*Diagnostic Pathology* 2013, Volume 8 Suppl 1
http://www.diagnosticpathology.org/supplements/8/S1

DIAGNOSTIC PATHOLOGY

MEETING ABSTRACTS Open Access

# Proceedings of the 11th European Congress on Telepathology and 5th International Congress on Virtual Microscopy

Venice, Italy. 6-9 June 2012

Edited by Vincenzo Della Mea and Roberto Mencarelli

Published: 30 September 2013

These abstracts are available online at http://www.diagnosticpathology.org/supplements/8/S1

## INTRODUCTION

### S1

### 2012 – The beginning of a new world for Digital Pathology

Vincenzo Della Mea[1*], Roberto Mencarelli[2]
[1]Dept. of Mathematics and Computer Science, University of Udine, Italy; [2]Clinical Pathology Department, Local Healthcare Authority, Rovigo, Italy
*Diagnostic Pathology* 2013, **8(Suppl 1):**S1

The series of the biennial European congresses on Telepathology reached with this Venice edition its 11th appointment: more than 20 years of research in a continuously growing field. It is also the fifth time the congress brings the additional name of International Congress of Virtual Microscopy: eight years ago the International Academy of Telepathology decided to recognize this way the importance of virtual microscopy, also called Whole Slide Imaging, in the renovation of the telepathology scenario. In the last congress the International Academy of Digital Pathology was also founded, to replace and continue the great work of the International Academy of Telepathology that organized most of this congress series. This step recollects under the single term "digital pathology" all the technologies that are tranforming the traditional, "analog" pathologist in a digital professional that work on digitized slides, with software tools that enable to better exploit his/her knowledge, leaving routine details to the computers and networks. Once slides are digitized, a whole bunch of applications come natural and available: telediagnosis, teleconsultation, e-learning, long term storage, up to image analysis, with the forthcoming field of digital immunohistochemistry. No need to carry out additional operations: slides are ready for further digital treatment, present and future.
Digital Pathology is made possible not only thanks to the research grown in the last many years and presented at every Congress of this series. Every year the processors power increases, memory becomes larger and larger, and thus computers become more and more adequate to such large, information-filled objects that are the so-called digital slides.
This 2012 Congress hosted more than 80 presentations from 23 countries of the world. It represents the current state of the art in the field of digital pathology: what can be read in the present proceedings will represent the future of Pathology in the short, mid and even long term, thus providing insights on a new world which the traditional and crucial work of the pathologist should remain under his/her own control, made easier and more productive through digital tools.

## PROCEEDINGS

### S2

### Recent development and perspectives of virtual slides (VS) and telepathology in Europe

Klaus Kayser[1*], Stephan Borkenfeld[2], Gian Kayser[3]
[1]UICC-TPCC, Institute of Pathology, Charite, Berlin, Germany; [2]International Academy of Telepathology (IAT), Heidelberg, Germany; [3]Institute of Pathology, University of Freiburg, Freiburg, Germany
E-mail: Klaus.kayser@charite.de
*Diagnostic Pathology* 2013, **8(Suppl 1):**S2

**Background:** The new telepathology and virtual slide (VS) technology undergo remarkable changes in development and implementation. What are the reasons? What do we have to expect in the near future?
**Present stage:** Telepathology started in Europe in the 1980s. It was implemented in closed communication systems and focussed on frozen section service [1,2]. Open systems embedded in open source software replaced the early communication about 10 years ago. They became to age, and nearly all of them closed [3]. A new era started with the **m**edical **e**lectronic **c**ommunication **e**xpert **s**ystem (MECES, http://www.diagnomx.eu) and the **V**irtual **I**nternational **P**athology **I**nstitute (VIPI,http://www.diagnomx.eu/vipi) that combine an internal communication network with external information nodes. It uses acoustic and visual information transfer as well as information sources at different levels such as access to libraries, image content information analysis, and diagnosis assistants [3]. It also incorporates VS, which are available from different companies. Although delivered in non-congruent formats certain medical platforms (MECES) and open access scientific journals (journal of diagnostic pathology (http://www.diagnosticpathology.org)) can handle VS via their specific viewers. VS implementation in routine tissue – based diagnosis is on its way. Most companies try to specifically connect their VS scanners to laboratory information systems (LIS) and/or to digital radiology imaging systems (picture archiving and communication system (PACS)) [4,5]. VS are provided with their own specific image analyzing system that focuses on evaluation of suitable immunohistochemistry scores such as Her2_neu or hormone receptors in breast cancer. Obligatory VS standards are still missing [6,7].
**Expectations:** The establishment of a mandatory VS standard related to PACS is on its way [8]. Recent development of so – called social forums (facebook, linkedin, youtube, etc.) has lead to new communication

standards that permit the extension of open access and open software forums to external nodes. They can be considered as a communication system with internal structures {discussion groups, communication pathways (images, sounds, movies, functions), language, data sets, etc.} equipped with flexible communicative surface {interpretation, measurements, quality assurance, standardization, language translation, and others}. These information transfers can be switched on and off [9].
There are two different ways to incorporate VS in routine tissue – based diagnosis, namely (a) direct implementation of VS scanners in the existing LIS with specific connection to the hospital information system (HIS), and (b) to create an open communication network that provides as flexible communication surface to HIS, LIS, VS, etc. Industry seems to prefer method (a) although method (b) offers great advantages [10].
**Conclusion:** Europe is involved in big changes that involve the world of tissue – based diagnosis [11]. Surgical pathology starts to gain in clinical significance and financial interest. It is promoted by predictive diagnosis and communicative approaches which have their roots in telemedicine and digitized images [12]. The digitalization of surgical pathology has irreversibly started with big investment; the out come of the footrace is promising; however, it remains still open.

**References**

1. Kayser K, Szymas J, Weinstein RS: **Telepathology: Telecommunication, Electronic Education and Publication in Pathology.** Berlin, Heidelberg, New Yorck: Springer 1999.
2. Kayser K, Molnar B, Weinstein RS: **Virtual Microscopy - Fundamentals - Applications - Perspectives of Electronic Tissue - based Diagnosis.** VSV Interdisciplinary Medical Publishing 2006.
3. Kayser K, Borkenfeld S, Djenouni A, Kayser G: **History and structures of telecommunication in pathology, focusing on open access platforms.** *DiagnPathol* 2011, **6**:110.
4. Weinstein RS, *et al*: **Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future.** *Hum Pathol* 2009, **40(8)**:1057-69.
5. Laurinavicius A, *et al*: **Digital image analysis in pathology: benefits and obligation.** *Anal Cell Pathol (Amst)* 2012, **35(2)**:75-8.
6. Rojo MG, Castro AM, Goncalves L: **COST Action "EuroTelepath": digital pathology integration in electronic health record, including primary care centres.** *DiagnPathol* 2011, **6(Suppl 1)**:S6.
7. Rojo MG, Daniel C, Schrader T: **Standardization efforts of digital pathology in Europe.** *Anal Cell Pathol (Amst)* 2012, **35(1)**:19-23.
8. Jara-Lazaro AR, *et al*: **Digital pathology: exploring its applications in diagnostic surgical pathology practice.** *Pathology* 2010, **42(6)**:512-8.
9. Kayser K, *et al*: **E-education in pathology including certification of e-institutions.** *Diagn Pathol* 2011, **6(Suppl 1)**:S11.
10. Kelly DF: **Veterinary pathology in the United Kingdom: past, present, and future.** *J Vet Med Educ* 2007, **34(4)**:383-9.
11. Kayser K: **Introduction of virtual microscopy in routine surgical pathology - a hypothesis and personal view from Europe.** *Diagn Pathol* 2012, **7(1)**:48.
12. Della Mea V: **25 years of telepathology research: a bibliometric analysis.** *Diagn Pathol* 2011, **6(Suppl 1)**:S26.

## S3

### Fifth generation telepathology systems. Workflow analysis of the robotic dynamic telepathology Component

BL Braunhut, AR Graham, LC Richter, PD Webster, EA Krupinski, AK Bhattacharyya, RS Weinstein*
Departments of Pathology and Medical Imaging and Arizona Telemedicine Program, University of Arizona College of Medicine, Tucson, AZ, USA
E-mail: ronaldw@u.arizona.edu
*Diagnostic Pathology* 2013, **8(Suppl 1):**S3

**Background:** Fifth generation telepathology systems are dual-modality systems (i.e., digital pathology systems) that combine whole slide imaging (WSI) and real-time dynamic robotic telepathology. Hybrid robotic dynamic/static image telepathology systems were the precursors of fifth generation telepathology systems and can be used as surrogates for fifth generation systems in workflow studies [1-4].
Variability in human performance was identified as a pathology issue in the first scientific paper on telepathology, published in 1987 [5]. In a study performed under highly controlled conditions, it was shown that individual pathologists had a range of thresholds for diagnosing breast cancer on frozen sections. The use of the "equivocal for malignancy" diagnostic category varied significantly among the six pathologists enrolled in an Receiver Operating Characteristic (ROC) video microscopy human performance study. Surprisingly, use of this category was essentially the same for both conventional light microscopy and video microscopy for each pathologist [5]. In another early study, Dunn *et al*., using a hybrid dynamic robotic/static image telepathology system, documented patterns of telepathology primary diagnoses case deferrals for seven telepathologists over a 12-year period of time. These seven telepathologists rendered provisional primary surgical pathology diagnoses on over 11,000 surgical pathology cases. Case deferral rates among the telepathologists varied from 2.5% to 32.7% [6,7].
In this workflow study, we compared the QA case deferral rates of general pathologists and subspecialty surgical pathologists staffing a telepathology-based QA program. Fifth generation telepathology systems are an extension of the hybrid dynamic robotic/static image system concept, replacing the gallery of individual static images with a single large whole slide imaging (WSI) file.
**Materials and methods: Telepathology-based quality assurance service:** UMC initiated a QA telepathology consultation service between Lake Havasu City, Arizona and Tucson, Arizona (approximately 300 miles away) in July 2005. The Arizona Telemedicine Program (ATP) provided the broadband telecommunications infrastructure for the service as well as a telepathology case coordinator. These QA services continued uninterrupted until October 2009 when HRMC changed direction and decided to outsource all of its laboratory services to a commercial reference laboratory for cost savings.
During this study, the HRMC pathology laboratory handled 3,000-4,000 surgical pathology cases per year and was locally staffed by a single pathologist. Surgical cases at HRMC were accessioned, grossed, embedded in paraffin and glass slides were produced in the HRMC laboratory. All new cancer cases and any other challenging surgical pathology cases were identified by the HRMC pathologist for telepathology QA review.
QA cases reviewed via telepathology had been diagnosed with a written report generated by the HRMC pathologist prior to telepathology. A remotely controllable hybrid robotic-dynamic telepathology system (Apollo PACS ®, Falls Church, VA) was used to transmit a stream of digital images, via the ATP network, from Lake Havasu City, Arizona to Tucson, Arizona. The case video images were reviewed in real-time and remotely navigated by the on-service triage telepathologist at UMC, in Tucson. The telepathologist, linked to the HRMC pathologist in Lake Havasu City, was able to collaborate face-to-face via real-time videoconferencing, a feature built into the Apollo system. At the time of each telepathology review, the HRMC pathology report and accompanying patient medical history/demographic data were made available to the telepathologist in Tucson.
Ten UMC telepathologists participated in the study. They had dual roles, functioning either as general telepathologists and/or as subspecialty surgical pathologists depending on the nature of the case. Nine surgical pathology subspecialties were represented, including dermatopathology, gastrointestinal/hepatic, renal/genitourinary, breast, thoracic, gynecologic and head/neck pathology. The CTP case workflow model, as described by Bhattacharyya *et al*., is shown schematically in Figure 1 [8].
**Study cases and data analysis:** Between July 2005 and October 2009, 1862 cases were transmitted from HRMC to UMC.
After completion of case accruals, data analysis commenced. The provisional surgical pathology reports (HRMC) and the final telepathology case reports (UMC) were compiled. Data compiled included date of telepathology review, name of telepathologist, specimen organ system, preliminary diagnosis, and final telepathology diagnosis. These data were uploaded into a Microsoft Excel spreadsheet and tallied for case load per telepathologist, number of cases per organ system, and deferral rates. The HRMC and UMC diagnoses were compared and classified by a senior pathologist as being either concordant or discordant. Discordant diagnoses were further sub-classified by the staff pathologist as a major discrepancy (one that would alter clinical management) or a minor discrepancy (one that would not change clinical management).
Cases were further evaluated with regards to the telepathologist's area of subspecialty surgical pathology expertise. A t-test for paired observations was used to determine if there was a statistically significant difference in deferral rates and a Chi-square test was done to determine if the distributions of deferral rates differed across readers (null hypotheses=deferral rate expected to be the same).

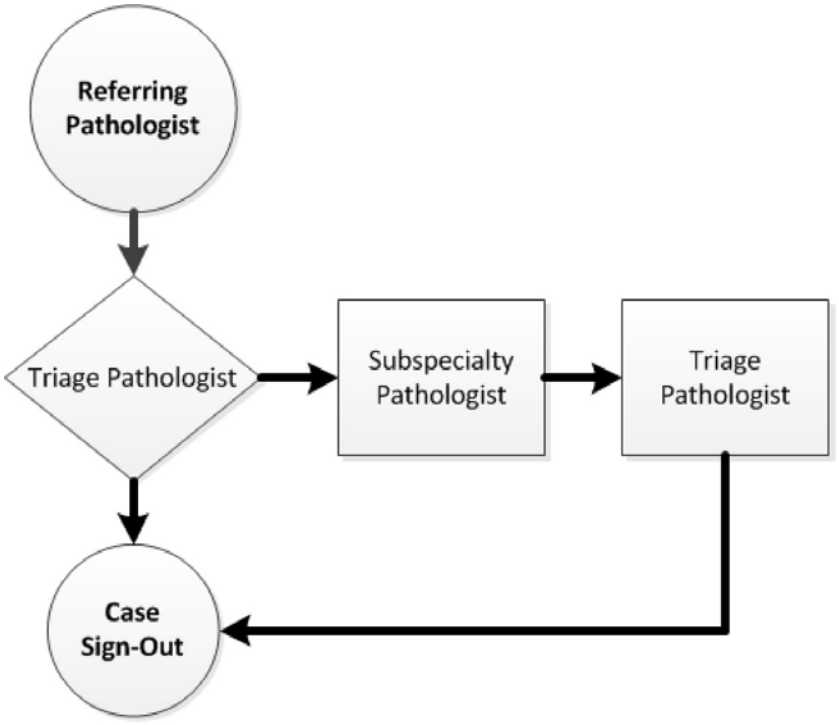



**Figure 1(abstract S3) CTP workflow model** Workflow for the Case Triage Practice (CTP) Model used in the Lake Havasu-Tucson, Arizona, telepathology quality assurance service. When triage pathologists consult with subspecialty pathologists, they retain responsibility for generating the final QA telepathology report. The service used an Apollo hybrid dynamic robotic/static image telepathology system. (Adapted from Bhattacharyya et al., 1995 [8]).

**Results and discussion:** Of the 1815 analytical cases, 1650 cases (90.91%) were signed out directly by the triage telepathologist. The remaining 165 cases (9.09%) were deferred for further analysis by a subspecialty surgical pathologist or for special studies such as immunohistochemistry.

**Levels of concordance:** Overall concordance was 94.27%. Concordance by specimen site/organ system ranged from 90.09% to 100%. The greatest discordance rate was seen with genitourinary cases with an overall discordance rate of 9.91% (7.66% major discrepancies, 2.25% minor discrepancies). Concordance on breast cases, cardiovascular cases, and head and neck cases was 100%. Table 1 includes 42 cases which were initially viewed by two telepathologists.

**Table 1(abstract S3) Concordance of HRMC and UMC diagnoses**

| Organ/site | Cases | Agree | Major disc. | Minor disc. |
|---|---|---|---|---|
| Gastrointestinal | 494 | 474 (95.95%) | 8 (1.62%) | 12 (2.43%) |
| Genitourinary | 444 | 400 (90.09%) | 34 (7.66%) | 10 (2.25%) |
| Skin | 246 | 230 (93.50%) | 5 (2.03%) | 11 (4.47%) |
| Lungs | 240 | 235 (97.92%) | 1 (0.42%) | 4 (1.67%) |
| Bone/soft tissue | 87 | 81 (93.10%) | 1 (1.15%) | 5 (5.75%) |
| Head/neck | 63 | 63 (100%) | 0 (0%) | 0 (0%) |
| Gynecological | 49 | 45 (91.84%) | 0 (0%) | 4 (8.16%) |
| Breast/axilla | 38 | 38 (100%) | 0 (0%) | 0 (0%) |
| Endocrine | 21 | 19 (90.48%) | 0 (0%) | 2 (9.52%) |
| Cardiovascular | 10 | 10 (100%) | 0 (0%) | 0 (0%) |
| Total | 1692 | 1595 (94.27%) | 49 (2.90%) | 48 (2.83%) |

**Deferral rates:** The case volumes per telepathologist ranged from 51 to 501 cases (average 182 cases). Deferral rates for individual telepathologists ranged from 4.79% to 21.26% (average 10.05%). Deferral rates were minimally changed by exclusion of cases within each telepathologist's subspecialty area and ranged from 4.94% to 21.81% (average 10.26%). These data are summarized in Table 2. A t-test showed no statistically significant difference in deferral rates for case triage telepathologists for cases outside their areas of subspecialty pathology expertise versus triage cases falling within their area of subspecialty surgical pathology expertise ($t = 0.032$, $p = 0.9754$). However, 8 out of 10 telepathologists deferred a lower percentage of telepathology cases that fell within their area of surgical pathology expertise which may represent a trend (Table 2). A Chi-square test for distribution of deferral rates across readers was statistically significant for general rates ($X^2 = 20.52$, $p < 0.05$) and subspecialty rates ($X^2 = 20.23$, $p < 0.05$).

This study evaluated a real-time telepathology QA program using the CTP case workflow model [8]. The overall concordance rate between primary (HRMC) and final (UMC) diagnosis was 94.27%. Of the discordant diagnoses, 2.90% represented major discrepancies and 2.83% represented minor discrepancies. In clinical practice, this discordance rate could be of concern [9,10].

Overall discordance rates were minimally changed with exclusion of cases within each telepathologist's subspecialty area. Conversely, UMC subspecialty surgical pathologists performed well in the role of general triage pathologist (Figure 1). This supports the use of subspecialty surgical pathologists as the general triage pathologist in a telepathology-based QA program.

**Conclusion:** This CTP workflow telepathology model, blending the services of university-based subspecialty surgical pathologists and a community-based general pathologist, provided a means for improving the quality of community-based laboratory services, presented opportunities for UMC-based community outreach, and increased the efficiency of a second opinion QA program. Service provider and user satisfaction were high.

Based on these data and observations, we suggest that the likelihood of a reviewing telepathologist agreeing or disagreeing with a diagnosis

**Table 2(abstract S3) Deferral rates**

| Pathologist | Total cases | Deferred cases | Total cases excluding pathologists subspecialty | Total deferred cases excluding pathologists subspecialty | Deferral rate overall | Deferral rate of cases within pathologists subspecialty | Deferral rate excluding pathologists subspecialty |
|---|---|---|---|---|---|---|---|
| Pathologist A | 501 | 24 | 344 | 17 | 4.79% | 4.46% | 4.94% |
| Pathologist B | 369 | 30 | 321 | 25 | 8.13% | 10.42% | 7.78% |
| Pathologist C | 188 | 24 | 150 | 22 | 14.79% | 5.26% | 14.67% |
| Pathologist D | 174 | 37 | 165 | 36 | 21.26% | 11.11% | 21.81% |
| Pathologist E | 166 | 12 | 161 | 12 | 7.23% | 0% | 7.45% |
| Pathologist F | 139 | 12 | 109 | 10 | 8.63% | 6.67% | 9.17% |
| Pathologist G | 85 | 9 | 83 | 9 | 10.59% | 0% | 10.84% |
| Pathologist H | 84 | 6 | 76 | 6 | 7.14% | 0% | 7.89% |
| Pathologist I | 58 | 7 | 50 | 5 | 12.07% | 25% | 10% |
| Pathologist J | 51 | 4 | 50 | 4 | 7.84% | 0% | 8% |

rendered at an outside hospital is not a function of the reviewers' expertise alone, but rather may be related to other human factors, even possibly the personality type of the telepathologist. For individual pathologists, there was a strong relationship between rates of case deferrals for cases within their own area of subspecialty expertise as compared with cases they handled that were outside of their area of subspecialty expertise. Low, intermediate, and high level users of the "case deferral" option could be identified. It would be of interest to conduct Myers-Briggs personality assessments on pathologists and to correlate personality assessment results with other quantifiable pathologist performance measures such as surgical pathology case deferral rates, rates of equivocation on malignant diagnoses, and others.

**List of abbreviations:** ATP: Arizona Telemedicine Program; CTP: Case Triage Practice; HRMC: Havasu Regional Medical Center; QA: quality assurance; SPP: subspecialty pathology practice; UMC: University Medical Center; WSI: whole slide imaging.

**Competing interests:** Dr. Weinstein is Scientific Director and has stock options in Apollo PACS®.

**Authors' contributions:** BB compiled, analyzed and reviewed data sets, and authored the first draft of the manuscript. AG reviewed, edited, and composed portions of the manuscript. LR collected data and provided technical support for the original telepathology consultation sessions. PW was clinical coordinator for the study and assisted with data compilation. EK reviewed the data and performed the statistical analysis. AB was Laboratory Medical Director overseeing the telepathology consultation service and provided telepathology support. RW conceived the study, authored sections of the manuscript, and provided telepathology expertise. All authors read and approved the final manuscript.

**Acknowledgements:** We thank Fangru Lian, MD, Raymond Nagle, MD-PhD, and Judith Pugh, MD for their pathology expertise and contributions to the review of data sets for this study.

