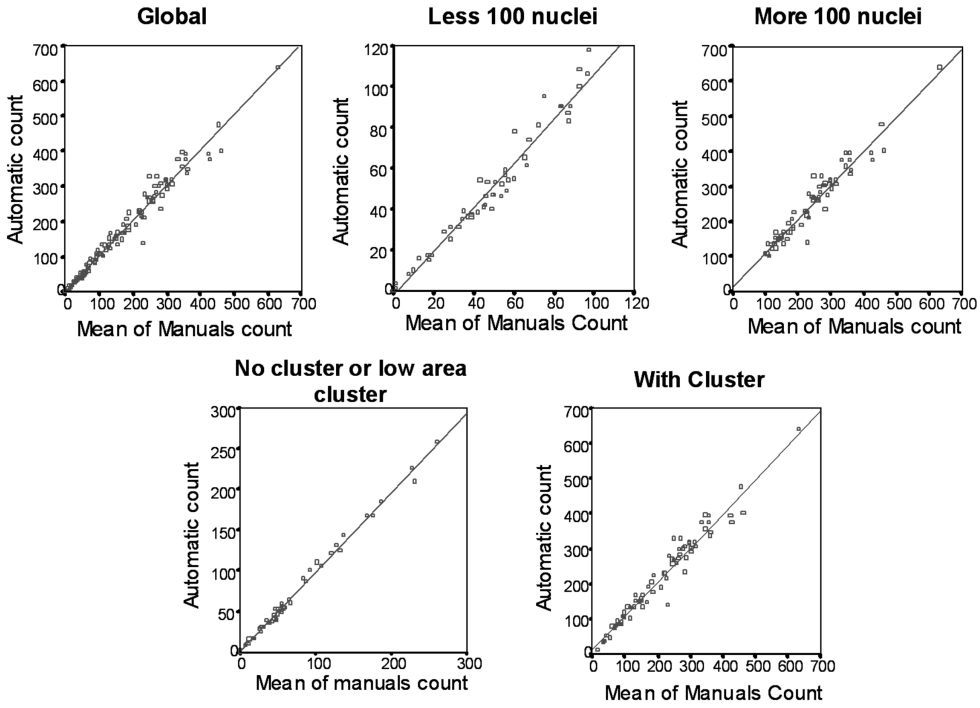
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Figure 2. Representation of immunohistochemical digital images with different cellular complexity from various nuclear markers. Images (**A**) and (**B**) correspond to Ki67 and ER, respectively, with a high nuclear concentration and containing high area clusters. Images (**C**) and (**E**) correspond to FOXP3 and Ki67 in images with high nuclear concentration and low area clusters. Finally, images (**D**) and (**F**) correspond to PR and FOXP3 images with low nuclei concentration and with no cluster and low area clusters. Panels **A**, **B**, **D**, and **E** correspond to breast tissues and panels **C** and **F** to lymph node tissues. Bar = 10 µm.