**References**

1. Weinstein RS, Graham AR, Lian F, Braunhut BL, Barker GR, Krupinski EA, Bhattacharyya AK: **Reconciliation of diverse telepathology system designs. Historical issues and implications for emerging markets and new applications.** *APMIS* 2012, **120(4)**:256-275.
2. Kaplan KJ, Weinstein RS, Pantanowitz L: **Telepathology.** *Pathology informatics: modern practice & theory for clinical laboratory computing* american society for clinical pathology press: Pantanowitz L, Balis U, Tuthill M 2012, 257-272.
3. Della Mea V, Cataldi P, Pertoldi B, Beltrami CA: **Combining dynamic- and static-robotic techniques for real-time telepathology.** *Telepathology* Berlin, Springer: Kumar S, Dunn BE 2009, 79-89.
4. Weinstein RS, Descour MR, Liang C, *et al*: **Telepathology overview: from concept to implementation.** *Hum Pathol* 2012, **32(12)**:1283-99.
5. Weinstein RS, Bloom KJ, Rozek LS: **Telepathology and the networking of pathology diagnostic services.** *Arch Pathol Lab Med* 1987, **111(7)**:646-52.
6. Dunn BE, Choi H, Almagro UA, Recla DL, Krupinski EA, Weinstein RS: **Routine surgical pathology in the Department of Veterans Affairs: experience-related improvements in pathologist performance in 2200 cases.** *Telemedicine J* 1999, **5(4)**:323-337.
7. Dunn BE, Choi H, Recla DL, Kerr SE, Wagenman BL: **Robotic surgical telepathology between the Iron Mountain and Milwaukee Department of Veterans Affairs Medical Centers: a 12-year experience.** *Hum Pathol* 2009, **40(8)**:1092-1099.
8. Bhattacharyya AK, Davis JR, Halliday BE, Graham AR, Leavitt SA, Martinez R, Rivas R, Weinstein RS: **Case triage model for the practice of telepathology.** *Telemedicine J* 1995, **1(1)**:9-17.
9. Weinstein RS, Graham AM, Richter LC, Barker GP, Krupinski EA, Lopez AM, Erps KA, Yagi Y, Gilbertson JR, Bhattacharyya AK: **Overview of telepathology, virtual microscopy and whole slide imagining: Prospects for the future.** *Hum Pathol* 2009, **40(8)**:1057-1069.
10. Kayser K, Szymas J, Weinstein RS: **Telepathology and Telemedicine: Communication, Electronic Education and Publication in e-Health.** VSV Interdisciplinary Medical Publishing, Berlin 2005, 1-257.

## S4

### Construction of web-based remote diagnosis system using virtual slide for routine pathology slides of the rural hospital in Japan

Ichiro Mori[1*†], Takashi Ozaki[2†], Yasuteru Muragaki[2†], Takatoshi Ibata[3†], Hiroshi Ueda[3†], Toshihito Shinagawa[4†], Yoshiyuki Osamura[1†]
[1]Department of Pathology, Mita hospital, International University of Health and Welfare, Minato-ku, Tokyo, Japan; [2]Department of Pathology, Wakayama Medical University, Wakayama, Japan; [3]Shingu Municipal Medical Center, Wakayama, Japan; [4]Kawasaki Municipal Ida Hospital, Kanagawa, Japan
E-mail: ichiro-m@iuhw.ac.jp

*Diagnostic Pathology* 2013, **8(Suppl 1):**S4

**Background:** In recent years, Japan has been short of pathologists. There are about 2,000 board-certified pathologists against 130 million people. Most pathologists are working in the urban area and only few are available in rural regions. Telepathology in Japan developed under these circumstances, mainly for intraoperative frozen section diagnosis where no pathologist is available. We have been performing frozen section telepathology for more than 10 years, and achieved fairly good results [1].

In this work, we constructed a web-based remote diagnosis system using virtual slide targeting routine pathology slides.

**Material and methods:** The hospital has about 2,000 histology specimens every year, and there are three pathology technicians in the pathology laboratory. Part-time pathologists used to visit the hospital twice a week. We started this project because one of the two pathologists moved, and could no longer go to the hospital. We used LINCE from CLARO (Aomori, Japan), a tiling type VS scanner. To scan slides, we fixed the objective lens to 40x. We limited the target to the biopsy specimens. Pathology technicians scaned paraffin slides and stored VS to the shared folder of their server. They also scanned the clinical application form, and saved the PDF files to the same folder. Personal information is excluded for security reasons. The pathologists opened a shared folder of the server through the

internet from Tokyo, 500 km from the hospital (see Figure 1). First, we checked the serial pathology numbers of VS and application form. Then we read the application form, observed VS, making a primary diagnosis report using word processor containing the serial pathology number in the report. The diagnosis reports were then stored to the same shared folder. This was done using remote VPN software. Technicians copied the report and pasted it to the reporting system to complete diagnosis.

In some difficult cases, we discussed with pathologist who visits the hospital once a week, and compared the diagnosis between VS and microscope.

**Results and discussion:** We started this system in July 2011, and continue to diagnose about 80 cases each month resulting in more than 800 primary diagnosis. The VS images were clear enough and we could diagnose most of the routine HE slides with no incident. No image degradation happened due to data transfer. Because VS are stored in the server and connected to the web, we can manage our time and location for diagnosis easily. I have also been able to view VS images from Europe using the internet at my hotel.

As a setting of VS scanner, we first tried to use 20x objective lens. But in some delicate cases, the images were frustrating for us. Since we are unable to predict the delicate cases, we fixed objective lens to 40x.

When performing a diagnosis on the PC, we usually open 5 windows, 1) clinical application form, 2) VS viewer, 3) word processor window to write diagnosis report, 4) shared VS data folder, and 5) folder to temporary store the diagnosis report. This is too much to display on 10 inch screen notebook PC, or a 22 inch desktop screen. As a result, we are using a dual display system (see Figure 2).

To check the discrepancies between direct microscope and remote VS diagnosis, we looked at some difficult cases and discussed together. The degrees of coincidence were fairly good.

The biggest issue was the data transfer speed. The response of the VS differed each time, sometimes it was really slow whilst at other times it was very fast. When comparing VS observation using broadband of 50-60 Mbps to a mobile telephone network of 1-5 Mbps, the difference was small. We also changed the internet connection of the hospital's VS server which resolved the issue dramatically. There should be bottleneck somewhere between the VS server and the web this time.

One big trouble was misunderstanding of the slide direction of VS. We put multiple small specimens on one glass slide, and numbered them from the label side. Because our VS viewer does not show thumbnail images with labels, a pathologist cannot find which side in #1 on viewer. Technicians usually set glass slides to the VS scanner with the slide label on the right hand side, which makes the right side specimen #1. One day, a glass slide was accidently placed onto the scanner the opposite way round. There were 4 colon specimens on the VS, and the 2$^{nd}$ specimen from the right showed cancer. The specimen #2 was taken from the ascending colon, but actually it was the specimen #3 taken from the descending colon. Luckily we noticed this before the operation.

According to the Japanese Medical Practitioners Act, primary pathology diagnosis is designated as a medical act, and medical act must take place in medical institution. Using this web-based system, we can make primary diagnosis from anywhere in the world. But in accordance to the law, we kept the report as draft when made outside of the medical institution and it was later finished at the hospital.

Because we cannot access the hospital's pathology reporting system directly from Tokyo, we cannot write the diagnosis directly. At least, this could help us because the pathology technicians can find simple mistakes in the report when they copy it.

**Conclusions:** There are many ideas reported for the usage of VS and Telepathology [2-5]. This time, we focused on the web-based primary

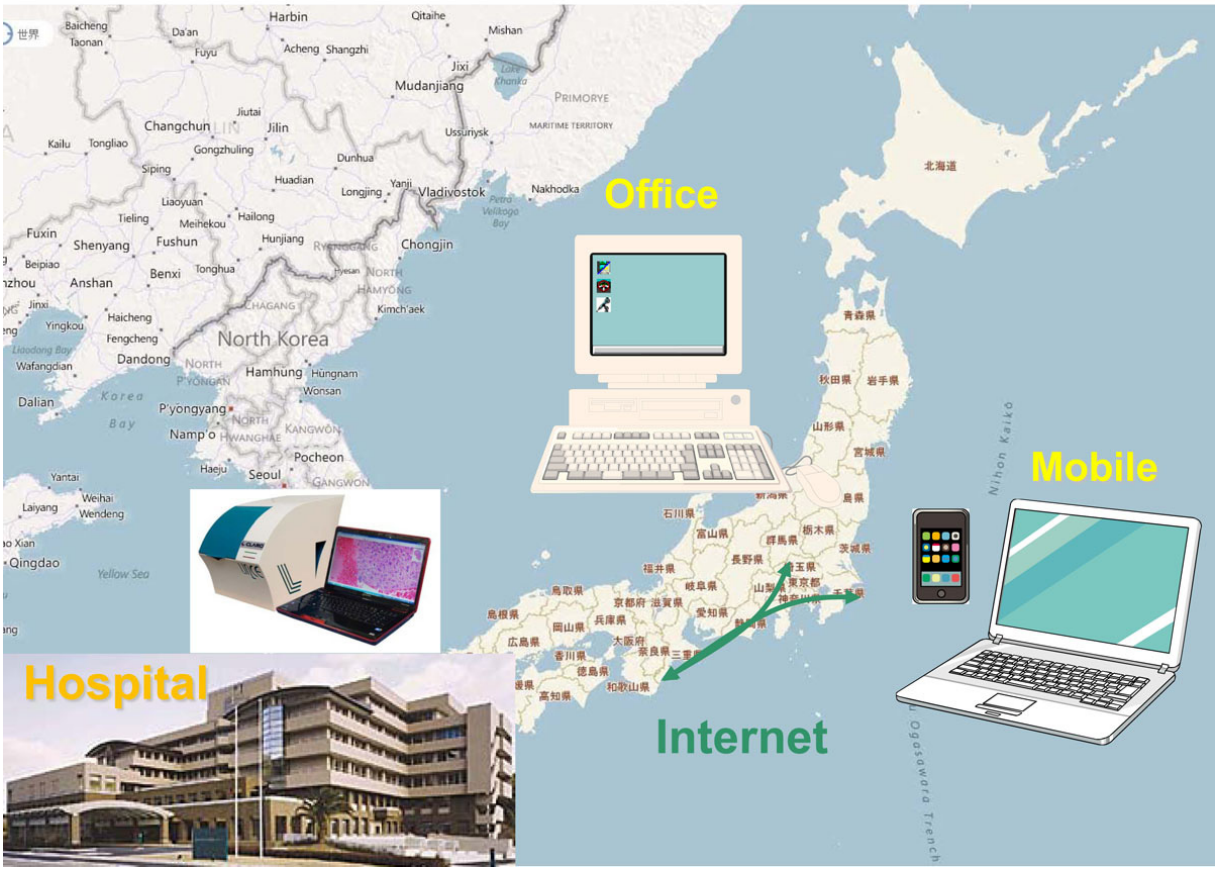



**Figure 1(abstract S4) VS telepathology of routine HE slides.** We made primary diagnosis of VS of routine HE slides of a hospital 500 km away from Tokyo through internet.

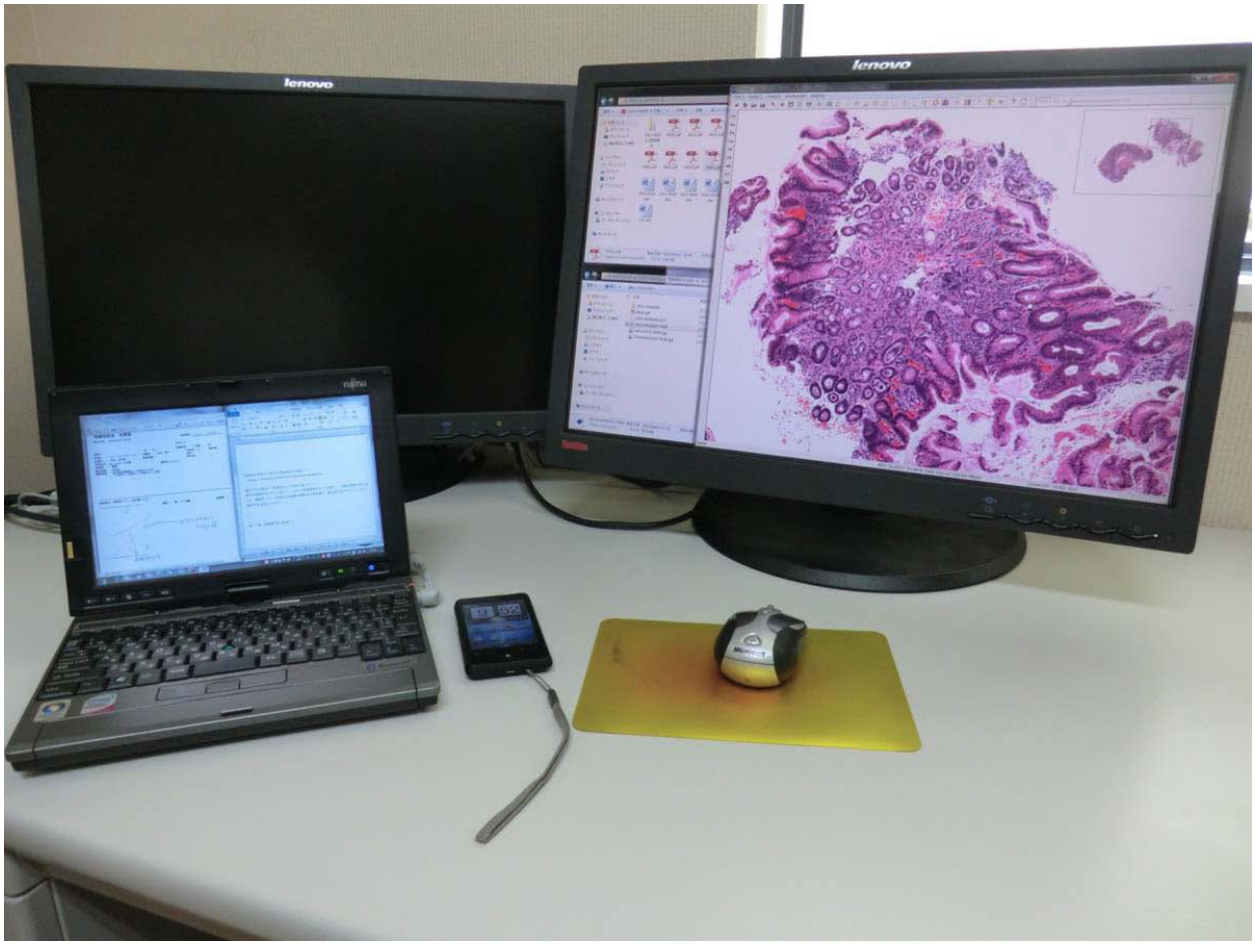



**Figure 2(abstract S4) Actual working desk of remote diagnosis** Notebook PC was connected to the web through mobile telephone network. Another display is also connected to the notebook PC, and diagnosis was done under dual display condition.

diagnosis. We constructed this system as a kind of emergency response for the absence of a pathologist, and it proved to be useful. There still many issues that remain, like the way to write diagnosis report to the system, clarification of where responsibility lies, and how to handle the large operation specimens or cytology slides, etc. In Japan, we have a guideline for intraoperative frozen section telepathology diagnosis written by the Japanese Research Society of Telepathology and Pathology Informatics [6]. We may require a guideline for web-based remote primary diagnosis using VS.

**Competing interests:** The authors declare that they have no competing interest.

**Authors' contributions:** IM did the most remote diagnosis. TO actually visit the hospital, look slides by microscope, and made many discussion with IM. YM participate to the VS telepathology and gave many suggestions. TI and HU are the pathology technicians of the hospital who prepare pathology slides, scan to VS, and copy the diagnosis to the system. TS and YO are experienced pathologists who gave suggestion to the design of the study, and helped to draft the manuscript.

**References**

1. Mori I, Wang X, Suzuma T, Yoshimura G, Sakurai T, Kakudo K: **Assessment of telepathological frozen section diagnosis of sentinel lymph node of breast cancer.** *Virchows Archiv* 2003, **443(3)**:302.
2. Wienert S, Beil M, Saeger K, Hufnagl P, Schrader T: **Integration and acceleration of virtual microscopy as the key to successful implementation into the routine diagnostic process.** *Diagn Pathol* 2009, **4(3)**:1-8.
3. Pantanowitz L, Valenstein PN, Evans AJ, Kaplan KJ, Pfeifer JD, Wilbur DC, Collins LC, Colgan TJ: **Review of the current state of whole slide imaging in pathology.** *J Pathol Inform* 2011, **2**:36.
4. Brachtel E, Yagi Y: **Digital imaging in pathology - current applications and challenges.** *J Biophotonics* 2012, **5(4)**:327-35.
5. Isaacs M, Lennerz JK, Yates S, Clermont W, Rossi J, Pfeifer JD: **Implementation of whole slide imaging in surgical pathology: A value added approach.** *J Pathol Inform* 2011, **2-39.**
6. Tsuchihashi Y, Sawai T: **Establishing Guidelines for Practical Telepathology in Japan.** *Gan No Rinsho* 2009, **51(9)**:721-725.

## S5

### Cell nuclei extraction from breast cancer histopathologyimages using colour, texture, scale and shape information

Antoine Veillard[1,2*], Maria S Kulikova[1], Daniel Racoceanu[1,2]
[1]IPAL– UMI CNRS, 1 Fusionopolis Way, #21-01 Connexis, 138632 Singapore; [2]School of Computing, National University of Singapore, 117417 Singapore
E-mail: veillard@comp.nus.edu.sg
*Diagnostic Pathology* 2013, **8(Suppl 1):**S5

**Background:** Standard cancer diagnosis and prognosis procedures such as the Nottingham Grading System for breast cancer incorporate a criterion based on cell morphology known as cytonuclear atypia. Therefore, algorithms able to precisely extract the cell nuclei are a requirement in computer-aided diagnosis applications.

However, unlike other modalities such as needle aspiration biopsy images, H&E stained surgical breast cancer slides are a particularly challenging image modality due to the heterogeneity of both the objects and background, low object-background contrast and frequent overlaps as illustrated in Figure 1. As a consequence, existing extraction methods which are largely reliant on color intensities do not perform well on such images.

**Materials and methods:** We propose a method based on the creation of a new image modality consisting in a grayscale map where the value of each pixel indicates its probability of belonging to a cell nuclei. This probability map is calculated from texture and scale information in addition to simple pixel color intensities. The resulting modality has a strong object-background contrast and evens out the irregularities within the nuclei or the background. The actual extraction is performed using an AC model with a nuclei shape prior included to deal with overlapping nuclei.

**Feature model:** First, a color deconvolution [1] is applied in order to separate the immunohistochemical stains from which 3 grayscale images are produced: a haematoxilin image, an eosin image and a third residual component orthogonal in RGB space. Next, local features based on Laws' texture measures [2] are computed for each pixel of the 3 obtained images. 5 different 1-dimensional convolution kernels ($L_5$ = (*1*, *4*, *6*, *4*, *1*), $E_5$ = (*–1*, *–2*, *0*, *2*, *1*), $W_5$ = (*–1*, *2*, *0*, *–2*, *1*), $S_5$ = (*–1*, *0*, *2*, *0*, *–1*) and $R_5$ = (*1*, *–4*, *6*, *–4*, *1*)) are used to compute 25 different 5 × 5 kernels by convolving avertical 1-dimensional kernel with a horizontal one. The 5 × 5 kernels are applied at every pixel to extract 25 features which are then combined into 15 rotationally invariant features after normalizing by the output of the $L_5^T \times L_5$ kernel and smoothing with a Gaussian kernel of standard deviation $\sigma = 1.5$ pixels.

The same process is repeated at 4 different scales after locally re-sampling the image using Lanczos-3 sinc kernels. Re-sampled images are locally computed around each pixel to allow the computation of the 15 texture features for the same pixel at different scales. Local texture features are computed at 1:1, 1:2, 1:4 and 1:8 scales for every pixel.

**Probability map:** The resulting 180-dimensional feature vector $x$ is used to compute the probability $p_n(x)$ of each pixel to belong to a cell nuclei. Let $\mu_n$ (resp. $\mu_b$) be the mean of the feature vectors for the pixels belonging to the nuclei (resp. to the background). A class dependent LDA is performed in order to find two directions in the feature space, $w_n$ and $w_b$, such that the projection of the classes on these directions has a maximum inter-class scatter over within-class scatter ratio. The estimated class probability associated with the feature vector x is then calculated from the linear scores $l_n = (x - \mu_n) \cdot w_n$ and $l_b = (x - \mu_b) \cdot w_b$ using the softmax function:

$$p_n(x) = \frac{\exp(l_n(x))}{\exp(l_n(x)) + \exp(l_b(x))}$$

The resulting probability map exhibits strong contrast between the objects and the background. Moreover, nuclei and background appear more homogeneous than in the original image. A post processing step is also applied to fill small holes still remaining in nuclei (larger holes are not removed to prevent the unintended deletion of interstices between different nuclei).

**AC model including shape prior:** The actual extraction of cell nuclei is performed from the probability map with an AC model with shape prior information. The total energy $E(\gamma)$ associated to a contour $\gamma$ is a weighted sum of an image term $E_i(\gamma)$ and a shape term $E_s(\gamma)$. The latter is itself the weighted sum of a smoothing term $E_{sm}(\gamma)$ and a shape prior term $E_{sp}(\gamma)$.

The shape prior term $E_{sp}(\gamma) = \frac{1}{2\pi} \sum f_k \left( \int \exp(ikt) \delta r(t) dt \right)^2$ allows to control the perturbations $\delta r(t)$ of a contour around a circle at different frequencies $k$ of the Fourier components by adjusting the coefficients $f_k$. Detailed formulas and explanations for this and the other energy terms can be found in the work of Kulikova *et al.* [3]. The shape prior information allows to properly extract overlapping nuclei according to their expected shape without arbitrarily discarding the overlapping parts.

The detection of nuclei is performed by a marked point process model the details about which the interested reader can find in [4]. An empirical study in [5] shows that this particular combination of MPP and AC over-performs other state-of-the-art methods for nuclei detection and extraction.

**Results and discussion:** The training set used for the LDA consists of 6 1024×1024 images where the nuclei have been manually delineated by a pathologist. Object and background parameters used in the AC model

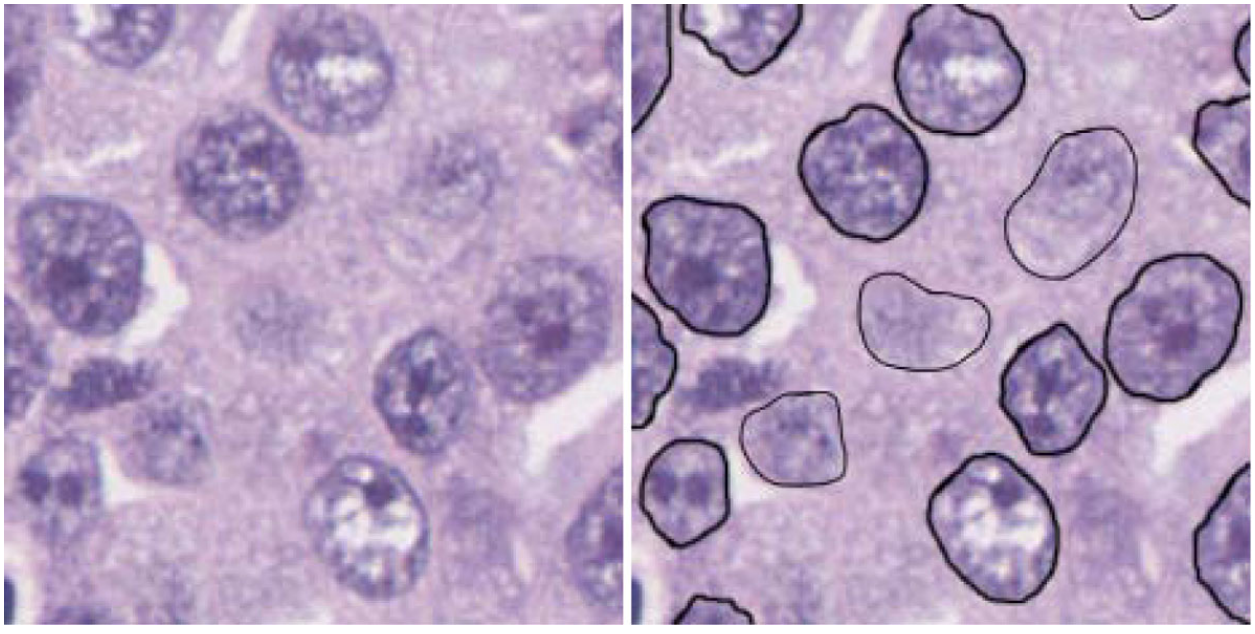


**Figure 1(abstract S5) Ground truth** (Left) magnified 250 × 250 region of a frame and (right) the same region with the nuclei delineated by a pathologist. Nuclei delineated with a thinner outline are hard to distinguish from the background. Dark and bright areas can indiscriminately occur inside and outside nuclei.

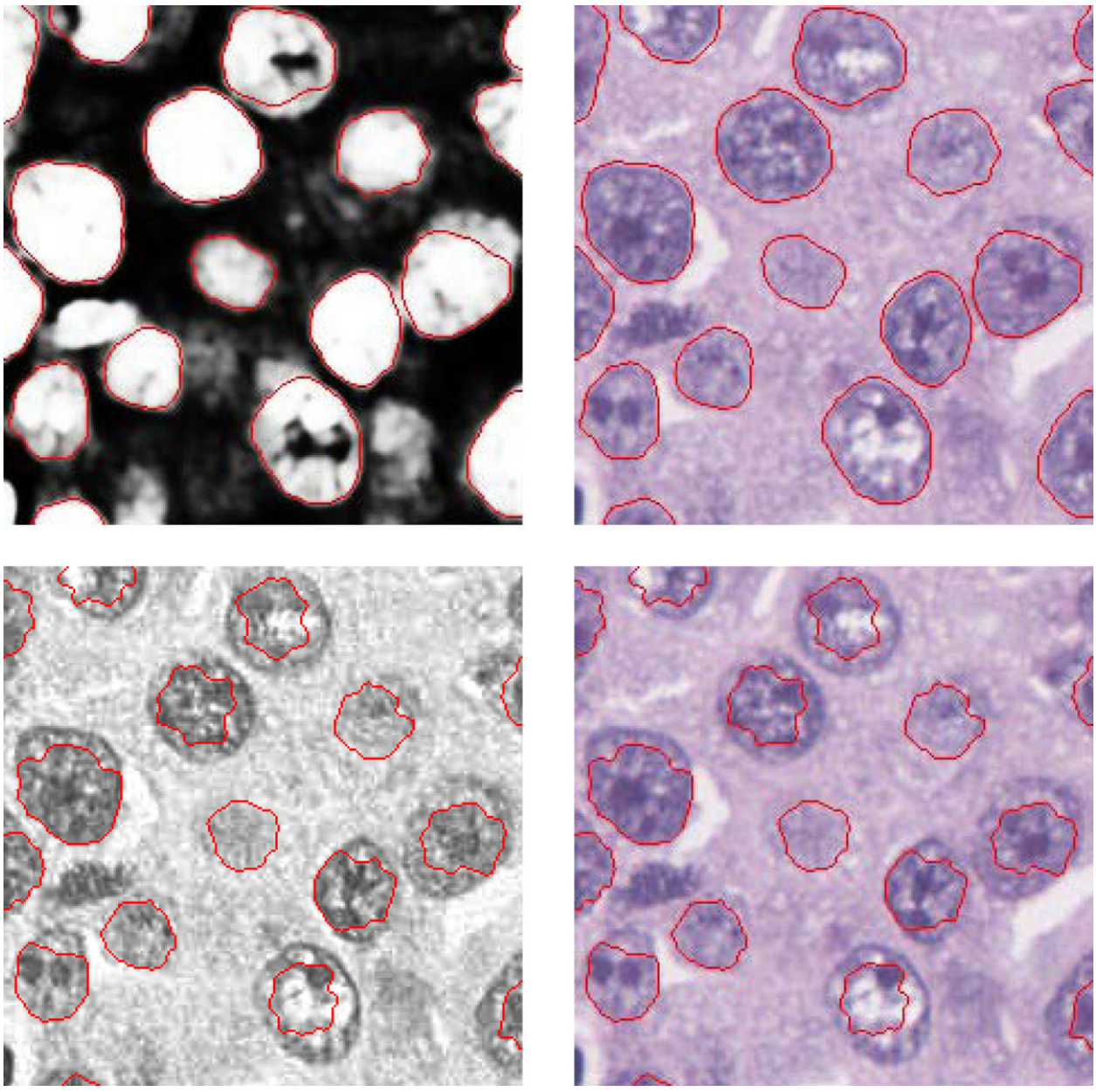


**Figure 2(abstract S5) Extraction results** The top row is obtained using the probability map and the bottom row is obtained using the haematoxilin channel after the image color deconvolution.

are also calculated from the training set. Weight parameters for the energy terms in the AC model are adjusted with a grid search. Images used for training are distinct from the images used for validation.

Figure 2 shows results obtained with the AC model applied to the probability map side-by-side with results obtained with the same AC model applied to the original image (in fact, the slightly better performing haematoxilin image from the color deconvolution was used instead of the red channel from the RGB image commonly used in other methods [6]). On the original image, the contours have a tendency to match irregularities within the cell nuclei rather than their actual boundaries. This problem is largely improved by using the probability map where the nuclei boundaries are much more salient and other irrelevant features are smoothed out.

**References**

1. Ruifrok AC, Johnston DA: **Quantification of histochemical staining by color deconvolution.** *Analytical and Quantitative Cytology and Histology* 2001, **23**:291-299.
2. Laws K: **Textured image segmentation.** *PhD thesis* University of Southern California 1980.
3. Kulikova M, Jermyn I, Descombes X, Zhizhina E, Zerubia J: **A marked point process model with strong prior shape information for extraction of multiple, arbitrarily-shaped objects.** *Proc. Signal-Image Technology and Internet-Based Systems* 2009.
4. Descombes X, Minlos R, Zhizhina E: **Object extractionusing a stochastic birth-and-death dynamics in continuum.** *J. Math. Imaging Vis* 2009, **33(3)**:347-359.

5. Kulikova MS, Veillard A, Roux L, Racoceanu D: **Nuclei extraction from histopathological images using a marked point process approach.** *Proc. SPIE Medical Imaging, San Diego, California, USA, 2012.*
6. Dalle J, Li H, Huang CH, Leow W, Racoceanu D, Putti TC: **Nuclear pleomorphism scoring by selective cell nuclei detection.** *Proc. Workshop on Applications of Computer Vision 2009.*

## S6

### Development of a nationwide telepathology consultation and quality control program for cancer diagnosis in China

Zhongjiu Zhang[1*], Guangming Gao[2], Chen Zhou[3]
[1]Bureau of Health Policy and Regulation, Ministry of Health, Beijing, China; [2]Division of Quality of Healthcare, Bureau of Health Policy and Regulation, Ministry of Health, Beijing, China; [3]British Columbia Cancer Agency, University of British Columbia, Vancouver, Canada
*Diagnostic Pathology* 2013, **8(Suppl 1):**S6

**Background:** China has the largest numbers of cancer patients worldwide but has limited pathology resources. According to a survey done by Chinese Pathologists Association, a large number of pathologists in clinical practice in China has only one year of formal or informal training. Only in recent year, formal pathology residency programs similar to those of Western countries are available in a couple of large medical schools, where limited numbers of well-trained and experienced pathologists are located. When having rare or complex cancer pathology cases, pathologists often have difficulty in making an accurate diagnosis. Mistakes in cancer diagnosis are not uncommon, which often leads to inadequate treatment of patients and medical legal problems. Thus, there is a great need for pathology consultation and quality control of cancer diagnosis in China.

In recent years, telepathology or digital pathology has been applied in many areas of pathology, including remote consultation and quality control [1]. Difficult pathology cases can be sent through internet or intranet to remote site for consultation [2]. Quality control can be performed on digitalized image or WSI by expert pathologists at remote sites [3]. The practice of telepathology can solve the problem of healthcare system with poor pathology resource, especially in developing countries [4]. In order to use pathology resource efficiently and to improve the quality of pathology diagnosis of cancer, the Ministry of Health of China developed and launched a nationwide telepathology consultation and quality control program for cancer diagnosis in China in 2011.

**Designs: Internet platform:** An internet based telepathology consultation platform (http://www.mpathology.cn) was designed to serve as the hub for the project, connecting hospitals and expert consultants. Selection an internet based platform or a website portal was due to the consideration that multiple hospitals/institutions and multiple expert pathologists from different institutions need to access the platform. A server was used for storage of images and data. Figure 1 shows the flowchart of the platform.

For telepathology consultation, referring pathologists from participating hospitals sent request for consults with attached WSI and related clinical information to the internet platform, the system would alert system information technologists by e-mails. The technologists would then check out the submitted materials to make sure the requesting form is filled out appropriately with the WSI attached. The system manager would then contact the expert pathologist for consultation.

For quality control, a designated expert pathologist would log into the platform, review the WSI of cancer case from specific hospitals, assesses the accuracy of the pathology diagnosis and quality of histology sections. The results would then be sent to provincial and national pathology quality control center.

All of the participating hospitals were supplied with a virtual microscope, Motic Virtual Microscopic Scanner (Motic Medical Diagnostic Systems Corporation, China). The virtual microscope and related software were tested and validated previously in a study involving a variety of 600 surgical pathology specimens in China. The study showed a high concordance rate (94.2%) between WSI and H/E glass slide [5].

**Expert consultants:** 356 pathologists with provincial and national reputation volunteered to participate in the program. 80 of them were selected to serve as expert consultants after passing an online assessment consists of 30 pathology consultation cases in WSI format. The names, affiliation of the pathologists and the areas of their subspecialty expertise were listed on the website (http://www.mpathology.cn). Experts were provided with a unique identifier and access code to the website. For telepathology consultation, the expert was instantly informed by the cell phone message and e-mail once he/she was requested for consult. The expert pathologist would use a computer or an Ipad to log into the website, check out the pending cases for consultation under his/her name, review WSI of the case, capture a representative image from WSI to included in the final report, type in pathologic features of the case, write up his/her opinion and provide a final diagnosis. After that, he/she would preview the final report, then sign the report with an electronic signature and then release the final report.

Once the final consultation report was released, the system sent an email to alert the system manager. The final consultation report was then sent by the system manger by e-mail or fax to the referring pathologist.

**Participating hospitals:** 87 hospitals in 17 provinces where national healthcare reform was carried out volunteered to participate the phase I of the program, all of the hospitals had more than 6000 pathology cases annually. Pathologists in those hospitals took an online diagnostic test of 50 cases of cancer. Those hospitals having pathologists with excellent test results were eliminated from the program. 60 hospitals were finally selected to participate in the phase I of the program. Participating hospitals were given a unique identification code and user name to login to the website platform. Pathologists of the participating hospitals are required to receive a short period of training, which includes operation of the virtual microscope and internet connection to the platform. Storage and security of the WSI, relevant clinical data, and consultation files are also to be addressed.

For telepathology consultation, each participating hospital is required to send WSI of at least 300 cases annually for second opinion. The number of cases is set at about 15% to 20% of cancer diagnosed each year in those hospitals. Hospitals can choose pathologists for second opinion or consultation from 80 expert consultants with different subspecialty expertise. For quality control of pathology diagnosis, the participating hospitals are required to submit WSI of every cancer diagnosed to the platform. Quality control is to be carried out by expert consultants every 3 month on 10% randomly selected cases. The accuracy of cancer diagnosis and quality of tissue section are to be reviewed, assessed, and submitted to provincial and national pathology quality control centers. The results of the phase I of the program are to be used to guide future deployment of the program across the country.

**Discussion and conclusion:** Since 2009, due to the wide spread use of high speed internet and availability of low cost virtual microscopes in China, telepathology has gained attention from academic institutions and government. In 2010, the Ministry of Health of China released an announcement encouraging hospitals to use telepathology for cancer diagnosis and planned to develop a nationwide telepathology consultation service and quality control program. The plan is to set up in each province a provincial telepathology consultation and pathology quality control network, which also connects to a central national network. The goal of the plan is to facilitate remote consultation for cancer diagnosis and to monitor the accuracy of pathological diagnosis of cancer. In 2011, the phase I of the plan consisting of 60 hospitals in 17 provinces was completed.

In China, we do not know the exact number of error in pathological diagnosis of cancer, however, we believe the rate is much higher than those reported in Western countries, as many pathologists in China are not as well trained as those in Western countries and many pathologists in China do not have subspecialty expertise. The project we design asks participating hospitals to submit all of their newly diagnosed cancer cases. 10% of those cases will be randomly selected for review. The 10% is set arbitrarily and is not a high percentage for newly diagnosed cancer cases, in comparison to prospectively secondly review of all newly diagnosed cancer cases or retrospective mandatory review of all cancer cases by some institutions in North America.

Second opinion or consultation is extremely important in pathology diagnosis, especially in cancer diagnosis. In a retrospectively review of consultation cases of urological malignancy, Wayment *et al.* [6] reported that a 10% disagreement, of which 8% is major. Matasar *et al.* [7] review the difference between second opinion and submitting diagnosis in lymphoma cases and found that major diagnostic revision was 17.8% in 2001 and 16.4% in 2006. In Asia or developing countries, the discrepancy

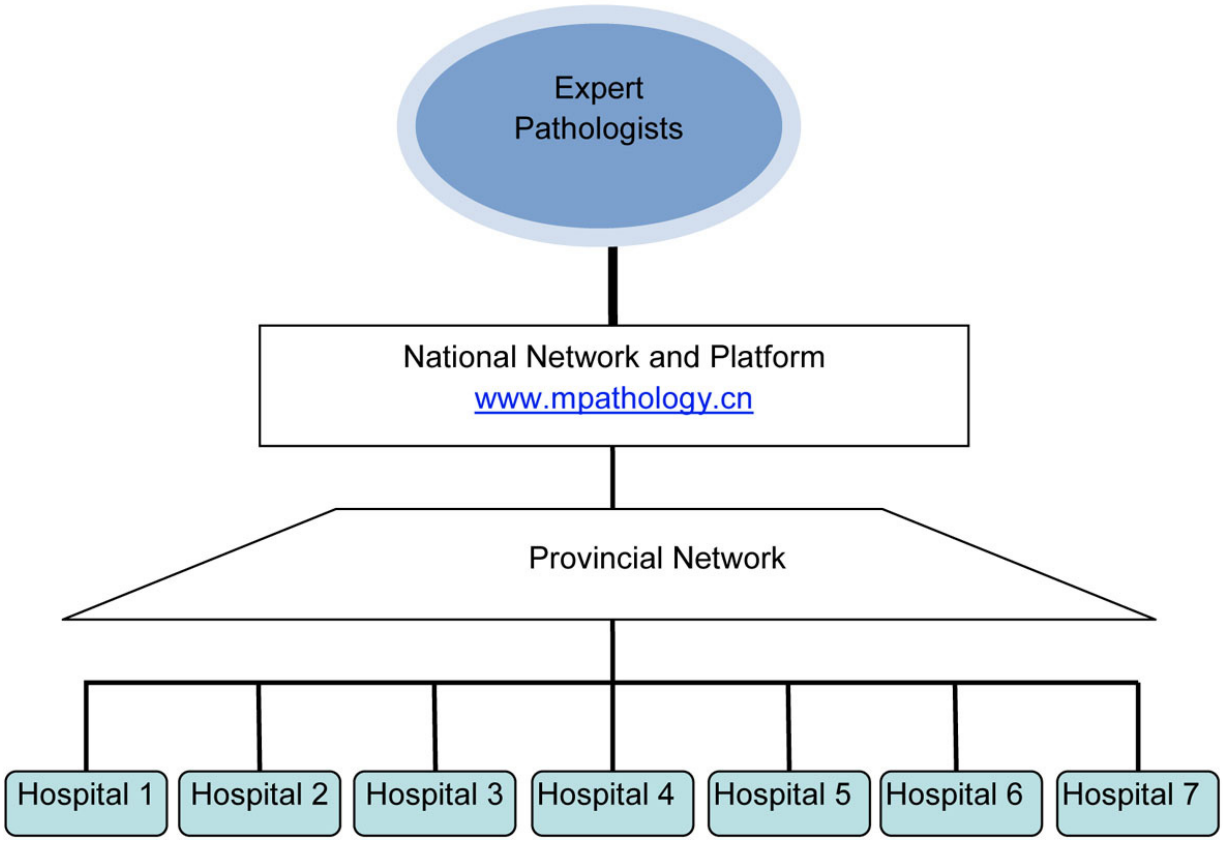



**Figure 1(abstract S6) Nation-wide Telepathology consultation and quality control network.** Figure 1 shows the flowchart of the internet platform and nation wide network. Participating hospitals send digital images or WSI of the histology slides of cancer through the internet to the platform, WSI is then stored in the server; expert pathologists from different institutions could access the website through the internet, review the WSI stored in the server and deliver a second opinion or make an assessment of quality of pathology diagnosis.

rate is similar or much higher. In a review of 673 consultation cases sent to a cancer center in Twain, Tsung [8] found a 16% of major disagreement between original diagnosis and second opinion. Hsu *et al.* [9] reported that among 2686 consultation cases, the tentative diagnosis and consultation diagnosis were discordant in 1,074 (64.3%) cases. Major discrepancy was seen in 205 (12.3%) cases, of which 66.8% were changed from malignant to benign, 21.0% were changed from benign to malignant. Based on these studies, the program we designed asks participating hospitals to submit 10-20% of their cases for consultation.

Through this program, we believe that in near future telepathology will be widely used in China for pathology consultation and quality control. The program will not only assists pathologists in dealing with diagnostically difficult pathology cases, but also benefits patients, who do not need to travel a long distance to large hospitals for a second opinion of cancer diagnosis, thus saving time and money for patients. The program will also guaranty the quality of pathology diagnosis of cancer in China.

**References**

1. Rocha R, Vassallo J, Soares F, Miller K, Gobbi H: **Digital slides: present status of a tool for consultation, teaching, and quality control in pathology.** *Pathol Res Pract* 2009, **205(11)**:735-41.
2. Wilbur DC, Madi K, Colvin RB, Duncan LM, Faquin WC, Ferry JA, *et al*: **Whole-slide imaging digital pathology as a platform for teleconsultation: a pilot study using paired subspecialist correlations.** 2010.
3. Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR: **Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies.** *Human pathology* 2006, **37(3)**:322-31.
4. Hitchcock CL: **The Future of Telepathology for the Developing World.** *Archives of pathology & laboratory medicine* 2011, **135(2)**:211-4.
5. Li X, Gong E, McNutt MA, Liu J, Li F, Li T, *et al*: **Assessment of diagnostic accuracy and feasibility of dynamic telepathology in China.** *Human pathology* 2008, **39(2)**:236-42.
6. Wayment RO, Bourne A, Kay P, Tarter TH: **Second opinion pathology in tertiary care of patients with urologic malignancies.** *Urol Oncol* 2011, **29(2)**:194-8.
7. Matasar MJ, Shi W, Silberstien J, Lin O, Busam KJ, Teruya-Feldstein J, *et al*: **Expert second-opinion pathology review of lymphoma in the era of the World Health Organization classification.** *Ann Oncol* 2011.
8. Tsung JSH: **Institutional pathology consultation.** *The American journal of surgical pathology* 2004, **28(3)**:399.
9. Hsu CY, Su IJ, Lin MC, Kuo TT, Jung SM, Ho DM: **Extra-departmental anatomic pathology expert consultation inTaiwan: a research grant supported 4-year experience.** *J Surg Oncol* 2010, **101(5)**:430-5.