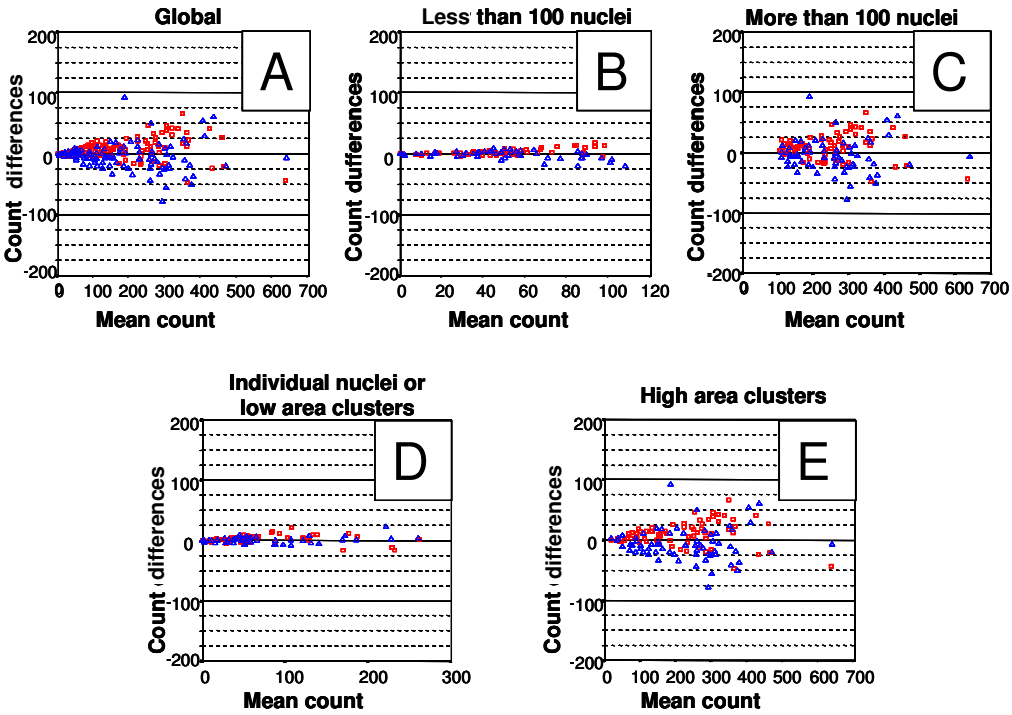
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Figure 3. Comparison of manual microscopic results (mean of two readings by the same observer) with those of the automated process by regression analysis by the different categories of image complexity.



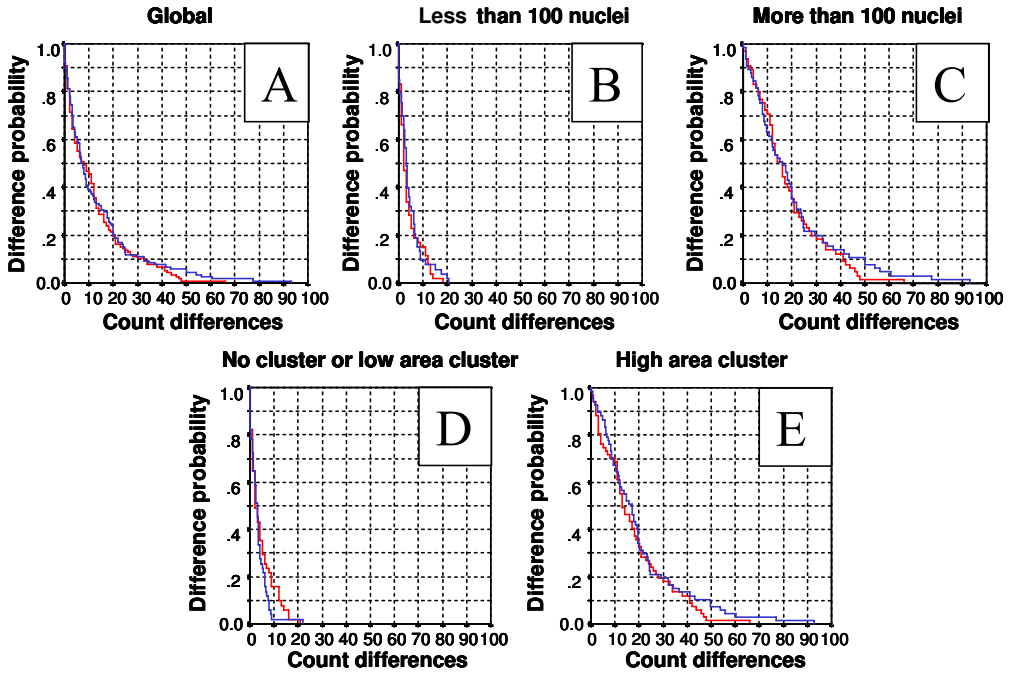
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Figure 4. Superimposed Bland–Altman graphs comparing differences between two manual microscopic readings (*red*), and between the mean of the manual microscopic readings and the automated method (*blue*) for the different subclassifications.