## S7

### Application and evaluation of teaching practical histology with the use of virtual microscopy

Eva Pospíšilová[†], Drahomíra Černochová[†], Radka Lichnovská[†], Běla Erdösová, Dimitrolos Krajčí[*†]
Department of Histology and Embryology, Faculty of Medicine and Dentistry, Palacký University, Hněvotínská 3, 77515 Olomouc, Czech Republic
E-mail: dimikra@gmail.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S7

**Background:** Emerging digital microscopy technology and newly developed scanning light microscopy systems enable histologists to transfer analog image data of entire slides into digital ones, and to

provide them to students in the class with identical contents and quality [1,2]. We have recently started a major innovation of our system of practical teaching using e-learning methods of delivery of high quality histology virtual slides and supporting documents [3]. We have created our own teaching environment in a computer-equipped and networked classroom dedicated to histology practical sessions. This method of teaching has been complemented with electronic in-course and final assessments of student's theoretical and practical knowledge and skills.

**Material and methods:** Glass histology slides have been scanned with Olympus dotSlide scanning system creating thousands of overlapping images saved in a proprietary file format. The scanning was done with standard 40x objective lens. Using dedicated viewer software (Olyvia, Olympus) the images were viewed on student's PC as a single image map at variable magnifications. Our classroom, dedicated to histology practical sessions, consists of 1 server PC (teacher) and 30 client's PC (students) [3]. The server PC has four accounts set up. Administrator account - with full administrator's rights, Lector account - with partially limited power user's rights. Student account and Exam account – with limited user's rights, enable remote access of student's workstations to the server. Under lector's login, the supporting documents and virtual slides can be exchanged, edited, added or deleted for incremental updates.

The student's PCs are standard ultraslim units running on Windows XP, SP3 operating system, under strict administrator's control with Microsoft Steady State software. Client computers have also Administrator, Lector, Student, and Exam accounts set up. Students can access only the last two accounts without any login password. When opening the Student account, they are presented with a welcome screen where they can select their preferable language of instruction for practical class (Czech or English).

The core of this virtual slide learning system is our own Database of Histology Practical saved on the X:\drive of the server PC. It is build up in MS Excel format having the first contents page, followed by practical pages devoted to all 24 topics of histology practical [3]. For electronic testing of student's practical knowledge we used special software Quizmaker '09 (Articulate) which is simple and easy to use by teachers having no programming skills. Since this software has a selective option to shuffle sequences of questions and also to shuffle all distracters in the quiz randomly on monitors of examined students, only one version of histology test was enough to prepare for one practical class. A time limit for display of each question and a total time allowed for a complete test were also settable.

**Results and discussion: Evaluation of practical sessions with virtual slides:** Students of histology practical classes readily accepted the use of computers for observation of virtual slides. Using our own specific set of evaluation questions, we have asked students to evaluate this new method of practical sessions. Majority of students in General Medicine and Dentistry specializations (93%) evaluated positively the use of virtual slides, as they allowed them to study and also to discuss various details of cells and tissues clearly at various magnifications. About half of students (58%) claimed that they benefited from using the attached supporting documents during practical sessions and almost all students (96%) downloaded these supporting documents to their external media for later self-revisions. The classical light microscopes and glass slides, that were available next to computers, were used by histology students occasionally, according to their reply.

Teachers benefited from a uniform quality of presented slides and also from a straightforward and easy personal communication with students in the class when personal guidance and explanation was needed at student's monitors. This was a highly beneficial feature when this system was used in large practical labs of 50 and more students participating simultaneously. PC-based classes of practical histology also provided an easy environment for computerized testing of student's practical knowledge of structures displayed on their monitors.

**Virtual slides versus classical light microscopy:** With the advent of virtual microscopy, the format of histology practical labs is changing towards the use of virtual slides viewed by individual students on their PCs. In the earlier-equipped histology practical labs the availability of both classical light microscopes and virtual slides is maintained up to date in order to give students a possibility to observe histology slides in both ways. According to survey of Drake et al. 2009 [4] regarding the laboratory experience of the 45 respondents, 13 reported that their laboratory used classical microscopes, 20 reported that their laboratory used virtual microscopy only, and 12 reported that their laboratory used a combination of microscopes and virtual microscopy. The authors conclude that histology lends itself to approaches that are more independent study friendly. This is especially true with the continuously increasing usage of virtual microscopy systems that students can access anywhere by computer.

**Conclusions:** The e-learning format of histology practical based on virtual slides is a didactically efficient method of teaching histology. It standardizes the set of histology slides and gives all students in a large teaching group equal opportunity to see the same high quality slides in practical sessions. It also gives students a new experience with the observation of histology structures on PC monitors, and enables them to take selected screen prints of virtual slides for their own study/research observations as if they would be using a complete research microscope with a digital camera. Students evaluate positively their use of virtual slides and accompanying supporting documents in histology practical sessions. Teachers benefit from a uniform quality of presented slides and also from a straightforward and easy didactic communication with students in the class.

**List of abbreviations:** PC: personal computer; MCQ: multiple choice question

**Competing interests:** The authors declare that they have no competing interests in relation to this work.

**Authors' contributions:** EP, DČ and RL participated in selection of glass slides for scanning, writing Czech versions of supporting documents, editing data in the Histology Practical Database, application and evaluation of virtual slides in practical sessions, and preparing MCQs for examinations. BE participated in editing English versions of supporting documents, application and evaluation of virtual slides in practical sessions, and preparing English versions of MCQs for examinations. DK conceived this project, planned and coordinated all works, scanned virtual slides, designed the Histology Practical Database, wrote English versions of supporting documents and trained the collaborating staff in use of the special software. All authors read and approved the final manuscript.

**Acknowledgements:** This project has been supported by European Social Fund, and Ministry of Education, Youth and Sports of the Czech Republic, grant no. CZ.1.07/2.2.00/28.0089 http://virtual-histology.upol.cz/o-projektu/informace-o-projektu.html

**References**

1. Gu Jiang, Ogilvie RW: **Virtual Microscopy and Virtual Slides in Teaching, Diagnosis, and Research.** *Advances in Pathology, Microscopy & Molecular Morphology* Taylor & Francis Routledge, London: Ogilvie Robert W and Jiang Gu 2005 [http://www.crcnetbase.com/isbn/9780849320675].
2. Romero E, Gómez F, Iregui M: **Virtual Microscopy in Medical Images: a Survey.** *Modern Research and Educational Topics in Microscopy* A. Méndes-Vilas and J. Díaz, Formatex 2007.
3. Krajčí D, Pospíšilová E, Černochová D, Kopečný T, Pop A: **Histology practical in electronic format of virtual slides.** *Mefanet report 04. Efficient multimedia teaching tools in medical education* Masaryk University, Brno, Czech Republic: D. Schwarz et al ISBN 978-80-210-5539 2011, 60-68.
4. Drake RL, McBride JM, Lachman N, Pawlina W: **Medical Education in the Anatomical Sciences: The Winds of Change Continue to Blow.** *Anat Sci Educ* 2009, **2**:253-259.

## S8

### The Eastern Quebec Telepathology Network: a support to the improvement to the public health care system

Bernard Têtu[1*], Marie-Pierre Gagnon[2], Geneviève Roch[3], Jean-Paul Fortin[4]
[1]Professor of pathology and Medical Director of The Eastern Quebec Telepathology Network; Departement of Pathology, Laval University, Québec, Canada; [2]Associate professor, Faculty of Nursing Sciences, Laval University, Québec, Canada; [3]Adjunct professor, Faculty of Nursing Sciences, Laval University, Québec, Canada; [4]Professor, Department of Social and Preventive Medicine, Laval University, Québec, Canada
E-mail: bernard.tetu@fmed.ulaval.ca
*Diagnostic Pathology* 2013, **8(Suppl 1):**S8

**Summary:** The Eastern Quebec Telepathology Network is aimed at providing uniform diagnostic telepathology services in a huge territory with a low population density. It has been designed to provide intraoperative

consultations (frozen sections) in smaller community hospitals and to allow pathologists working alone to rapidly obtain a second opinion. This study provides an interim evaluation of the benefits of the network. The network involves 24 sites, of which seven are devoid of a pathology laboratory, three have no pathologist and six have one pathologist on site. Since the beginning of the implementation, the coverage in pathology in this territory was improved: 1) telepathology allowed pathologists to work outside of their office and to provide surgeons with faster diagnoses of urgent biopsies; 2) expected interruptions of the intraoperative consultation coverage were avoided; 3) intraoperative consultations were provided to surgeons in hospitals devoid of pathology laboratory; 4) expert opinions were obtained with reduced isolation for pathologists working alone and improved turn-around time; 5) merging of smaller laboratories resulted in a more stable pathology coverage and an attractive effect on the recruitment of young pathologists; 6) videoconferencing and macroscopy station allowed real-time communication between a pathologist and the remote technician for macroscopic description; and 7) several technical procedures were standardized (staining, sectioning, reporting). In conclusion, the Eastern Quebec Telepathology Network was designed to improve medical care to patients. In a short period of time, an improvement of the organization of health cares and of the delivery of services is already apparent.

**Background:** Canada is a huge country with a low population density. The province of Quebec is the largest but the second most populated province with its 7,957,600 inhabitants. The territory covered by the Eastern Quebec Telepathology Network spans over 452,600 $km^2$ in which 1,7 million inhabitants live (figure 1). In certain areas, the density is as low as 0.4 inhabitants/$km^2$.

In 2004, a telepathology project team was asked to perform a survey exploring and evaluating specific needs for telepathology in Eastern Quebec. This survey revealed that, due to the lack of consistent pathology coverage in several smaller community hospitals, certain surgeries were required to be postponed, several patients were transferred to regional hospitals and two-step surgeries had to be performed when an intraoperative consultation was needed and no pathologist was available on site. This situation also proved to be a major limitation for the recruitment of young surgeons trained in centers where access to an expert pathologist is never an issue. Furthermore, it was clear that younger pathologists in early practice felt insecure and were often reluctant to work alone because of the difficulty in rapidly obtaining a second opinion from a colleague. Practicing pathologists complained that they could hardly be absent without disturbing the surgical unit. Recent literature indicates that telepathology may be instrumental in reforming the health care system and improving the quality of care [1]. In this context, the Eastern Quebec Telepathology Network has been created and was aimed at improving pathology coverage in remote communities by sharing regional expertise to ensure consistent intraoperative consultation and access to expert opinions.

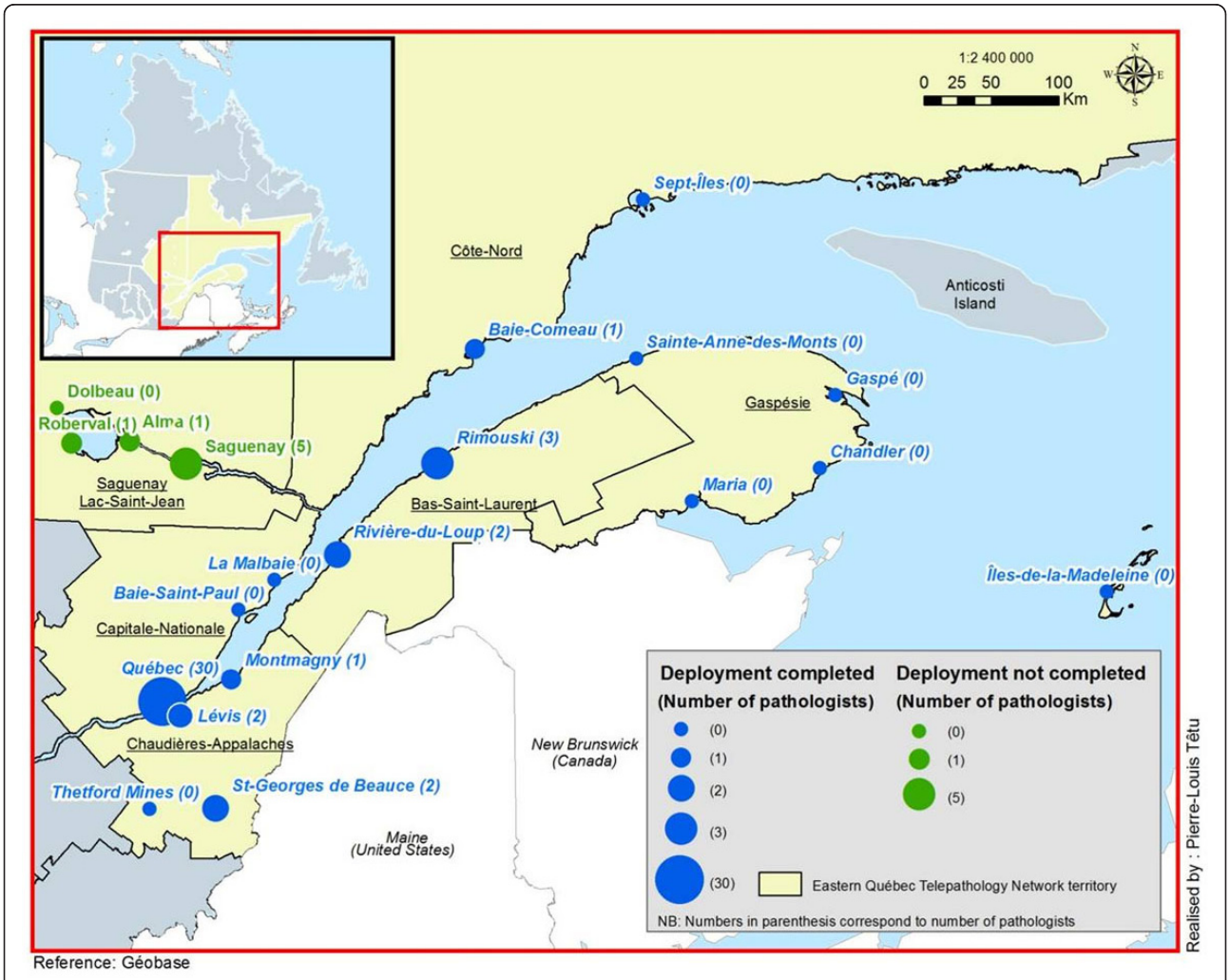



**Figure 1(abstract S8)** Territory covered by the Eastern Quebec Telepathology Network. (Adapted from figure 1 published in Médecine/Science, 2012; 28:993-9 and used with permission from the editor).

The Quebec Ministry of Health and Canada Health Infoway agreed to financially support this innovative initiative with the objective of providing the population with uniform pathology services. Deployment began in January 2010 and two years later, the implementation and training of technical and medical staff are nearly complete and the network is functional. Recently, we identified a number of challenges which, while not threatening the deployment, nevertheless required a strategy to insure that the pace of the implementation was maintained [2]. The objective of this paper is to share our experience and provide an interim evaluation of the benefits of the network as its implementation is underway.

**Methods:** Each hospital of the network is equipped with a whole slide scanner (Nanozoomers, Hamamatsu Photonics, *Shizuoka Prefecture*, *Japan*), a macroscopy station (PathStand 40, Diagnostic Instruments, *Sterling Height*, *USA*) and two videoconferencing devices (PCS-XG80DS Codec, Sony, *Minato*, *Tokyo*, *Japan*) equipped with a drawing tablet (Bamboo CTE-450K, WACOM, *Otone*, *Saitama*, *Japan*). Those equipments were obtained from Olympus Canada.Inc (*Markham*, Canada). The viewer and case management and collaboration system selected is mScope (Aurora Interactive Ltd., *Montreal*, *Canada*). All hospitals in the province of Quebec are linked together by a private, dedicated information network which ensures confidentiality and avoids use of the Internet (RITM: réseau intégré de télécommunications multimédias). The project has been implemented in a total of 24 sites, of which seven are devoid of a pathology laboratory. Of the remaining 17 sites, three have no pathologist, six have one pathologist and eight have two or more pathologists (figure 1).

Each of the 24 sites was visited by the project team prior to the deployment. The project team met the local medical (surgeons and pathologists), technical (informatics and biomedical) and administrative teams in order to better understand their needs, provide project information and identify potential locations for the equipment. A second visit was performed one year later in each center to gather invaluable information regarding the benefits and challenges met in the course of implementation.

**Results and discussion:** The Eastern Quebec Telepathology Network is currently the most ambitious telepathology project in Canada [3] and ranks among the most important in the world in terms of both the number of sites and geographic coverage. Essentially, the activities include remote intraoperative consultations (frozen sections), expert opinions between pathologists, remote assistance to macroscopic description, rapid return of immunohistochemistry and primary diagnosis on paraffin sections, including the request by surgeons of rapid diagnoses for urgent medical decision. Updated data on the use of telepathology in the Network for the period of September 2010 to January 2012 is provided in table 1.

Basically, thus far, we have found that this network encouraged more collaboration between surgeons, technologists and pathologists in a region and resulted in a better overall regional organization of medical care. More specifically, the main benefits identified to date have been:

1) Most slides were scanned for primary diagnosis (7108 slides scanned). Indeed, telepathology allowed pathologists to work outside of their office and to provide surgeons with faster diagnoses of urgent biopsies for faster planning of patients care, despite the lack of an on-site pathologist;

2) Expected interruptions of the intraoperative consultation coverage were avoided (473 slides scanned) with maintenance of local surgical and pathology technical activities. The process was intended to allow pathologists and surgeons to follow the same steps as if they were working in a single hospital. Our experience to date [2] shows that the time required is competitive with the current situation in a single hospital and compares favorably with data from the literature [4];

3) Intraoperative consultations were provided to oncologic surgeons in hospitals devoid of pathology laboratory. This is an innovative application of this technology and required proper training to technologists with limited training in histology;

4) Many expert opinions were obtained (505 slides scanned) and immunohistochemical analyses were returned more rapidly (149 slides scanned), supporting pathologists working alone and resulting in a significant reduction of the turn-around time. The system allows a referring pathologist to request an opinion from any other pathologist in the network. Indeed, it is estimated that 10 to 20% of oncologic cases must be validated by more than one pathologist [5] and certain quality assurance programs require that 10% of cases be reviewed by more than one pathologist [6]. Current literature shows that expert opinion is an area with enormous growth potential for telepathology [7];

5) The merging of smaller laboratories in a sub-region was encouraged with the result of a more stable intraoperative consultation coverage and an attractive effect on the recruitment of young pathologists in this sub-region. Prior studies by two of us (MPG, JPF) showed that telemedicine technologies may help to attract physicians to and retain them in remote regions by contributing to better working conditions [8,9];

6) The videoconferencing and macroscopy station allowed for real-time communication between a pathologist and the remote technician or pathologist assistant for macroscopic description of specimens (166 sessions) resulting in a decreased need for specimen transportation;

7) Several technical procedures such as staining, sectioning and reporting had to be standardized in different institutions working together. There is increasing literature suggesting that technical improvements and standardization enhance the quality of images by telepathology [10].

**Conclusion:** The Eastern Quebec Telepathology Network is successful and improves medical care and of the delivery of services to patients in this region. In the course of implementation, the objectives that we had settled were not only met but were also exceeded. The learning experience of such a project may be helpful to any organization intending to implement a public health/patient-oriented telepathology network.

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** All authors contributed to the development, writing and approval of the final manuscript.

**Acknowledgements:** We want to thank Mr. François Boilard, project manager of the Eastern Quebec Telepathology project, for his kind collaboration at gathering statistics about telepathology.

**Table 1(abstract S8) Analyses performed by telepathology between September 2010 and June 2012**

| Analysis | Number of slides scanned |
|---|---|
| Primary diagnosis, including urgent interpretation | 7108 |
| Intraoperative consultations (frozen sections) | 473 |
| Expert opinions between pathologists | 505 |
| Assistance to macroscopic description | 166 |
| Immunohistochemistry | 149 |

**References**

1. Bashshur RL, Shannon GW, Krupinski EA, Grigsby J, Kvedar JC, Weinstein RS, Sanders JH, Rheuban KS, Nesbitt TS, Alverson DC, Merrell RC, Linkous JD, Ferguson AS, Waters RJ, Stachura ME, Ellis DG, Antoniotti NM, Johnston B, Doarn CR, Yellowlees P, Normandin S, Tracy J: **National telemedicine initiatives: essential to healthcare reform.** *Telemed J E Health* 2009, **15**:600-10.
2. Tetu B, Fortin JP, Gagnon MP, Louahlia S: **The challenges of implementing a "patient-oriented" telepathology network; the Eastern Quebec telepathology project experience.** *Anal Cell Pathol (Amst)* 2012, **35**:11-8.
3. **Quebec telepathology project aims to be the largest in Canada.** *Healthcare Technology 2010* 2010, **15(7)**:13-15.
4. Evans AJ, Chetty R, Clarke BA, Croul S, Ghazarian DM, Kiehl TR, Perez Ordonez B, Ilaalagan S, Asa SL: **Primary frozen section diagnosis by robotic microscopy and virtual slide telepathology: the University Health Network experience.** *Hum Pathol* 2009, **40**:1070-81.
5. Dietel M, Nguyen-Dobinsky TN, Hufnagl P: **The UICC Telepathology Consultation Center. International Union Against Cancer. A global approach to improving consultation for pathologists in cancer diagnosis.** *Cancer* 2000, **89**:187-91.
6. Nakhleh RE, Bekeris LG, Souers RJ, Meier FA, Tworek JA: **Surgical pathology case reviews before sign-out: a College of American Pathologists Q-Probes study of 45 laboratories.** *Arch Pathol Lab Med* 2010, **134**:740-3.
7. Weinstein RS, Graham AR, Lian F, Braunhut BL, Barker GR, Krupinski EA, Bhattacharyya AK: **Reconciliation of diverse telepathology system designs. Historic issues and implications for emerging markets and new applications.** *APMIS* 2012, **120**:256-75.

8. Duplantie J, Gagnon MP, Fortin JP, Landry R: **Telehealth and the recruitment and retention of physicians in rural and remote regions: a Delphi study.** *Can J Rural Med* 2007, **12**:30-6.
9. Gagnon MP, Duplantie J, Fortin JP, Landry R: **Exploring the effects of telehealth on medical human resources supply: a qualitative case study in remote regions.** *BMC Health Serv Res* 2007, **7**:6.
10. Martina JD, Simmons C, Jukic DM: **High-definition hematoxylin and eosin staining in a transition to digital pathology.** *J Pathol Inform* 2011, **2**:45.

## S9

### Out-of-sample extension of diffusion maps in a computer aided diagnosis system. Application to breast cancer virtual slide images

Belhomme Philippe[1*], Oger Myriam[2], Michels Jean-Jacques[2], Plancoulaine Benoît[1]
[1]BioTICLA-HIQ EA 4656, Université de Caen Basse-Normandie, Caen, France; [2]BioTICLA-HIQ EA 4656, CLCC François Baclesse, Caen, France
E-mail: philippe.belhomme@unicaen.fr
*Diagnostic Pathology* 2013, **8(Suppl 1):**S9

**Introduction:** While the pathologist population tends to dramatically drop, the number of pathological cases to be examined increases sharply, mainly due to early screening campaigns; developing automated systems would thus be useful to help pathologists in their daily work. As Virtual Microscopy (VM) is more and more introduced in pathology departments [1] where it holds immense potential despite the large amounts of data to be managed, its combination with image processing techniques can allow to find objective criteria for differential diagnosis or to quantify prognostic markers. Thus, many works try to develop computer-aided diagnosis systems (CADS) based on image retrieval and classification [2,3]. The first step consists in building a knowledge database involving many features extracted from a set of well-known images; it is an 'off-line' procedure conducted once. These features are represented by vectors of non-linear data acting as a signature for the original images. In a second step, signatures are obtained from new unknown images to analyze and compared with the database; it is an 'on-line' procedure. Because of tumor heterogeneity, it is essential to build knowledge databases containing representative features of the multiple morphological types of lesions before considering to implement a CADS. But, as it is almost impossible for a pathologist to manually segment large virtual slide images (VSI), the usual practice consists in manually selecting some 'representative areas'. A bias is then introduced in the process as this choice is obviously subjective. It is then mandatory to find wiser solutions leading to an unbiased collection of these 'representative areas' (and later called 'patches'). In a previous work [4], we have proposed an original strategy: starting from a collection of breast cancer VSI, then taking advantage of stereological sampling methods and diffusion maps, a knowledge database is obtained from a reduced number of patches that are representative of given histological types. The sampling tools offered by stereology are well-suited in this context [5]. Systematic sampling starting from a random point with a fixed periodic interval is able to reduce the area to be analyzed, while preserving the collection of distinctive regions encountered in a tumor. However, even if the working area becomes smaller, the number of selected patches can be very large and may include many redundant elements. A data reduction has then to be conducted. Among the available methods, the diffusion maps technique [6,7] has been retained since it provides a very attractive framework for processing and visualizing huge non-linear bulk data. Diffusion maps belongs to unsupervised learning algorithms dealing with a spectral analysis of non-linear data, providing a clustering only for given training points with no straightforward extension for out-of-sample cases. The work presented here focuses on a way to get around this problem and explains how unknown VSI can be classified by considering the diffusion maps as a learning eigenfunction of a data-dependent kernel. It makes use of the Nyström formula to estimate diffusion coordinates of new data [8]. An application on histological types of breast cancer is presented with VSI of Invasive Ductal Carcinoma and Mastosis.

**Materials:** VSI come from histological sections of breast tumors stained in the same laboratory according to the Hematoxylin-Eosin-Safron protocol and acquired with the same digital scanner (a ScanScope CS from Aperio Technologies). The aim being to develop a generalized CADS, it is mandatory to manage color calibration of each device used along the process, from histological staining up to image acquisition [9]. For this study, we have collected image patches from two histological types: Invasive Ductal Carcinoma (IDC) and Mastosis (Ma) with patches from the 'normal' morphology for further be able to remove non-informative patches. VSI have been acquired at X20 (0.5 µm per pixel) and stored in TIFF 6.0 file format (compression 30%). The tools are developed in Python language with the help of specialized modules (PIL: Python Imaging Library, SciPy and mathplotlib).

**Methods: Stereology:** In order to reduce the expertise workload and to obtain a reliable ground truth, a stereological test grid for point counting is over-imposed onto VSI in the ImageScope viewer [10]. The grid step has been set to 1000 x 1000 pixels (3500 points in average per image). The pathologist has then to determine which histological class is associated with the local areas centered on grid points; 30 possibilities are proposed for breast tumors. A simple mark has to be drawn on a grid point in the overlay layer whose name corresponds to his choice. Each area is then extracted at the plain resolution and stored as an uncompressed TIFF image. These areas (also called 'patches') are squares of size 400 x 400 pixels. This size has been chosen according to the representative structures encountered in breast tumors and allows to expertise only 16% of a VSI.

**Features extraction:** For each patch, some statistical features are computed and embedded in a vector with its histological type and its coordinates in the stereological grid. At this stage of the study, all features are obtained from global measurements on patches computed on *RGB* color components (reduced to 64 values) and from the two first components (*H*, *E*) of the color deconvolution specific to Hematoxylin and Eosin staining [11]. For any given component *X*, the computed features are: *X*, *X* reverse sorting, cumulative_*X*, 20%-40%-60%-80% quantiles of cumulative_*X*, mean_*X*, median_*X*, mode_*X*, Skewness_*X*, Kurtosis_*X*, PearsonModeSkewness_*X*, that is a total of 13 data. Three of them are themselves histograms with 64 values but will provide a single measure after computing the distance between two signatures.With the 5 components (*R*, *G*, *B*, *H*, *E*) 65 measures will be taken into account for a patch but 1010 values will be stored in its signature. Considering the sparse numerical range of features, the symmetric Kullback-Leibler distance has been retained for its ability to easily manage such values, while remaining fast to implement. The distance between two vectors $p_1$, $p_2$ of length $n$ is then defined by:

```
<?xml version="1.0" encoding="utf-8"?>
<kml xmlns="http://earth.google.com/kml/2.1">
  <Document>
    <name>059-2.jpg</name>
    <description>Telepathology Congress</description>

    <LookAt>
      <longitude>-0.40500000000000</longitude>
      <latitude>39.47503845807844</latitude>
      <altitude>500</altitude>
      <range>2000</range>
      <tilt>0</tilt>
      <heading>0</heading>
      <altitudeMode>absolute</altitudeMode>
    </LookAt>

    <NetworkLink>
      <name>0/0/0.png</name>
      <Link>
        <href>http://digipat.org/vs/GE/059-2/0/0/0.kmz</href>
      </Link>
    </NetworkLink>
  </Document>
</kml>
```

**Data reduction:** This study aims to develop a CADS whose one component is a visualization tool showing relations between breast cancer images, stored in a knowledge database, and new images presented to the system.

**Table 1(abstract S9) Computation time on a PC (dual core)**

| Patch number | Features extraction (in seconds) | Spectral analysis (in seconds) |
|---|---|---|
| 250 | 46 | 17 |
| 500 | 98 | 69 |
| 1000 | 180 | 308 |
| 2000 | 407 | 1429 |

Typically, these relations may be expressed as a connected graph in a 3D space where we hope to find 30 distinctive clusters corresponding to histological types or sub-types. It is therefore mandatory to reduce dimensionality from n (65 dimensions in our example) to 3. The signatures being non linear data, it is not appropriate to perform a principal component analysis (PCA). Belkin [5] and Coifman [6] have shown that methods based on Spectral Connectivity Analysis (SCA) such as diffusion maps, involving eigenvalues and eigenvectors from a normalized graph Laplacian, are well suited to non linear data. Let $X=\{x_1,x_2,...,x_n\}$ be a set of $n$ patches that we estimate as a fully connected graph $G$, that means a distance function is computed for each pair $\{x_i,x_j\}$. A $nxn$ kernel $P$ is obtained from a Gaussian function whose coefficients are given by:

$$p(x_i,x_j)=\frac{w(x_i,x_j)}{d(x_i)} \text{ with } d(x_i)=\sum_{x_k\in X} w(x_i,x_k) \text{ and } w(x_i,x_j)=e^{-\left(\frac{D_{KL}(x_i,x_j)}{\epsilon}\right)}.$$

In fact, $p(x_i,x_j)$ may be viewed as the transition kernel of the Markov chain on $G$. In other words, $p(x_i,x_j)$ defines the transition probability for going from $x_i$ to $x_j$ in one time step. The eigenvectors $\Pi_k$ of $P$, ordered by decreasing positive eigenvalues, give the practical observation space axes. It must be noticed that $\Pi_0$ is never used since linked to eigenvalue $\lfloor=1$ (i.e. the data set mean or trivial solution). Projection is then done along $(\Pi_1,\Pi_2,\Pi_3)$ for a 3D visualization. Choosing $\Sigma$ in $w(x_i,x_j)$ is an empirical task which should permit a moderate decrease of the exponential; some works use the median value of all distances $D_{KL}(x_i,x_j)$ where other use the mean distance obtained from the k nearest neighbors of a subset of $X$ [6].

**Out-of-sample Nyström extension:** SCA techniques share one major characteristic that is to compute the spectrum of a positive definite kernel. It is known that the eigenvalue decomposition of a matrix $P\in\mathbb{R}^{n\times n}$ can be computed no faster than $O(n^3)$; this limits SCA techniques to moderately sized problems [12]. Fortunately, the Nyström extension, originally applied for finding numerical solutions of integral equations, can be used to compute eigenvectors and eigenvalues of a sub-matrix formed by $m$ columns of $P$ randomly subsampled and then extended to the remaining $n$-$m$ columns [8]. Given an $nxn$ matrix $P$ and an integer $m<n$. Let call $P^{(m)}$ the matrix formed by $m$ columns of $P$ that is the graph Laplacian of a set $Y\subset X$ with $|Y|=m$. $Y$ is then a training set. The orthonormal matrix of eigenvectors $U^{(m)}$ and their associated eigenvalues in a diagonal matrix $\ddot{E}^{(m)}$ are classically obtained from $P^{(m)}$ by solving: $P^{(m)}U^{(m)}=\ddot{E}^{(m)}U^{(m)}$. This step has to be run once and then may be considered as an 'off-line' procedure. The Nyström formula allows to obtain the approximate eigenvectors of all the set $X$ by:

$$\widehat{u_i}=\sqrt{\frac{m}{n}}\frac{1}{\lambda_i^{(m)}}P_{N,M}u_i^{(m)}$$

where $\lambda_i^{(m)}$ and $u_i^{(m)}$ are the $i^{th}$ diagonal entry and $i^{th}$ column of $\ddot{E}^{(m)}$ and $U^{(m)}$ respectively. $P_{N,M}$ is a $nxm$ sub-matrix of the complete graph obtained from distances $w(x_i,x_j)$. Its computation is an 'on-line' procedure having to be

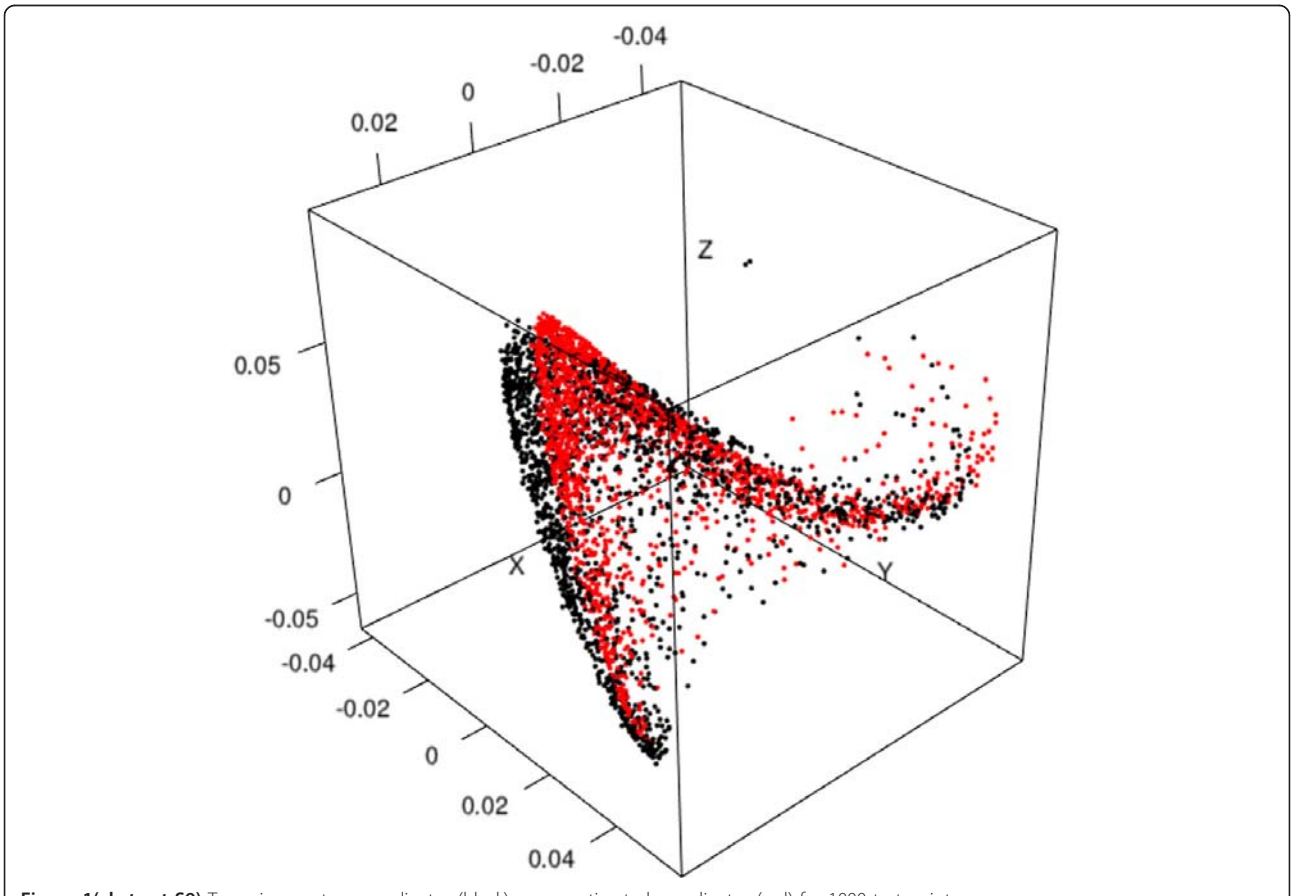



**Figure 1(abstract S9)** True eigenvectors coordinates (black) versus estimated coordinates (red) for 1000 test points.

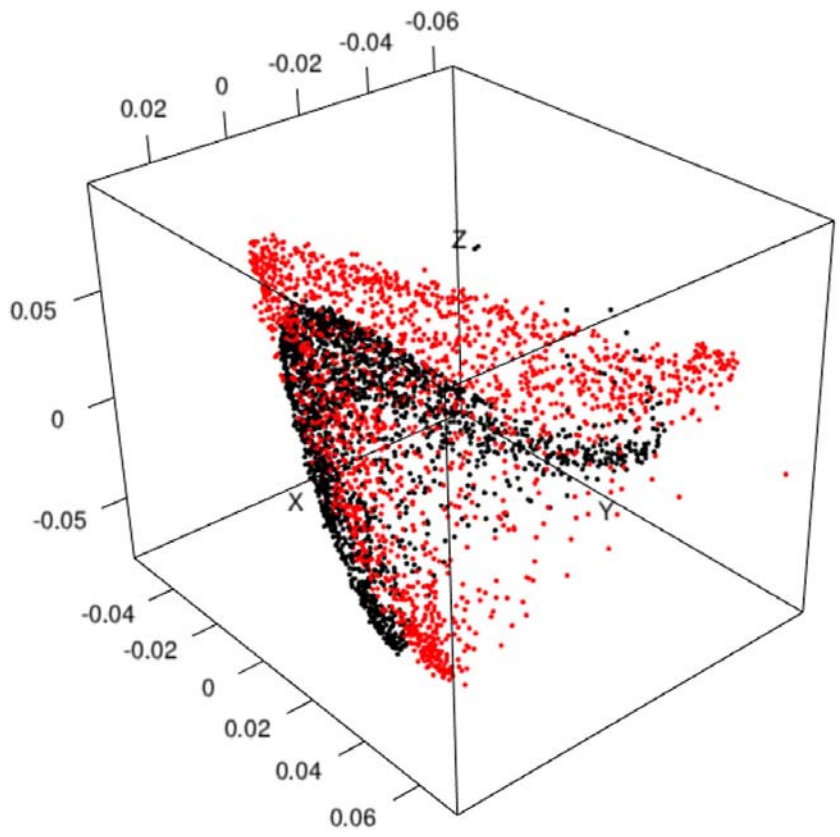



**Figure 2(abstract S9)** True eigenvectors coordinates (black) versus estimated coordinates (red) for 500 test points.

conducted for each new test set (*X**Y*). For a 3D visualization, the second to fourth columns are used (the first one being the trivial solution).

**Results and discussion:** To illustrate the out-of-sample extension to diffusion maps, 7 VSI of breast cancer cases have been used. Their mean size is 80 000 × 42 000 pixels$^2$. A total number of 1857 patches, classified as Mastosis (919 Ma), Invasive Ductal Carcinoma (812 IDC) and Normal (126 Nor) have been extracted from their inner stereological test grid. At first, table 1 shows that features extraction is $O(n)$ while the spectral analysis is close to $O(n^3)$; it has to be noticed that the latter involves both eigenvectors decomposition and code for managing the CADS. Figure 1 illustrates the projection of patches with their true eigenvectors (in black) and their estimated coordinates obtained from 1000 patches (in red). The visual comparison shows that computing a classical Euclidean distance between two points should be equivalent in both cases. Figure 2 shows the same approach from only 500 patches. Besides a shift between clouds of points, a rescaling is visible but the main shape is still preserved. To confirm this assertion we have analyzed for each patch the histological type of their nearest neighbor. This has been done both with the true eigenvectors and the estimated coordinates. In our application four cases are considered: a 'Ma' patch may be associated with another 'Ma' or 'IDC' whereas a 'IDC' patch may be associated with 'IDC' or 'Ma'. When a patch is close to the 'normal' type, we consider it as non-informative. Table 2 shows that the Nyström extension allows to obtain very similar results than the true eigenvectors (row 'reference').

**Table 2(abstract S9) Histological type in the nearest neighborhood**

| Number of test points | Ma | | IDC | |
|---|---|---|---|---|
| | Ma | IDC | Ma | IDC |
| 500 | 70.7% | 18.0% | 15.6% | 78.8% |
| 1000 | 72.2% | 17.6% | 15.3% | 80.4% |
| reference | 73.9% | 17.1% | 14.8% | 82.1% |

**Conclusion:** This work is the second part of a CADS we aim to develop based on an original strategy starting from VS and leading to an unbiased knowledge database containing reference patches of breast tumors. The first part has been presented in [4]. We have shown that combining stereological sampling and data reduction based on diffusion maps offers an interesting general framework. The results illustrated here are a proof of concept of the second part that is to classify new unknown patches. About 400 high resolution VS are now available in our lab; the benign and malignant breast tumors are classified into 30 histological types and subtypes. We plan to project some reference patches extracted from these 30 classes in the same 3D space, in order to build clusters, and then to classify a new unknown VS previously split in patches. But the spectral decomposition is very CPU intensive and managing for example 30 000 patches at a time (1 000 per histological type) would rapidly become impossible to compute. The Nyström extension seems to provide a good approximation of eigenvectors which then allow to reduce this computational burden.

**References**

1. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, Erps KA, Bhattacharyya AK, Yagi Y, Gilbertson JR: **Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future.** *Hum Pathol* 2009, **40(8)**:1057-69.
2. Kayser K, Gortler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, Kayser G: **AI (artificial intelligence) in histopathology-from image analysis to automated diagnosis.** *Folia Histochem Cytobiol* 2009, **47(3)**:355-61.
3. Oger M, Belhomme P, Klossa J, Michels JJ, Elmoataz A: **Automated region of interest retrieval and classification using spectral analysis.** *Diagnostic Pathology* 2008, **3(Suppl 1)**:S17.
4. Belhomme P, Oger M, Michels JJ, Plancoulaine B, Herlin P: **Towards a computer aided diagnosis system dedicated to virtual microscopy based**

on stereology sampling and diffusion maps. *Diagnostic Pathology* 2011, **6(Suppl 1)**:S3.
5. Baddeley A, Jensen EB: **Stereology for Statisticians.** Chapman & Hall/CRC 2005.
6. Belkin M, Niyogi P: **Laplacian eigenmaps for dimensionality reduction and data representation.** *Neural Computation* 2003, **15**:1373-1396.
7. Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker S: **Geometric diffusions as a tool for harmonics analysis and structure definition of data: Diffusion maps.** *Proceedings of the National Academy of Sciences* 2005, **102(21)**:7426-7431.
8. Williams C, Seeger M: **Using the Nyström method to speed up kernel machines.** *Advances in Neural Information Processing Systems* 2001, **13**:682-688.
9. Kayser K, Borkenfeld S, Gortler J, Kayser G: **Image standardization in tissue–based diagnosis.** *Diagnostic Pathology* 2010, **5(Suppl 1)**:S13.
10. Herlin P: **Computer-Assisted Stereology for Pathology Applications.** *Science Webinar series* 2009 [http://www.aperio.com].
11. Ruifrok AC, Johnston DA: **Quantification of histochemical staining by color deconvolution.** *Anal Quant Cytol Histol* 2001, **23**:291-299.
12. Homrighausen D, McDonald DJ: **Spectral approxima-tions in machine learning.** 2011, ArXiv:1107.4340.