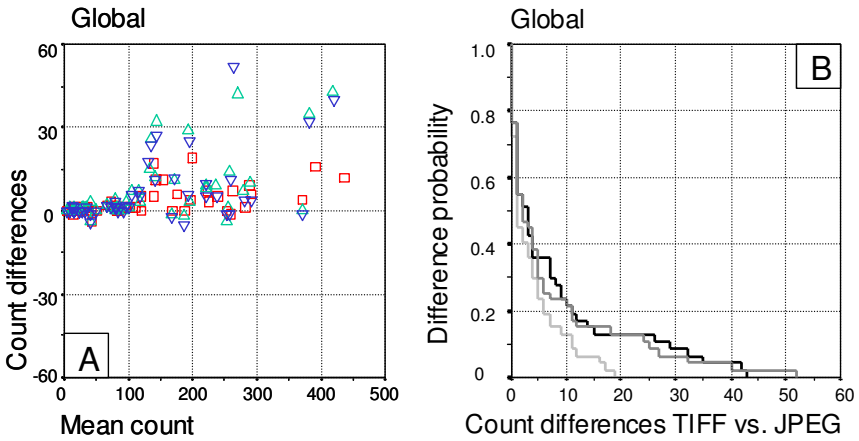
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Figure 5. Superimposed Kaplan–Meier curves comparing the probability of difference between the two manual microscopic readings (*red*), and between the mean of manual microscopic readings and the automated method (*blue*) for the different subclassifications.



UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

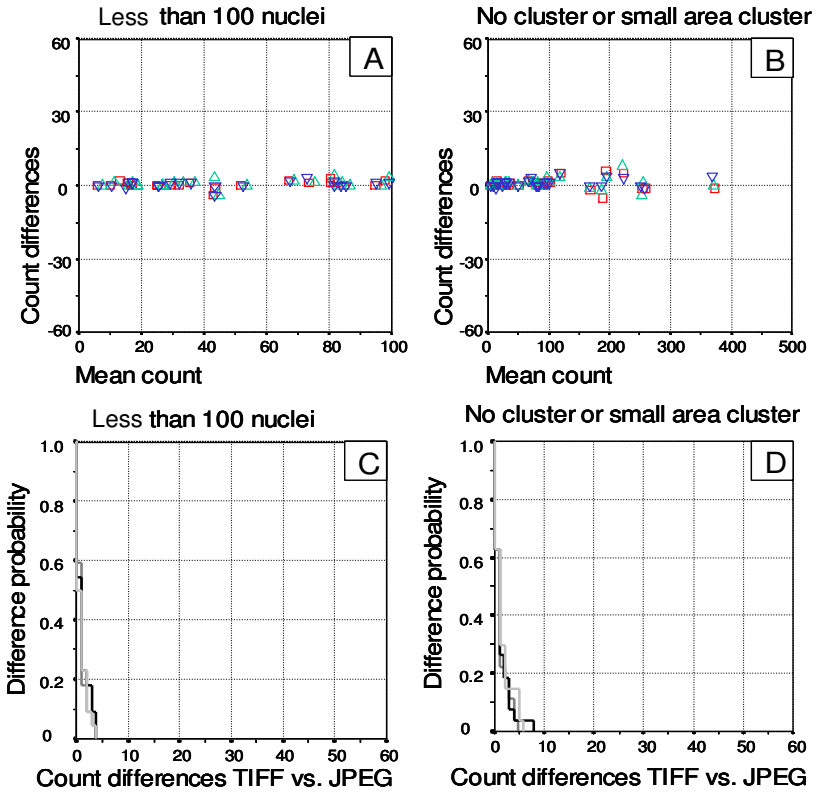
Figure 6. Global results obtained by automatic quantification. The Bland-Altman graph (A) represents the difference between each pair of results (Y axis) and the mean of both counts (X axis) for each comparison. The Kaplan-Meier curves (B) represent the conditional probability of observing differences between the TIFF and the JPEG images at different levels of compression. The X axis represents count differences between the number of nuclei in TIFF and JPEG images at 3X, 23X and 46X compression. The Y axis represents the probability of observing these count differences. **A.** Superimposed Bland-Altman graphs comparing count differences between TIFF and JPEG images. TIFF vs.3X compression JPEG (□), TIFF vs. 23X compression JPEG (△), TIFF vs. 46X compression (▽). **B.** Superimposed Kaplan- Meier curves comparing the probability of difference between TIFF and JPEG image counts. TIFF vs. 3X compression JPEG (—) TIFF vs. 23X compression JPEG (—), TIFF vs. 46X compression JPEG (—).



UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Figure 7. Results obtained by automatic quantification from images with ≤100 nuclei/image, with no cluster or small-area cluster. **A.** Superimposed Bland-Altman graphs comparing count differences between TIFF and JPEG images. TIFF vs. 3X compression JPEG (□), TIFF vs. 23X compression JPEG (△), TIFF vs. 46X compression JPEG (▽).

B. Superimposed Kaplan-Meier curves comparing the probability of difference between TIFF and JPEG image counts. TIFF vs. 3X compression JPEG (——), TIFF vs. 23X compression JPEG (——), TIFF vs. 46X compression JPEG (——).

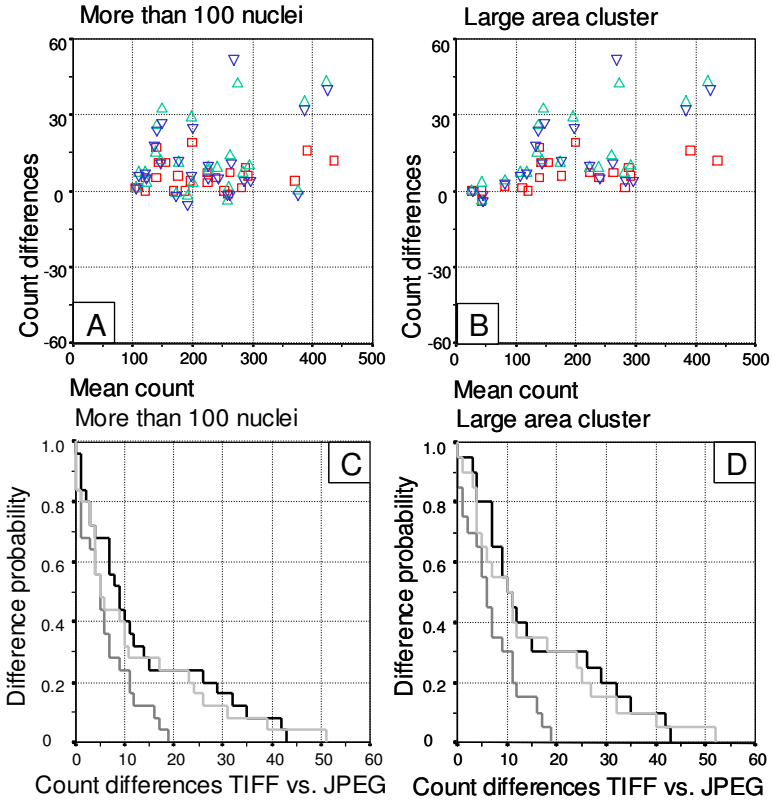


UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Figure 8. Results obtained with automatic quantification images with >100 nuclei/field or with large-area clusters.