## S10

### Telepathology consultation in China using whole slide image and an internet based platform

Chen Zhou[1*], Amir Rahemtulla[1], Liu Tong Hua[2], Huaqiang Shi[3]
[1]Department of Pathology, University of British Columbia, Vancouver, Canada; [2]Department of Pathology, Peking Union Medical College, Beijing, China; [3]Department of Pathology, General Hospital of Air Force, Beijing, China

*Diagnostic Pathology* 2013, **8(Suppl 1):**S10

**Background:** Telepathology is a particularly useful pathological tool suitable for developing countries such as China, which generates a large number of pathology specimens each year due to the size of its population but has a shortage of well trained and experienced pathologists in primary cares hospitals and hospitals within under-developed regions. Pathologists in these hospitals often have difficulty diagnosing challenging pathology cases. Telepathology, especially second opinion or teleconsultation is one of the solutions to this problem [1-3]. We reported an internet based open telepathology consultation platform using whole slide images (WSI) in China. The results from the telepathology consultation service since inception have been analyzed and summarized for this publication. We believe that our experiences will help to promote telepathology in China and in other developing countries.

**Methods:** The telepathology consultation cases used in this report represent pathology consultation sent to the telepathology consultation platform (http://www.mpathology.cn/mpcc/) from the beginning of the service in July 2008 to May 30, 2011. The cases submitted for teleconsultation were from 29 institutions, which were equipped with a virtual microscope, Motic Virtual Microscopic Scanner (Motic, China). The equipment and related software used in this report were validated in a previous telepathology study using a variety of 600 surgical pathology specimens [4]. When the participating hospital had a pathology case requiring consultation, a referring pathologist logged into the website, http://www.mpathology.cn/mpcc/ with a secure user name and password and filled out an online request form which included the patient's name, age, and relevant clinical information, gross findings, immunohistochemistry results, a preliminary diagnosis and the name of expert pathologist chosen for consultation. Referring pathologist then scanned, uploaded and attached WSI of one or several representative H/E slides as well as relevant immunohistochemistry stained slides to the request form, and then sent these to the platform.

An internet based telepathology platform was used and a server was used for storage of images and data. The system was maintained by two information technology (IT) technicians and one system manager. When referring pathologists from submitting hospital sent request for consults to the platform, IT technicians would be alerted by e-mails. The technician would then exam the completeness of submitted materials and system manager would contact the expert pathologist for consultation.

A panel of 84 Chinese pathology experts was invited and agreed to participate in the teleconsultation service. The names, affiliation of the pathologists and the areas of their subspecialty expertise were listed on the website (http://www.mpathology.cn/mpcc/). The platform covered the medical malpractice insurance for the expert consultants. When an expert pathologist was requested for consultation, the expert was instantly noticed by the cell phone message and e-mail. The expert pathologist would use a computer or an Ipad to log into the website, review WSI, capture a representative image, write the pathologic diagnosis, sign the report with an electronic signature and then release the final report.

Once the final report was released, the system sent an email to alert the system manager who then sent the final consultation report by e-mail or fax to the referring pathologist.

Statistical analysis was performed using SPSS 7.0 edition.

**Results: Consultation cases submitted by hospitals:** The number of cases submitted increased from 17 cases in 2008 to 587 cases in 2010. 935/1022 cases (91%) were sent in 2010 and the first 5 months of 2011.

A total of 29 hospitals were participated in the telepathology consultation service. The number of participating hospitals was 3 in 2008, 9 in 2009, and 21 in 2010; a 7-folds increase from 2008 to 2010. The mean number of cases submitted by each hospital was 35 cases in the past 3 years. Ten hospitals accounted for 90% of the cases submitted, with the other 19 hospitals accounting for only 10% of the submitted cases.

The average number of WSI per case was 1.8 with a range from 1 to 11. 62.0% of the cases only had one WSI, 27.6% had two WSI, and 10.4% had 3 or more WSI. 203 (19.8%) of the cases had included immunohistochemistry slides.

**Turnaround time:** The average time needed from transmitting teleconsultation request form with attached WSI image to the server to the release of teleconsultation pathology report by expert consultants was 38 hours with a range of less than 1 hour to 323 hours. The teleconsultation report was released within 12 hours for 43.2% of cases; 24 hours for 65.9% of cases and 48 hours for 79% of cases (Table 1).

**Expert consultants:** Although a list of 84 expert consultants was available for teleconsultation, only 43 expert consultants were requested by submitting hospitals for teleconsultation. Among these 43 experts, 23 consultants were requested most frequently, accounting for 95% of the consultation cases.

**Pathology site of the submitting cases:** The common sites of the pathology were gynecologic (24.2%), gastrointestinal/liver/pancreatic (14.7%), and lung (13.1%), each accounting for more than 10% of cases. 810 (79.3%) out of 1022 were neoplastic pathology. Among them, 193 (21.3%) were benign tumors but 637 (78.7%) were malignant.

**Agreement of second opinion with preliminary diagnosis provided by referring pathologists:** 302 (29.5%) out of 1022 cases were not given a preliminary pathology diagnosis by submitting hospitals. Among 720 cases with a preliminary pathology diagnosis, 122 (16.9%) of the cases received a consultation report which was not in agreement with the preliminary pathology diagnosis (Table 2). Local hospital pathologists could not render a preliminary diagnosis or made a wrong preliminary diagnosis in 424 cases, representing 41.5% of the total teleconsultation cases.

**Discussion:** Our results indicated that telepathology consultation in China has gained traction since 2010. More than 90% of consultation cases submitted during the 3 years of this study were in 2010 and the first 5 months of 2011. The main reason is the governmental support for telepathology consultations. In 2010, the Ministry of Health in China released a document encouraging the use of telepathology across the country; other reasons are the availabilities of high speed internet and 3G networks and low cost commercial virtual microscopes in China after 2009.

Ideally, it is would be best to submit WSI of all tissue slides to the expert pathologist for teleconsultation as in traditional pathology consults.

**Table 1(abstract S10) Turnaround time for teleconsultation**

| Time (hours) | No. Cases | Percent(%) |
|---|---|---|
| **<12** | 441 | 43.2 |
| **12-24** | 232 | 22.7 |
| **24-48** | 134 | 13.1 |
| **48-168** | 151 | 14.8 |
| **>168** | 64 | 6.2 |

**Table 2(abstract S10) Agreement of expert opinions with preliminary diagnosis**

| Preliminary diagnosis | No. Cases | Agreement of expert opinions with preliminary diagnosis | | | |
|---|---|---|---|---|---|
| | | Concordance | General concordance | No concordance | No explanation |
| **Yes** | 720 | 282 | 218 | 122 | 98 |
| **No** | 302 | - | - | - | - |
| **Total** | 1022 | 282 | 219 | 122 | 98 |

However, submitting and viewing all digitalized WSI from a single case is time consuming and more expensive. In this report, about 90% cases had ≤2 WSI. Our report indicates that using WSI in teleconsultation is effective and that most cases need only 1 or 2 WSI.

An internet based open platform used in this study has several advantages. First, expert pathologists are from multiple institutions and enjoy national reputations in their subspecialties. Second, the platform can be accessed by expert pathologists at anytime or places where internet or a Wi-Fi connection is available. Third, a centralized open platform can accept a large number of consultation cases, creating an economic model for the long-term survival. In several previous reports, the telepathology consultation services were either closed systems or offered only local service [5,6]. Such systems often had limited cases and were economically not viable.

One of the advantages of teleconsultation is a much shorter turnaround time(TAT), as compared with a mean of 6 days TAT in traditional pathology consultation [7]. Our results show that the average TAT was 38 hours and that about 2/3 of cases had a TAT within 24 hours. Approximately 10% of cases had a TAT longer than 7 days, this is due to expert pathologists requested additional slides or immunohistochemistry staining for some cases before rendering a final diagnosis.

Our analysis showed that although 84 expert consultants have been listed on the website, only 43 of them had participated in the consultation service and 23 of them completed 90% of the consultation work. The result indicated that a large telepathology consultation service may operate using only about 25 pathologists with different subspecialty expertise, reducing the cost of malpractice insurance and the difficulty for referring pathologists to select an expert pathologist.

In this report, we found that gynecologic specimens, gastrointestinal and live/pancreatic specimens and lung specimens were the most common specimens sent for consultation. A high proportion of head and neck, soft tissue, and hematopoietic pathology cases were also among the more frequent consultation specimens. Unlike consultation cases in Western countries, a low proportion of cases consisted of skin pathology [8]. This is similar to the findings of a previous telepathology study in China [4].

In our study, about 30% of cases did not come with a preliminary pathology diagnosis. This may be related to the degree of difficulty of those submitted cases, and/or due to the lack of pathology training and experience of submitting pathologists in China. Among those having preliminary pathology diagnosis, 16.9% were not in agreement with the expert diagnosis. This is in agreement with previous reports [9,10].

Our analysis showed that consultation cases without preliminary diagnosis and with the wrong preliminary diagnosis accounted for 41.5% of teleconsultation cases. This indicated pathologic diagnosis was valuable and clinically significant in 40% of the teleconsultation cases. This finding demonstrated that telepathology has an important and practical role in pathology consultation in China.

**Competing interests:** None.

**Authors' contributions:** CZ: participating the study, analysis and summary of the data and writing up manuscript
AR: direct analysis and writing up manuscript
LTH: directing the study and reviewing the manuscript
HS, participating and supervising the study, and reviewing manuscript

**Acknowledgments:** The authors would like to thank Huang Rong Sheng for providing some of the material and Xiajun Zhu, Zhichuang Huang, Liqun Zhang for providing technical support for the telepathology consultation platform.

**References**

1. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, *et al*: Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future. *Human pathology* 2009, **40(8)**:1057-69.
2. Hitchcock CL: The Future of Telepathology for the Developing World. *Archives of pathology & laboratory medicine* 2011, **135(2)**:211-4.
3. Wamala D, Katamba A, Dworak O: Feasibility and diagnostic accuracy of Internet-based dynamic telepathology between Uganda and Germany. *Journal of telemedicine and telecare* 2011, jtt. 2010.100609 v1.
4. Li X, Gong E, McNutt MA, Liu J, Li F, Li T, *et al*: Assessment of diagnostic accuracy and feasibility of dynamic telepathology in China. *Human pathology* 2008, **39(2)**:236-42.
5. Dietel M, Nguyen©\Dobinsky TN, Hufnagl P: The uicc telepathology consultation center. *Cancer* 2000, **89(1)**:187-91.
6. Brauchli K, Oberholzer M: The iPath telemedicine platform. *J Telemed Telecare* 2005, **11(Suppl 2)**:S3-7.
7. Azam M, Nakhleh RE: Surgical pathology extradepartmental consultation practices. *Arch Pathol Lab Med* 2002, **126(4)**:405-12.
8. Zembowicz A, Ahmad A, Lyle SR: A Comprehensive Analysis of a Web-Based Dermatopathology Second Opinion Consultation Practice. *Archives of pathology & laboratory medicine* 2011, **135(3)**:379-83.
9. Hsu CY, Su IJ, Lin MC, Kuo TT, Jung SM, Ho DM: Extra-departmental anatomic pathology expert consultation inTaiwan: a research grant supported 4-year experience. *J Surg Oncol* 2010, **101(5)**:430-5.
10. Tsung JSH: Institutional pathology consultation. *The American journal of surgical pathology* 2004, **28(3)**:399.

## S11

### Virtual microscopy with Google-Earth: a step in the way for compatibility

Luis Alfaro[1*], Maria Jose Roca[2], Pablo Catala[3]
[1]Department of Pathology. Fundacion Oftalmologica del Mediterraneo. Valencia. Spain; [2]Department of Pathology. Hospital Arnau de Vilanova. Valencia Spain; [3]IT department. Fundacion Oftalmologica del Mediterraneo. Valencia. Spain
E-mail: lalfaro@comv.es
*Diagnostic Pathology* 2013, **8(Suppl 1):**S11

**Background:** Advances in the field of virtual microscopy are continuously growing. Many companies have introduced equipments with very good image quality levels, and increased speed in the scanning processes. However also a wide variety of image formats, software viewers and servers have appeared, with lack of compatibility in the managements of virtual slides.

Commercial solutions for virtual microscopy tend to be rigid and difficult of customize, probably to protect the developments, but this leads to a detachment in the management of the images and a difficulty in becoming familiar with this technology. Handling virtual slides in a similar way as we do with conventional pictures taken from digital cameras surely would bring to virtual microscopy a much wider number of pathologist.

In this approach we tried to adapt to virtual microscopy simplistic solutions employed in digital photography as software oriented for panoramic images [1]. Panoramic images share with virtual slides their huge size and a similar way to be generated stitching smaller images [2]. Software for panoramic images can be adapted for virtual microscopy.

Google Earth is a well-known software oriented as a geographic information system working in a way similar to virtual microscopy, zooming and panning images, and moving along huge files. It is widely distributed, installed in may computers and can be useful to share virtual slides and employed as a viewer of virtual slides [3].

There are developments for virtual microscopy based in the use of the API (application programming interface) of Google Maps, such as the NYU School of Medicine Virtual Microscope [4]. These options of high complexity require a team of programmers and computer support, not available in all

situations. However, it is possible for pathologists to use Google Earth as a viewer of our virtual slides more easily, and without programming knowledge.

**Material and methods:** With the aim of testing the value of Google Earth as a software for the handling of virtual slides we selected 20 pathology cases. Glass slides of 10 were scanned with a·3D-Histech Panoramic Midi, and the other 10 slides with an Aperio XT. Original virtual slides were exported into .jpg flat files with Aperio ScanScope software. We generated kml files, the file format to display information in Google Earth, and the compatible pyramidal tiles structure.

The software employed was GDAL, an open source library for geospatial data images, with the utility gdal2tiles, and its graphical interface variant (MapTiler). All the cases were uploaded into two servers, our own server at the hospital, and a external sever hosting web pages. HTML web pages were built linking the cases with the kml files. MIME types were defined in both servers in order to lead kml files be opened with Google Earth.

**Results and discussion:** Cases were accessible anywhere from the Internet through its web address and opened directly with Google Earth. All functions to allow diagnostic, consultation, educational... purposes were available. Image quality obtained was equivalent to any other specific viewer for virtual slides. Speed in serving files was related with lines capacity and hosting server performance, and not with software.

Google Earth uses a specific file type (KML/KMZ) to define specific locations. kmz files are compressed versions of the kml with zip compression. When clicking on these files Google Earth opens and shows the location defined in the file with specific spatial coordinates.

KML files have many features, which can be reviewed at the program tutorials [5]. They allow among them, to insert pictures that are embedded on the landscapes of Google Earth. This feature, called "PhotoOverlay" is often used associated with Google Earth position marks, but also supports the use of very large photographs, with many megapixels as used in our virtual slides. The procedure for visualization of these giant photographs is the usual breaking them into small portions, and arrange them in a pyramidal structure. Each image of the pyramid is divided into tiles, so that only the parts wished to see need to be charged at every moment, and with the known functions of zoom and panning.

Kml files have a syntax similar to html files. They can be created manually and adding a link to the virtual slide with the PhotoOverlay tag is not difficult. Anyway there are available kml file generators [6,7].

The organization of the virtual slide in a set of pyramidal tiles although can also be generated manually, needs in practice an automated system to avoid a long and repetitive process. Several tools are designed for this purpose. We selected MapTiler [8] because of its easiness of use and the simultaneous generation of the associated kml file. The main handicap of this software is that it becomes very slow even in the most modern computers when facing very big virtual slides.

A similar alternative for the generation of the pyramidal image is gdal2tiles.py a software that generates a directory with the small tiles from the fragmentation of the virtual slide an also the kml file to be opened with Google Earth [9]. It's a command line software based on python programming language not as easy to be used.

An example of a virtual slide prepared to be seen with Google Earth is provided in Figure 1.

Table 1 shows the syntax of the simplified KML file prepared for the example.

A certain limitation in the use of this alternatives is that original virtual slides need to be exported into flat jpg files which will be the template to generate the pyramidal structure suitable for Google Earth. The well

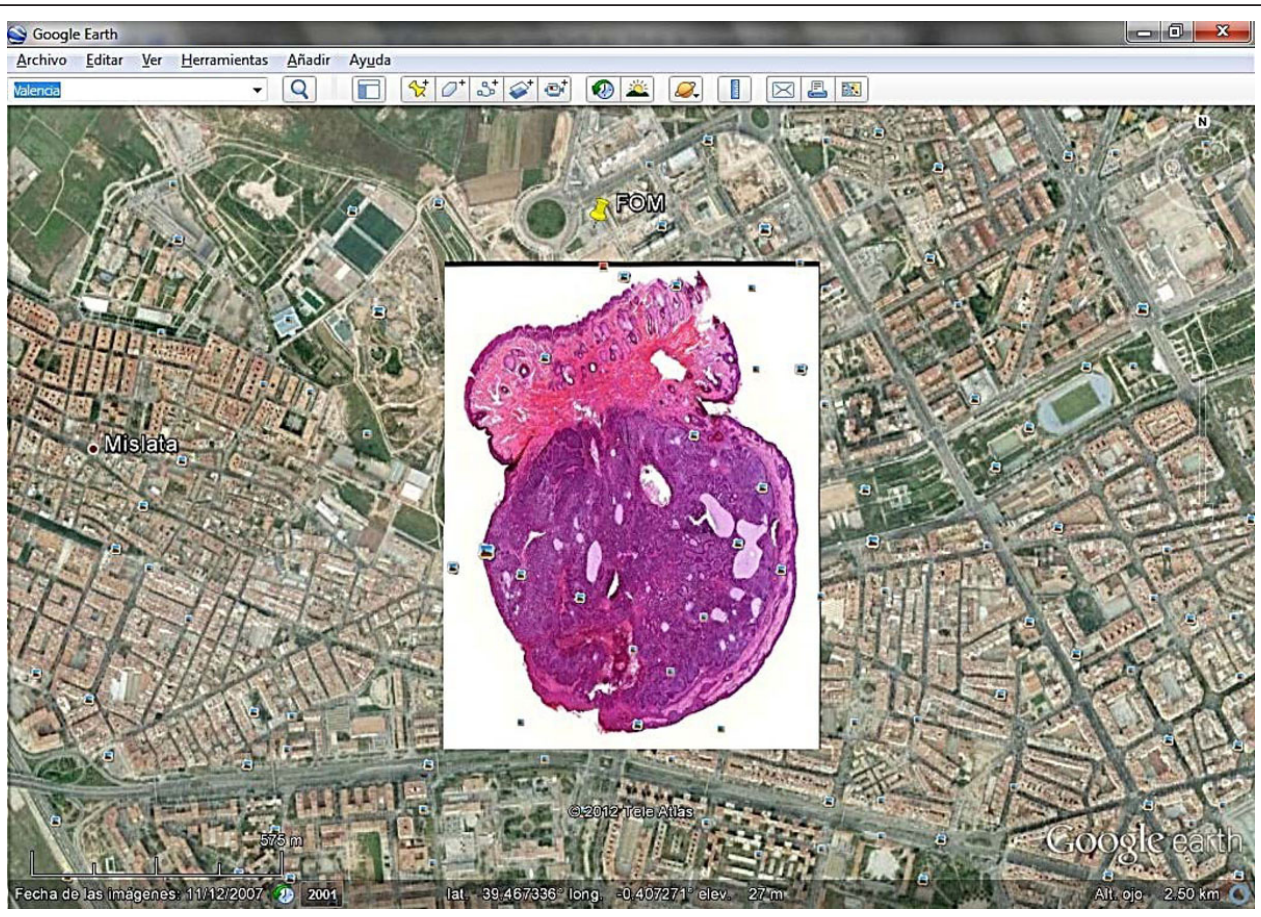



**Figure 1(abstract S11)** Google Earth showing a virtual slide.

**Table 1(abstract S11) KML file prepared to open a virtual slide with Google Earth**

```
<?xml version="1.0" encoding="utf-8"?>
<kml xmlns="http://earth.google.com/kml/2.1">
  <Document>
    <name>059-2.jpg</name>
    <description>Telepathology Congress</description>

    <LookAt>
      <longitude>-0.40500000000000</longitude>
      <latitude>39.47503845807844</latitude>
      <altitude>500</altitude>
      <range>2000</range>
      <tilt>0</tilt>
      <heading>0</heading>
      <altitudeMode>absolute</altitudeMode>
    </LookAt>

    <NetworkLink>
      <name>0/0/0.png</name>
      <Link>
        <href>http://digipat.org/vs/GE/059-2/0/0/0.kmz</href>
      </Link>
    </NetworkLink>
  </Document>
</kml>
```

know restriction of jpg files to a maximum size of 65.000 pixels ($2^{16}$) can be insufficient for big virtual slides. JPG2000 format is not available for the analyzed software, and conventional Tiff files have also a size limitation of 4 GB.

**Conclusions:** Google Earth is widely distributed and can be a good choice to avoid compatibility limitations in virtual microcopy. Even most viewers for virtual slides are free, sometimes especially for remote consultation, it is not possible to expect a remote pathologist having installed all different browsers. Besides many pathologist in hospitals have not administrations rights to install software at their computers. Viewing slides with Google Earth requires not technical skills and any pathologist can use it easily. Exporting virtual slides and generating the tiles and files to serve them for Google Earth requires a bit more of knowledge in information technologies, however it is possible for a pathologist with some experience in virtual microscopy to do without technical assistance. No programming abilities are needed further the generation of html files to link the slides in the servers, and the full process can be semi-automated. Google Earth is also a very dynamic software with frequent actualization, and many working groups introducing new improvements and can be suitable to be used in virtual microscopy.

**List of abbreviations:** GB: Gigabyte; GDAL: Geospatial Data Abstraction Library; HTML: HyperText Markup Language; JPG (JPEG): Joint Photographic Experts Group; KML: Keyhole Markup Language; MIME: Multipurpose Internet Mail Extensions; TIFF: Tagged Image File Format

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** LA Conception, design and initial manuscript writing
MJR manuscript reviewing and important intellectual contribution
PC Technical advice and support

**References**

1. Alfaro L, Poblet E, Catalá P, Navea A, García-Rojo MJ: **Compatibilización de equipos de microscopía virtual: análisis de alternativas con software de imágenes panorámicas.** *Rev Esp Patol* 2011, **44(1)**:8-16.
2. Alfaro L, Roca MJ: **Manual generation of virtual slides: a simplistic alternative for small biopsies [abstract].** *Virchows Arch* 2011, **459(Suppl 1)**:312.
3. Alfaro L, Poblet E, Roca MJ, Catala P, Navea A: **Google Earth and panoramic photo software in the management of virtual slides. [abstract].** *Mod Pathol* 2011, **24(S1)**:339A.
4. **NYU School of Medicine Virtual Microscope.** ©NYU. [http://cloud.med.nyu.edu/virtualmicroscope/].
5. Google developers: **Keyhole Markup Language.** *KML tutorial* [https://developers.google.com/kml/documentation/kml_tut?hl=en].
6. **KML file creator.** *Free Maps tools* [http://www.freemaptools.com/kml-file-creator.htm].
7. **KML generator 2.05.** [http://www.madsencircuits.com/kmlgenerator.html].
8. MapTiler: **Map Tile Cutter.** *Map Overlay Generator for Google Maps and Google Earth* [http://www.maptiler.org/].
9. **GDAL Utilities: gdal2tiles.py.** [http://www.gdal.org/gdal2tiles.html].

## S12

### Diagnostic reproducibility on whole digital slide in cytology and histology in oncologic screenings

Stefania Lega[1], Paola Crucitti[1], Paola Pierotti[1], Roberta Rapezzi[1], Priscilla Sassoli de' Bianchi[2†], Carlo Naldoni[2†], Arrigo Bondi[1*]
[1]Anatomia Patologica Ospedale Maggiore, Azienda USL di Bologna, Italy; [2]Sanita' e Politiche Sociali – Regione Emilia-Romagna, Italy
E-mail: arrigo.bondi@ausl.bologna.it
*Diagnostic Pathology* 2013, **8(Suppl 1):**S12

**Background:** Diagnostic reproducibility and accuracy in cytology and histology are major issues in Oncologic Screenings of cervix, breast and colorectal cancer: it can be achieved by procedures and programs for quality assurance (QA). The slides set standard represents the most used method to compare diagnostic proficiency, the chance of interpreting microscopic digital photographs provided an interesting alternative to reading conventional microscope slides.

The whole digital slide observed in a computer screen is a third, interesting, option to reach the purpose. In fact all the information on conventional sample are transferred into a file, easily archived, cataloged, duplicated or advice for quality control, but is especially available at a distance and from multiple locations simultaneously with drastic reduction of time needed to achieve proficiency test reproducibility [1].

The production of digital slides with modern scanners is relatively simple and quick. All suppliers offer publishing services into private or public networks server [2] and software able to track scanned cases comprehensive database to build large casistic archives on the net [3]. While tools are already available for a teleconference discussion of cases with vision of cytological preparations on line [4,5], educational programs with integrated digital slides are poorly developed nor self-evaluation or proficiency tests for continuing education and professional updating are easily accessible.

A project on Virtual Microscopy and Digital Pathology has been conducted in Emilia-Romagna (Italy) with the objective to promote quality in diagnostic cytology and histology of Screenings by testing a different system involving pathologists and cytologists using digital slides, with a faster mechanism and reproducible than standard diagnostic sets and by distance training operators with a final consensus diagnosis meeting.

The aim has been reached with the realization of a management system for cytological and histological whole-slides digital images and related clinical data and the building of a picture archive and communication system (PACS) among pathologists of our Region. This must be backed by software for the realization of network slide seminars to perform periodic tests of diagnostic reproducibility and proficiency test. The cases, collected and properly cataloged in an online, extensive and systematic digital archive of slides, easily accessible, with diagnoses discussed in clinical-pathology audit and validated by experts, can be used as diagnostic reference tool (casistic Atlas online). The cataloging and indexing is performed with NAP codes, a Nomenclature derived from SNOMED [6], which contains terms in Italian and English and encompasses extensive synonyms and complex searches.

**Material and methods:** The cancer screening survey group of the Emilia Romagna Region (Italy) set up a picture, archive and communication system (PACS) devoted to pathologists for cooperative diagnosis, didactics and training, teleconsulting, documentation of rare cases and pilot experiences; furthermore selected cases are catalogued in the PACS with

the aim of the check of the diagnostic concordance in the regional oncologic screenings (cervix, breast and colon). The PACS system is composed by two Aperio scanner and an adequate internet server where the described programs operate (see Figure 1) [7].

The slides have been digitalized using an Aperio scanner, 20x for histology and 40x for cytology and an internet server was used to store the files, arranged into a Spectrum database (Aperio). An e-learning platform (Docebo)[8] has been used to built an interface for the applicants: cases and slides were considered "teaching objects" for the educational software (Seminars) and appropriate questioning forms have been designed with the diagnostic occurrences of the Bethesda System 2011 for cytology and the CIN options for histology for the cervical cancer and of International Guidelines for breast and colorectal cancer.

At the present the diagnostic reproducibility has been performed in colorectal and cervical cancer screening (Bologna October 2010, Bologna June 2011), and for breast cancer is ongoing (Bologna June 2012). In all three Seminars a number of cases have been selected by a committee of Pathologists from Regional Units.

Colorectal cancer screening was certainly the first courses for pathologists performed with these features in our region and maybe in Italy.

Three Regional Units (Bologna, Cesena and Ferrara) were involved by sending representative histological cases of all main diagnostic occurrences to test the diagnostic reproducibility, 28 histological cases were represented. A day interval was let to study slides, then a consensus conference has been programmed in the same day.

In cervical cancer screening, the second Seminar of QA, to test the diagnostic reproducibility, 30 cytological and 30 histological cases have been selected by a committee of Pathologists among the cases proposed by all the Regional Units. All main diagnostic occurrences were represented, basic clinical information and relevant follow up information were available; the cases have been completely anonymized for the participant.

A 30 days interval was let to study the slides, then a consensus conference has been programmed. Before the meeting each participant received a report with the interpretation of the committee and her/his diagnosis inserted in the form.

**Results and discussion:** 15 Pathologists of Regional Units attended the **colon-rectal QA** and the diagnostic reproducibility have been evaluated matching them with the final diagnosis reached during the consensus conference. The observed agreement was 69% and the overall performance of the participating readers was assessed with a statistical analysis using Cohen's kappa: the average value was 0.64 (substantial).

95 cytologists and 32 histopathologists have been involved in the **cervical cancer screening QA**.

The diagnostic reproducibility has been evaluated using the final diagnosis reached in the Consensus Conference: in 2 out of 30 cytological cases the diagnosis was different from the opinion of the committee, while all histology diagnoses were in agreement. The overall performance of the participating readers is reported in table 1.

**Conclusion:** Whole digital slide is suitable for proficiency tests and the internet e-learning platform allow to share cases and to get the answers

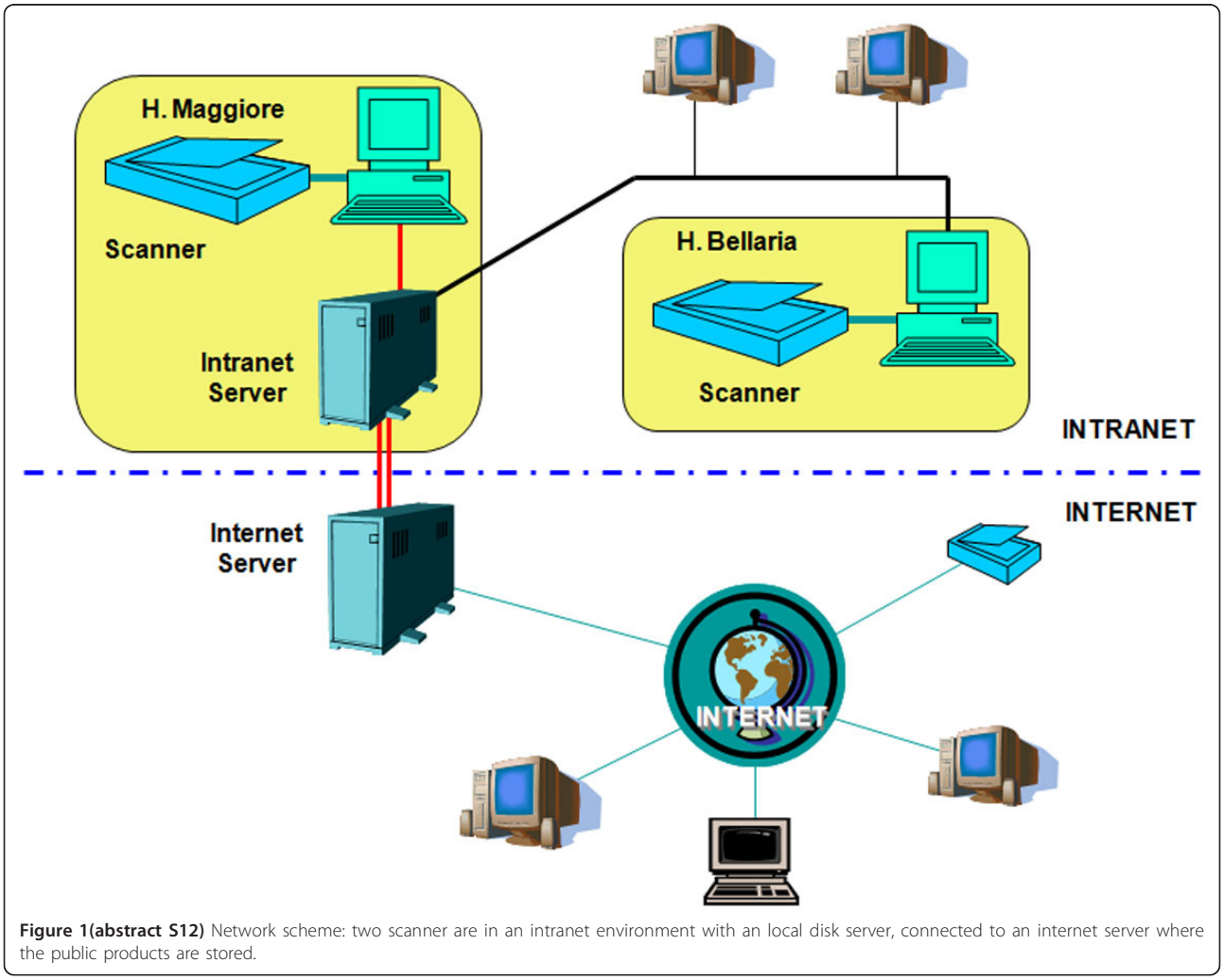



**Figure 1(abstract S12)** Network scheme: two scanner are in an intranet environment with an local disk server, connected to an internet server where the public products are stored.

**Table 1(abstract S12) Screening PAP test - Distribution of readers' diagnosis**

| | | Agreement in Omogeneous groups of diagnoses | | | |
|---|---|---|---|---|---|
| | | **neg** | **ASCUS L SIL** | **ASC-H H SIL / Ca sq** | **AGC AIS / ADK** |
| **readers' diagnosis** | **neg** | **85,1** | 15 | 6,6 | 2,1 |
| | **ASCUS/ L SIL** | 5,1 | **72,4** | 20,2 | 3,6 |
| | **ASC-H/ H SIL/ Ca sq** | 2,9 | 11,9 | **67,7** | 14 |
| | **AGC / AIS/ ADK** | 2,9 | 0,3 | 5,2 | **79,2** |
| | **advanced tumor** | 3,5 | 0,2 | 0,2 | 0,7 |
| | **no answer** | 0,5 | 0,2 | 0,1 | 0,4 |
| | | **100%** | **100%** | **100%** | **100%** |
| | Observed agreement | | 73% | | |
| | Cohen's Kappa | | 0.63 | (substantial) | |

from participants, in a easier way than in the circulation of a set of conventional slides.
The quality of whole slides is diagnostic, approaching optical microscopic resolution.
In a cytological environment the difficulty to get a perfect focus on all fields, the wider area of slide to examine and the higher number of diagnostic classes may justify a worse agreement of the users and a poorer performance (lower Cohen's kappa) than histology.
We have produced an integrated environment that includes many of the modern aspect of digital pathology that can be shared with the PACS system in many laboratories of the Region, including quality promotion and control of image interpretation in cytology and histology applied to cancer prevention screenings.
**Competing interests:** Authors declare that they have no competing interests.

**References**

1. Demichelis F, Della Mea V, Forti S, Dalla Palma P, Beltrami CA: **Digital storage of glass slides for quality assurance in histopathology and cytopathology.** *J Telemed Telecare* 2002, **8**:138-142.
2. Wilbur DC, Madi K, Colvin RB, Duncan LM, Faquin WC, Ferry JA, Frosch MP, Houser SL, Kradin RL, Lauwers GY, *et al*: **Whole-slide imaging digital pathology as a platform for teleconsultation: a pilot study using paired subspecialist correlations.** *Arch Pathol Lab Med* 2009, **133**:1949-1953.
3. Huisman A, Looijen A, van den Brink SM, van Diest PJ: **Creation of a fully digital pathology slide archive by high-volume tissue slide scanning.** *Hum Pathol* 2010, **41**:751-757.
4. Saysell E, Routley C: **Telemedicine in community-based palliative care: evaluation of a videolink teleconference project.** *Int J Palliat Nurs* 2003, **9**:489-495.
5. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, Erps KA, Bhattacharyya AK, Yagi Y, Gilbertson JR: **Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future.** *Hum Pathol* 2009, **40**:1057-1069.
6. College of American Pathologists: **SNOMED - Systematized Nomenclature of Medicine.** Skokie, Ill., USA: College of American Pathologists 1979.
7. Bondi A, Pierotti P, Crucitti P, Lega S: **The virtual slide in the promotion of cytologic and hystologic quality in oncologic screenings.** *Ann Ist Super Sanita* 2010, **46**:144-150.
8. **Docebo: e-learning open source platform.** 2010.

## S13

### A multistep image analysis method to increase automated identification efficiency in immunohistochemical nuclear markers with a high background level

Marylène Lejeune[1], Vanessa Gestí[3], Barbara Tomás[3], Anna Korzyńska[4], Albert Roso[1], Cristina Callau[1], Ramon Bosch[3], Jordi Baucells[5], Joaquín Jaén[3], Carlos López[1,2*]

[1]Molecular Biology and Research Section, Hospital de Tortosa Verge de la Cinta, IISPV, URV, Tortosa, Spain; [2]Unitat de Suport a la Recerca de la Gerencia Territorial Terres de l'Ebre, IDIAP Jordi Gol, IISPV, URV, UAB, Tortosa, Spain; [3]Department of Pathology, Hospital de Tortosa Verge de la Cinta, IISPV, Tortosa, Spain; [4]Laboratory of Processing Systems of Microscopic Image Information, Nalecz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, Warsaw, Poland; [5]Department of Informatics, Hospital de Tortosa Verge de la Cinta, IISPV, Tortosa, Spain
E-mail: clpclp3@gmail.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S13

**Background:** In anatomical and surgical pathology, the customary method of manual observation and measurement of immunohistochemically stained markers from microscopic images is tedious, expensive and time consuming. There is great demand for automated procedures for analyzing digital images (DIs) of these markers [1] given that they reduce human variability in the evaluation of stained markers [2,3] and increase the speed and efficiency of the analysis [4]. Computerized DI analysis software generally involves a stained objects/nuclei segmentation method to detect and quantify the number of positively stained markers in combination with the standard evaluation of their morphometric and/or densitometric features [5,6]. However, automatic segmentation often fails due to the presence of spurious stain deposits in tissue sections (background). The "removal" of the background from noisy DIs, so that only the objects of interest are identified, is difficult due to the color values of pixels in the nuclei and background overlapping during the color segmentation processes.
We previously developed an automated macro that allows quantification of several nuclear markers in various neoplasic tissues [7]. In an attempt to standardize the immunohistochemical analysis and to improve cell detection, we propose a new procedure that quantifies only positively stained nuclei even though they have a similar color to that of the surrounding tissue. The aim of this work was to develop a single automated procedure that allows images to be analyzed irrespective of whether the spurious stain deposit in background is present or absent. The multistep process includes algorithms that permit this discrimination so that the appropriate procedure for optimal quantification can then be applied.
**Materials and methods: Images:** Histological sections of lymphomas and breast cancer tissues, previously immunohistochemically stained with standardized protocols [8,9], were selected from the archives of the Department of Pathology of the Hospital de Tortosa Verge de la Cinta, Catalonia, Spain. Staining was performed with monoclonal antibodies directed against the nuclear protein estrogen receptors (ERs; clone NCL-ER-6F11, Novocastra, Newcastle upon Tyne, UK), progesterone receptors (PRs; NCL-PGR-312, Novocastra), Ki-67 (clone MIB-1, Dako, Carpinteria, CA, USA) and FOXP3 (clone FOXP3-236A/E7, CNIO, Spain). The entire process was standardized to ensure high reproducibility and brown staining homogeneity, which are very important requirements for image analysis [10]. This study received institutional review board approval.
**Image capture:** Stained 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. One hundred digital images were captured with a Leica DFC320 digital camera connected to a computer and controlled with the Leica IM50 v4.0 program. TIFF format DIs, with a resolution of 1392 x 1040 pixels (1.4 Mpixels) in RGB 24 true-color format, were

selected on the basis of the presence or absence of the spurious stain deposits of the background, otherwise ensuring a variety of concentrations and distributions of stained nuclei (Figure 1).

**Procedure developed in the new procedure:** The new automated multistep DI analysis procedure was developed with Image-Pro® Plus 5.0 software (Media Cybernetic, Silver Spring, USA). We had previously developed an automated macro to quantify stained nuclei in images without background using an RGB color model and iterative morphological segmentation [7]. This macro makes use of a wide color range to detect positive nuclei from the darkest to the lightest positive brown color pixel and applies a mask to displace the pixel color values of negative objects outside the segmentation color range of the positive nuclei. However, in DIs with a background this macro does not segment DIs correctly due to the similar color values of the positive nuclei and the background (Figure 2).

We therefore developed a multistep procedure that discriminates DIs as a function of the presence or absence of background and that enables the more appropriate of the three macros to be applied directly. First, a mask overlaps objects with negative and light positive-intensity pixels (Figure 3a) so that only positive objects with the darkest range of color (objects map 1) can be selected (Figure 3b). Then, using a mask only for negative objects, the next two steps select the clearer positive objects (objects map 2 and 3) using different color ranges and morphological ranges of area and roundness (Figure 3c). The objects map 4 is the sum of all the positive objects in the three previous steps. At the end of the positive selection, another step with a discriminative algorithm is applied in which the non-selected brown color is segmented with two different ranges of brown (Figure 3d). If no brown color is detected with the second range (area 2 = 0), the algorithm determines that there is no background in DIs that are automatically analyzed with the old macro (Figure 3e). On the other hand, when DIs with background were detected (area 2 > 0), they are submitted to another algorithm that calculate the ratio of the area of the two color ranges (Figure 3f). DIs with a low-level background (low ratio) are automatically analyzed with a restrictive macro with the first three steps of the new procedure (Figure 3g). DIs with a high background level (high ratio) are automatically analyzed with the new procedure, supplemented with a final step that selects the clearest objects using color and morphological segmentation (Figure 3h). The three macros detect only positive objects (individual nuclei and clusters). Information about the area of the positive objects detected with these macros is automatically exported to an Excel datasheet (Figure 3i) containing several algorithms so that the number of definitive positive nuclei in those selected objects can be estimated, as previously described [7].

**Quantification and statistical analysis:** The mean of two manual counts made by two trained observers was taken as the reference value (gold standard) to validate the results obtained with the old macro and the newly proposed procedure. The comparisons made were: manual 1 versus manual 2 readings; mean of manual readings versus old macro reading; mean of manual readings versus new procedure reading; and the old macro versus the new procedure readings. The extents of agreement between the manual and automatic results were evaluated with Bland-Altman and Kaplan-Meier analyses with their corresponding graphs. Bland-Altman graphs illustrate the differences between the compared methods with respect to the mean of each paired count. Kaplan-Meier curves portray the conditional probability of observing differences between results obtained from the methods compared. All statistical analyses were carried out with SPSS 19.0.