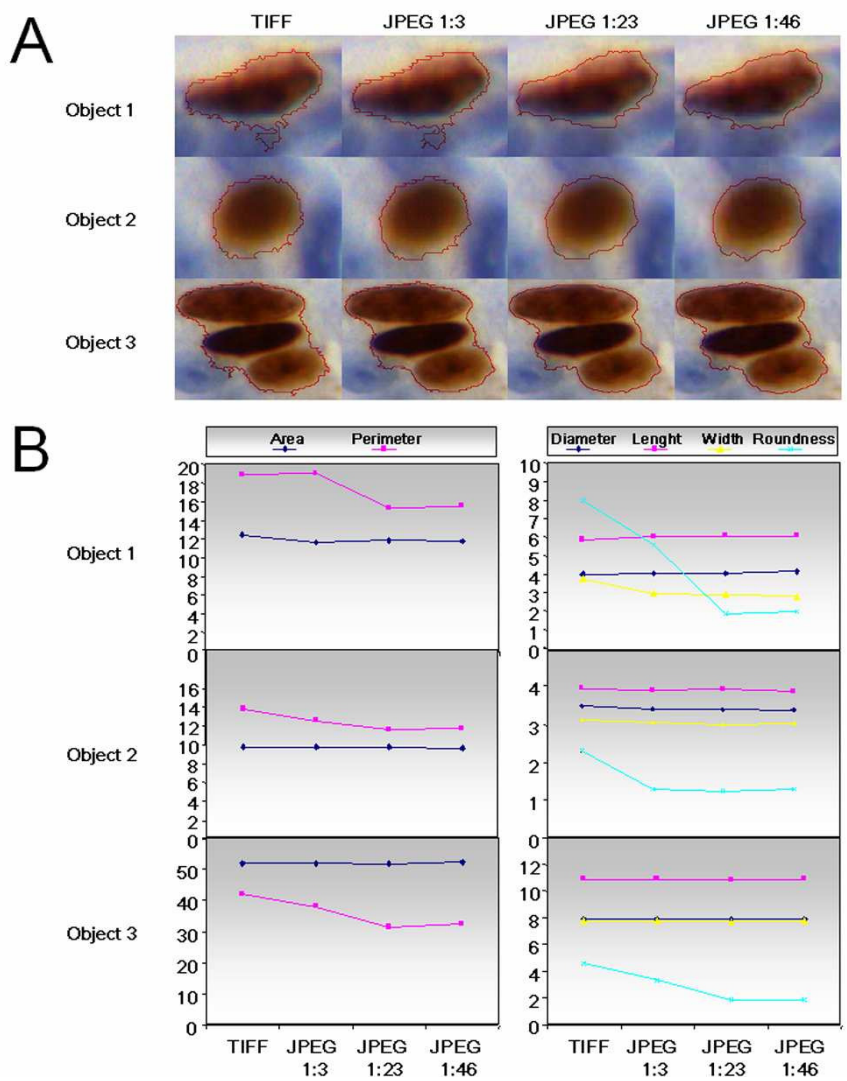
A. Superimposed Bland-Altman graphs comparing count differences between TIFF and JPEG images. TIFF vs. 3X compression JPEG(□), TIFF vs. 23X compression JPEG(▽), TIFF vs. 46X compression JPEG(△).

B. Superimposed Kaplan-Meier curves comparing the probability of difference between TIFF and JPEG image counts. TIFF vs. 3X compression JPEG (——), TIFF vs. 23X compression JPEG (——), TIFF vs. 46X compression JPEG (——).



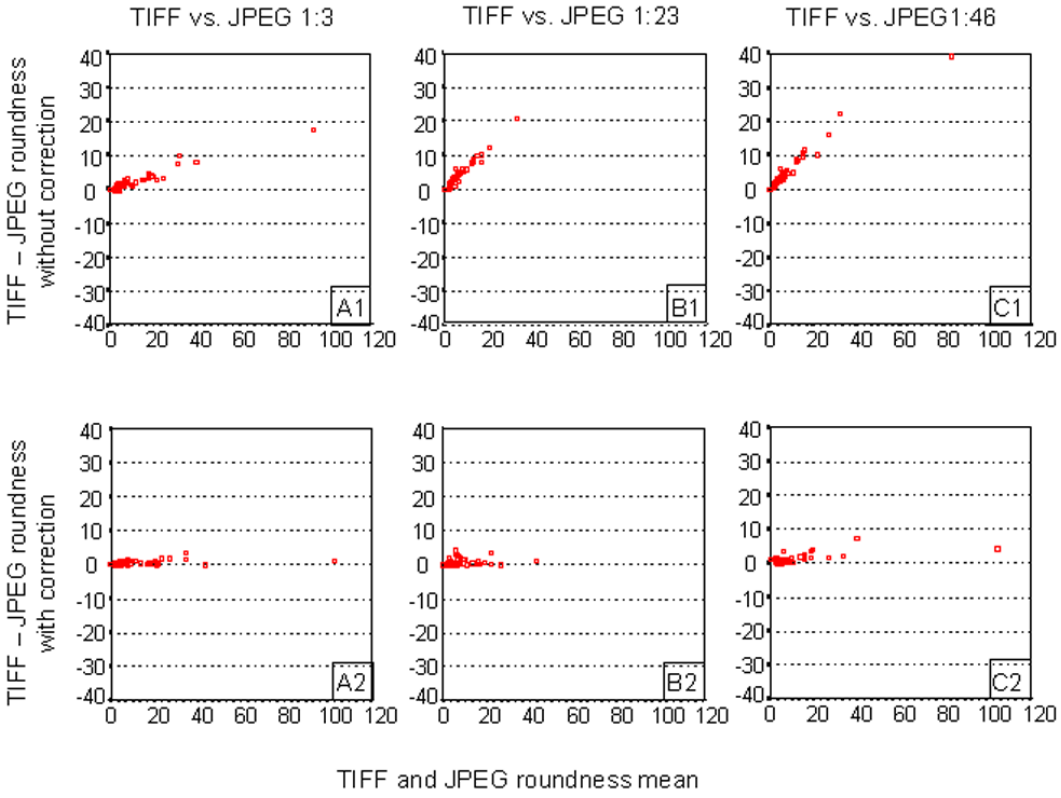
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Figure 9. Representation of the tracing limits of two individuals (objects 1 and 2) and one cluster objects (object 3) produced by region-based color segmentation algorithm in TIFF original image and the different JPEG compressed image (a). Variations in the values of each parameter are presented in second part of the figure for the 3 objects (b).



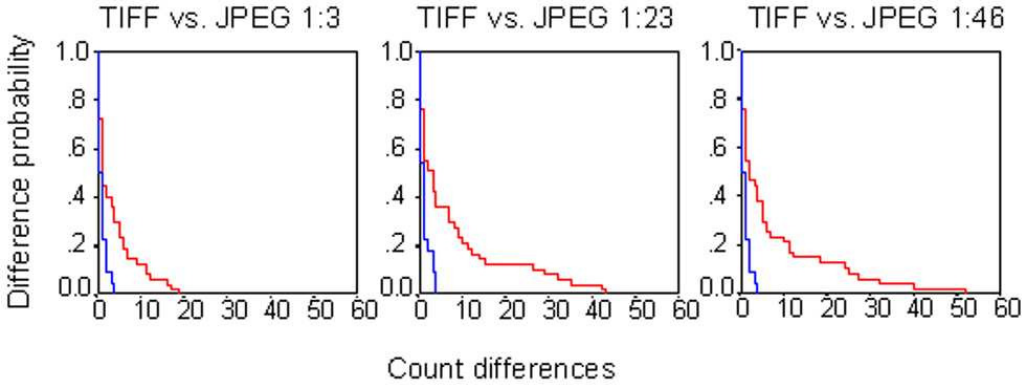
UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Figure 10. Bland–Altman graphs comparing roundness differences between objects analyzed in TIFF and JPEG formats. The upper row (a1, b1, c1) corresponds to the differences observed with the uncorrected roundness values and the lower row shows the differences from the corrected values (a2, b2, c2).



UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL:T-315-2011

Figure 11. Superimposed Kaplan–Meier curves comparing the probability of difference between the counts made in TIFF and JPEG formats. The red line represents the differences in the definitive count between TIFF and JPEG compressed images with uncorrected roundness. The blue line represents the differences between TIFF and JPEG compressed images with corrected roundness.



UNIVERSITAT ROVIRA I VIRGILI
DEVELOPMENT OF A NEW AUTOMATED METHOD FOR THE QUANTIFICATION OF NUCLEAR IMMUNOHISTOCHEMICAL MARKERS
Carlos López Pablo
ISBN.978-84-694-1253-4/DL.T-315-2011

Figure 12. Example of borderline positive pixels where in the right image (JPEG) is showed how a cell cluster could be separated from the right original image (TIFF) due to the effects of image compression, which alters the color values of positive pixels.