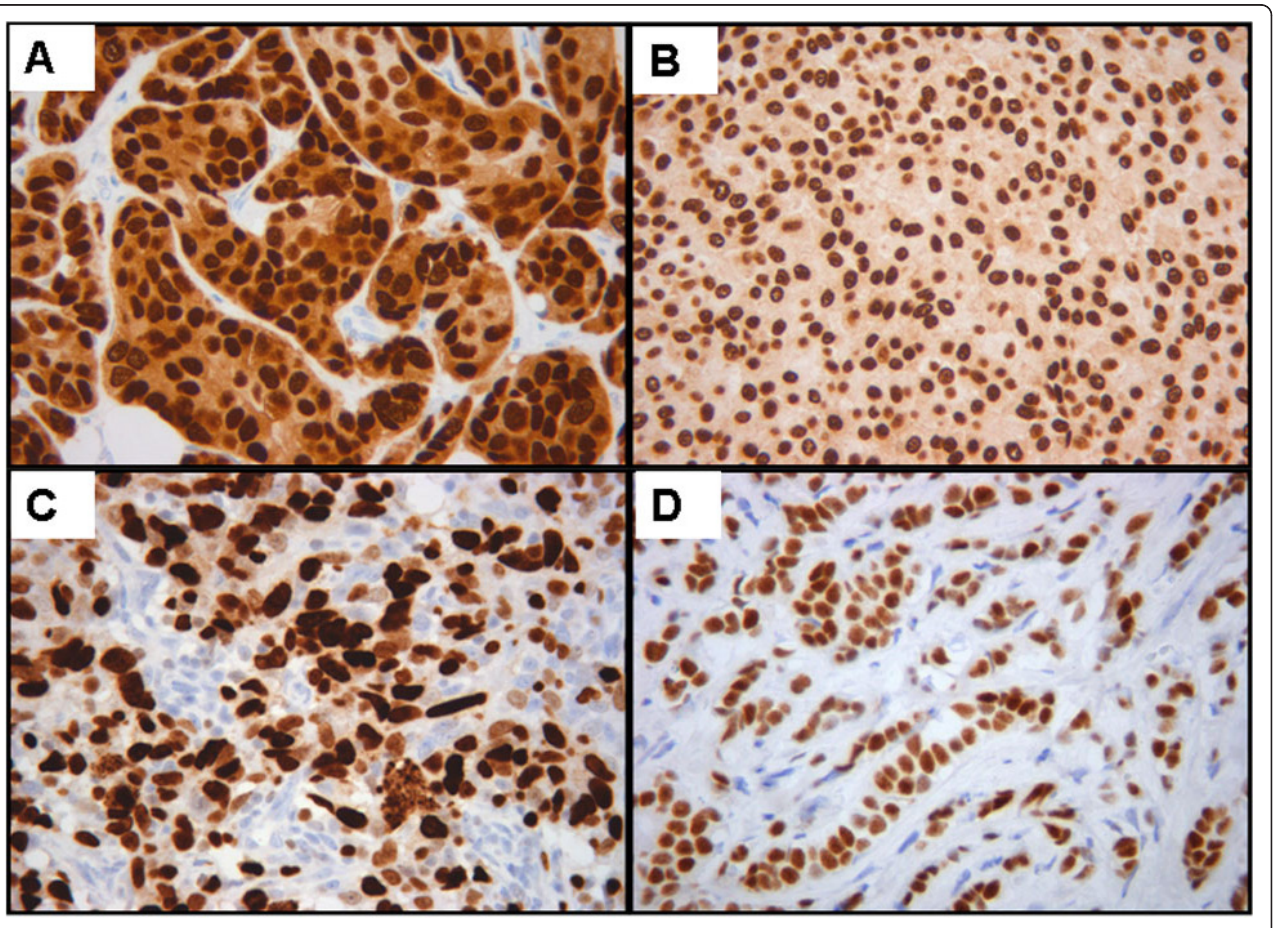



**Figure 1(abstract S13)** Illustration of original digital images of immunohistochemically stained nuclear markers with different background level (A, B) and without background (C, D).

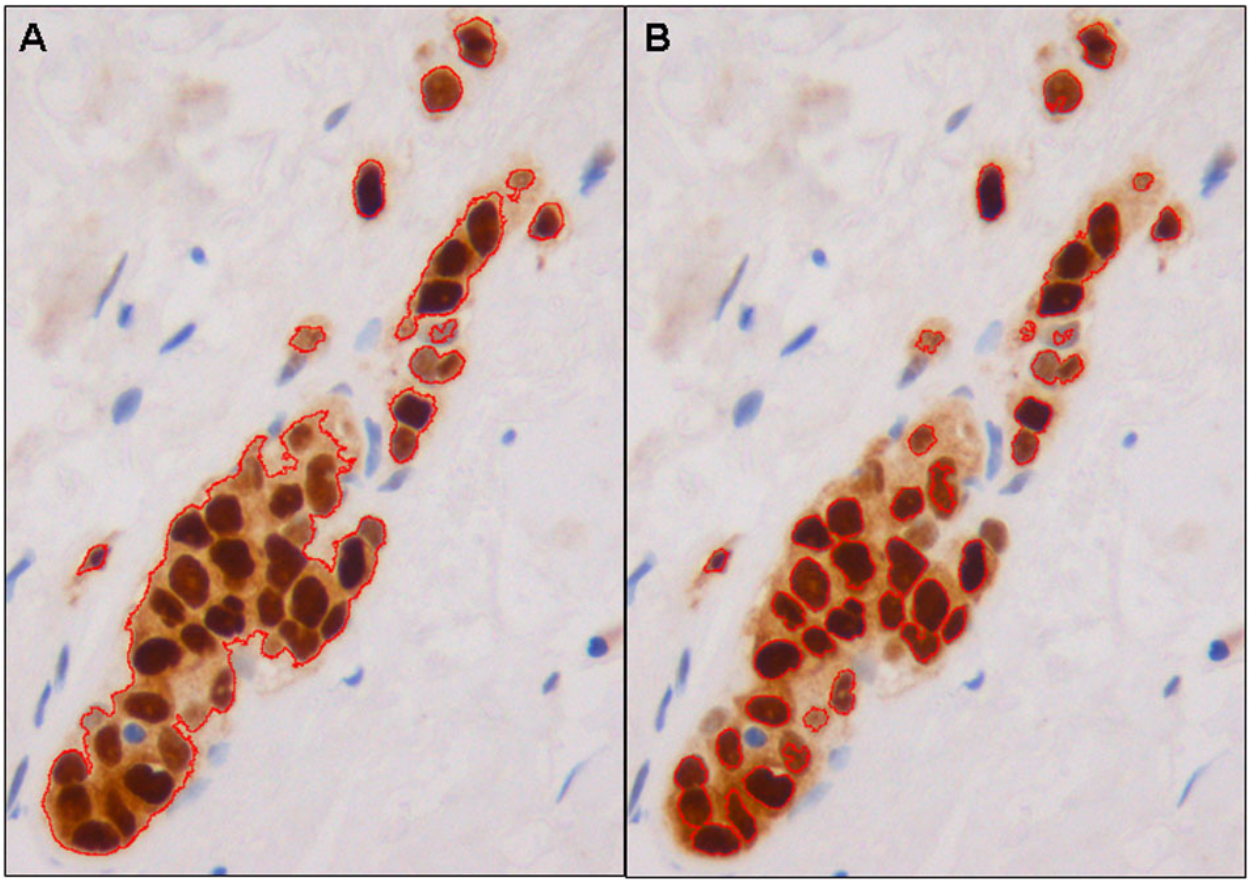



**Figure 2(abstract S13)** Example of active contour segmentation in digital images with positively stained nuclear markers. Evident differences in the contours obtained with the old macro (A) and those obtained with the new procedure (B).

**Results and discussion:** Overall, the Bland-Altman graph (Figure 4) indicates that count differences obtained comparing the new procedure vs. the mean of the manual counts (green circle) were lower than the old macro vs. the mean of the manual counts (red circle). Nevertheless, greater accuracy was obtained from images with background (Figure 5A1) where the probability of observing a difference of 50 nuclei between the new procedure and the manual quantification (22.7%, green curve) is less than half that between the old macro and manual quantification (47.4%; red curve).

As previously demonstrated, image complexity relative to the number of positively stained nuclei may affect the automated nuclear quantification [8]. In the present study, DIs were divided as before into a low-complexity group (≤100 positively stained nuclei/images) and a high-complexity group (>100 positively stained nuclei/image). As observed in Figure 5B2, the small differences in the counts between the manual method vs. the old macro and between manual counts vs. the new procedure were similar in low-complexity images. However, in high-complexity images (Figure 5B1), larger count differences were observed, although those between the manual counts vs. the new procedure were much lower than those between the manual counts vs. the old macro.

However, when images were grouped by background and complexity (Figure 6), the differences increased more with the complexity of the DIs than with the presence of the background. High-complexity DIs with a background (Figure 6A3) gave the greatest differences between manual counts vs. old macro (red curve), with those between manual counts and the new procedure being the next biggest (green curve). High-complexity images without background (Figure 6A4) gave the same difference for the comparison between the old macro and the new procedure with the manual readings (red and green curves).

A general point regarding the proposed method is that, for high-complexity DIs with background (Figure 6A3), the quantification of nuclear markers obtained with the new procedure (green curve) was closer to the gold standard (manual method, black curve) than with the old macro (red curve). For high-complexity DIs without background (Figure 6A4), the two automated methods gave the same results (red and green curves). The presence or absence of background did not appear have a great influence on the quantification of nuclear markers in low-complexity DIs (Figure 6A1-A2), probably because these images have a lower background level than the other images.

**Conclusions:** Despite specific and careful preparation of tissues and the use of blocking buffers, strong background staining can sometimes mask the detection of the target antigen during automated analysis. The results of the method presented in this paper are promising since the selective identification of brown color ranges and morphological parameters of selected objects in DIs enables the background to be discriminated during the automated localization and quantification of specific stained nuclei. The principle of this approach is applicable to all quantitative nuclear signals and should prove useful in a variety of tumor specimens, irrespective of the immunohistochemical techniques employed.

**List of abbreviations used:** DIs: digital images; ERs: estrogen receptors; FOXP3: forkhead box protein 3; PRs: progesterone receptors; TIFF: Tagged Image File Format

**Competing interests:** The authors declare that they have no competing interests.

**Author contributions:** ML, CL and JJ conceived the study. CL designed and implemented the algorithms. JB and AK provided IT support. RB and JJ carried out the manual quantification, discussed the methods and modified them. VG, BT and CC realized the automated quantification with

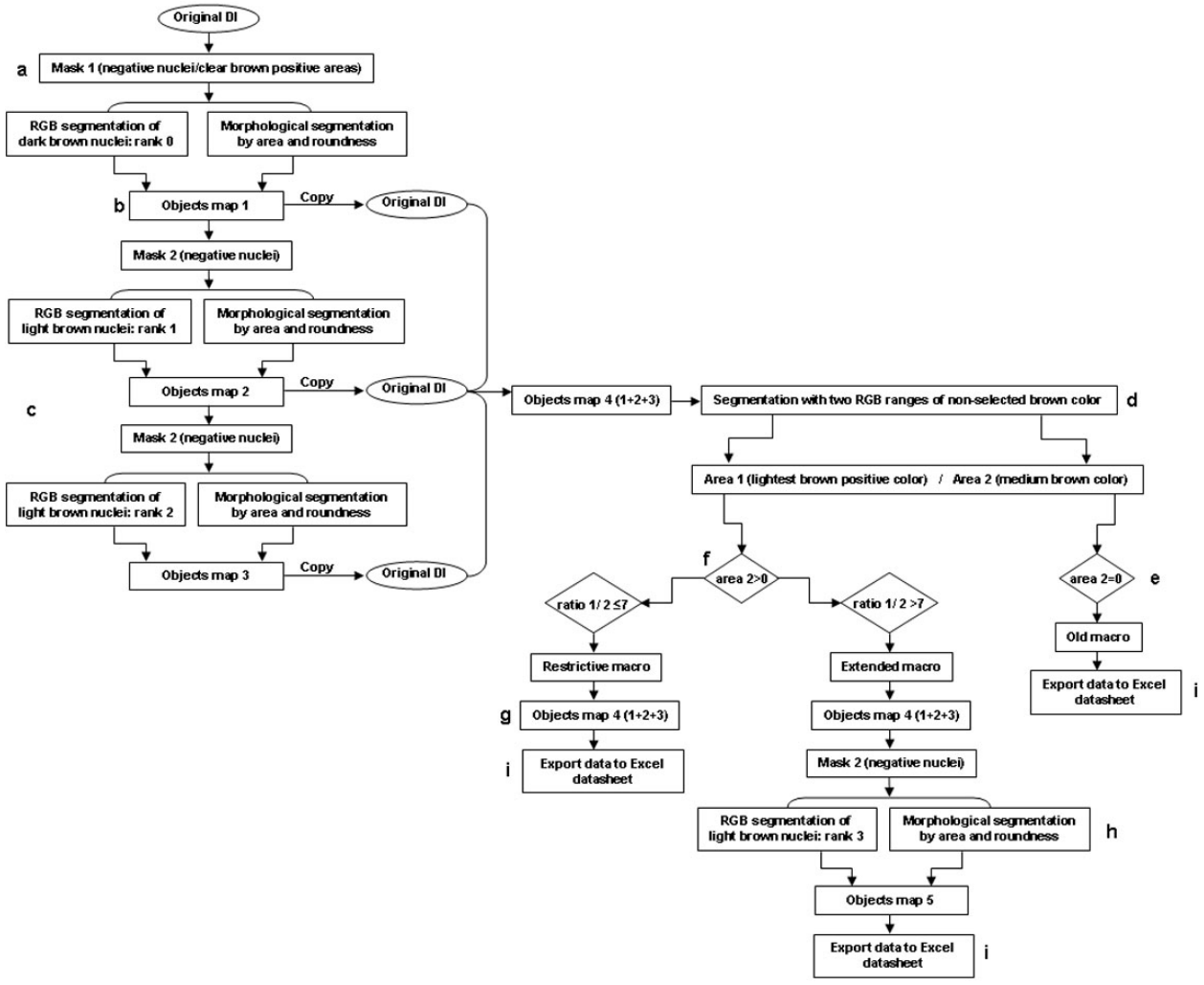



**Figure 3(abstract S13)** Schematic representation of the automated multistep procedure. Image analysis was carried out with the Image-Pro® Plus 5.0 program and data from selected positive objects are exported to an Excel datasheet in which appropriate algorithms are implemented, enabling the definitive number of positive cells to be calculated.

the two automated methods and designed the database. AR analyzed the data. ML, CL and AK drafted the manuscript. All authors revised the manuscript and approved the final version.

**References**

1. Rojo MG, Bueno G, Slodkowska J: **Review of imaging solutions for integrated quantitative immunohistochemistry in the Pathology daily practice.** *Folia Histochem Cytobiol* 2009, **47(3)**:349-354.
2. Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM: **Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy.** *Arch Pathol Lab Med* 2011, **135(2)**:233-242.
3. Lejeune M, Jaén J, Pons L, López C, Salvadó MT, Bosch R, García M, Escrivà P, Baucells J, Cugat X, Álvaro T: **Quantification of diverse subcellular immunohistochemical markers with clinicobiological relevancies: validation of a new computer-assisted image analysis procedure.** *J Anat* 2008, **212(6)**:868-878.
4. Bolton KL, Garcia-Closas M, Pfeiffer RM, Duggan MA, Howat WJ, Hewitt WJ, Yang XR, Cornelison R, Anzick SL, Meltzer P, Davis S, Lenz P, Figueroa JD, Pharoah PD, Shermand ME: **Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies.** *Cancer Epidemiol Biomarkers Prev* 2010, **19(4)**:992-999.
5. Markiewicz T, Wisniewski P, Osowski S, Patera J, Kozlowski W, Koktysz R: **Comparative analysis of methods for accurate recognition of cells through nuclei staining of Ki-67 in neuroblastoma and estrogen/progesterone status staining in breast cancer.** *Anal Quant Cytol Histol* 2009, **31(1)**:49-62.
6. Di Cataldo S, Ficarra E, Acquaviva A, Macii E: **Automated segmentation of tissue images for computerized IHC analysis.** *Comput Methods Programs in Biomed* 2010, **100(1)**:1-15.
7. Lopez C, Lejeune M, Salvado 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, **129(3)**:379-387.
8. Lopez C, Lejeune M, Escriva P, 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, **15(6)**:794-798.
9. Alvaro-Naranjo T, Lejeune M, Salvado-Usach MT, Bosch-Príncep R, Reverter-Branchat G, Jaén-Martínez J, Pons-Ferré LE: **Tumor-infiltrating cells as a prognostic factor in Hodgkin's lymphoma: a quantitative tissue microarray study in a large retrospective cohort of 267 patients.** *Leuk Lymphoma* 2005, **46(11)**:1581-1591.

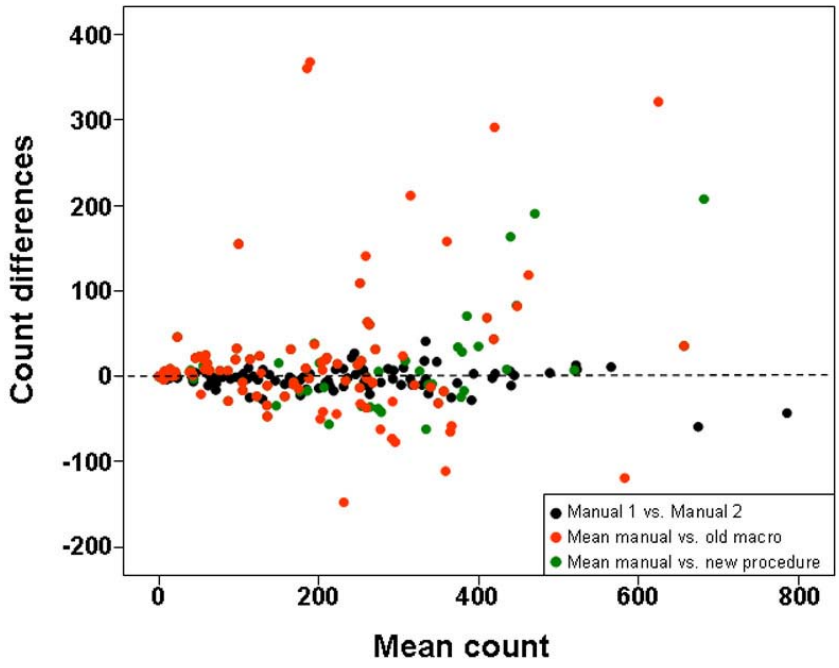



**Figure 4(abstract S13)** Bland-Altman graph showing differences between the two manual quantification methods (in black), the mean of the manual counts and the old macro results (in red) and the mean of the manual counts vs. the new procedure results (in green). The Y axis represents the difference between the results from pairs of methods, and the X axis represents the mean of both counts for each comparison.

10. Seidal T, Balaton AJ, Battifora H: **Interpretation and quantification of immunostains.** *Am J Surg Pathol* 2001, **25(9)**:1204-1207.

## S14

### Is HER2 amplification predictable by digital immunohistochemistry?

Tamás Micsik[1*], Gábor Kiszler[2], Daniel Szabó[2], László Krecsák[3], Tibor Krenács[1], Béla Molnár[2]

[1]Ist Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary; [2]3DHistech Ltd., Budapest, Hungary; [3]H-1063 Budapest, Podmaniczky u. 63, Hungary
E-mail: micsikt@gmail.com

*Diagnostic Pathology* 2013, **8(Suppl 1):**S14

**Background:** In the last decade anti-HER2 treatment became one of the best examples for targeted treatment. Since the aggressive behavior of HER2-positive breast cancers could have been successfully reduced by trastuzumab therapy, HER2 positive breast cancers recently show improving prognosis. According to a four-tiered classification of international clinical guidelines, cases with strong and complete staining (IHC 3+) with anti-HER2 antibodies are eligible for trastuzumab therapy. The cases with complete, but moderate anti-HER2 stainings (2+ or equivocal) should be further investigated with (F)ISH-technique to determine HER2-amplification [1]. Negative and IHC 3+ cases are easy to interpret semi-quantitatively on routine immunohistology, it is hard to conclude on the equivocal cases, sill, anti-HER2 therapy is indicated upon the predictive pathology report of HER2-expression and interobserver variability of IHC-interpretation still remains rather high [2]. Furthermore, the response rate of patients to the rather expensive tratuzumab therapy that might be accompanied by side effects is still only about 50% [3]. The rapidly developing digital pathology solutions have promised better ways of archiving, documenting and standardizing immunohistochemistry including image analysis of HER2 detection to improve the efficacy of targeted anti-HER2 therapy [4].
*MembraneQuant* application of *Pannoramic Viewer* platform (3DHISTECH, Budapest, Hungary) offers standardized way for semi-automated scoring of membrane-staining. Our aim was to validate *MembraneQuant* application against semi-quantitative routine scoring of HER2 IHC slides in order to improve prediction of HER2 gene amplification status.

**Materials and methods: Patients:** *We selected invasive breast cancers from year 2002-to 2005 from the* archive of the 1st Department of Pathology and Experimental Cancer Research of the Semmelweis University, Budapest, Hungary. 100 invasive ductal carcinomas were used in TMAs (created with *TMA Master*, 3DHISTECH Ltd, Budapest, Hungary) of 2mm cores of formalin-fixed paraffin embedded breast cancer specimens from females aged 26-86 years. The survey was performed with the permission of the *Ethical Committee*. TMA slides were used for HER2 IHC according to manufacturer's protocol on a Bond-max™ fully automated staining system (Leica Microsystems GmbH, Germany), using PATHWAY® HER-2/*neu* (clone 4B5, Ventana, USA), whereas their duplicates were used for HER2-FISH testing by the Rembrandt Her2/Neu - Cen 17 FISH kit (PP-C801K.5206, Biomedica kft. Budapest).
**Digitalization platform:** Digitalization was performed using a *Pannoramic SCAN* digital slide scanner (3DHISTECH) using Zeiss plan-apochromat objective (magnification: 20X, Numerical aperture: 0.8) and Hitachi (HV-F22CL) 3CCD progressive scan color camera (resolution: 0.2325 μm/pixel). For fluorescent slides Zeiss (AxioCam MRm Rev. 3) monochrome microscopy camera (resolution: 0.322500 μm/pixel) was used which has high spectral sensitivity.
**Digital evaluation of slides:** In our application a color deconvolution algorithm was applied to separate the signal of immunoreactive cell membranes in the brown channel (DAB signal) from the counterstain blue channel (hematoxilin signal). The color deconvolution algorithm generates two different grayscale images which are separately processed. The membrane detection algorithm runs on the brown, whereas the cell nuclei detection on the blue channel. The immune-negative epithelial cells have no membrane stain so these cells are to be detected on their cell nuclei on hematoxilin signal. The processing of the cell nuclei segmentation is similar to the *NuclearQuant* application which has been previously described [5].
Cell membrane immunostained slides can be described as connected pixel curve of local minima of intensity in DAB image. The intensity based linking algorithm was developed to segment the image into adjacent spots whose border potentially marks the middle of the membranes lines (skeleton). Some false curves can appear at local minima where is no actual immunreaction presented, therefore adjacent spots which have suboptimal features should be merged. Merging criteria are based on area

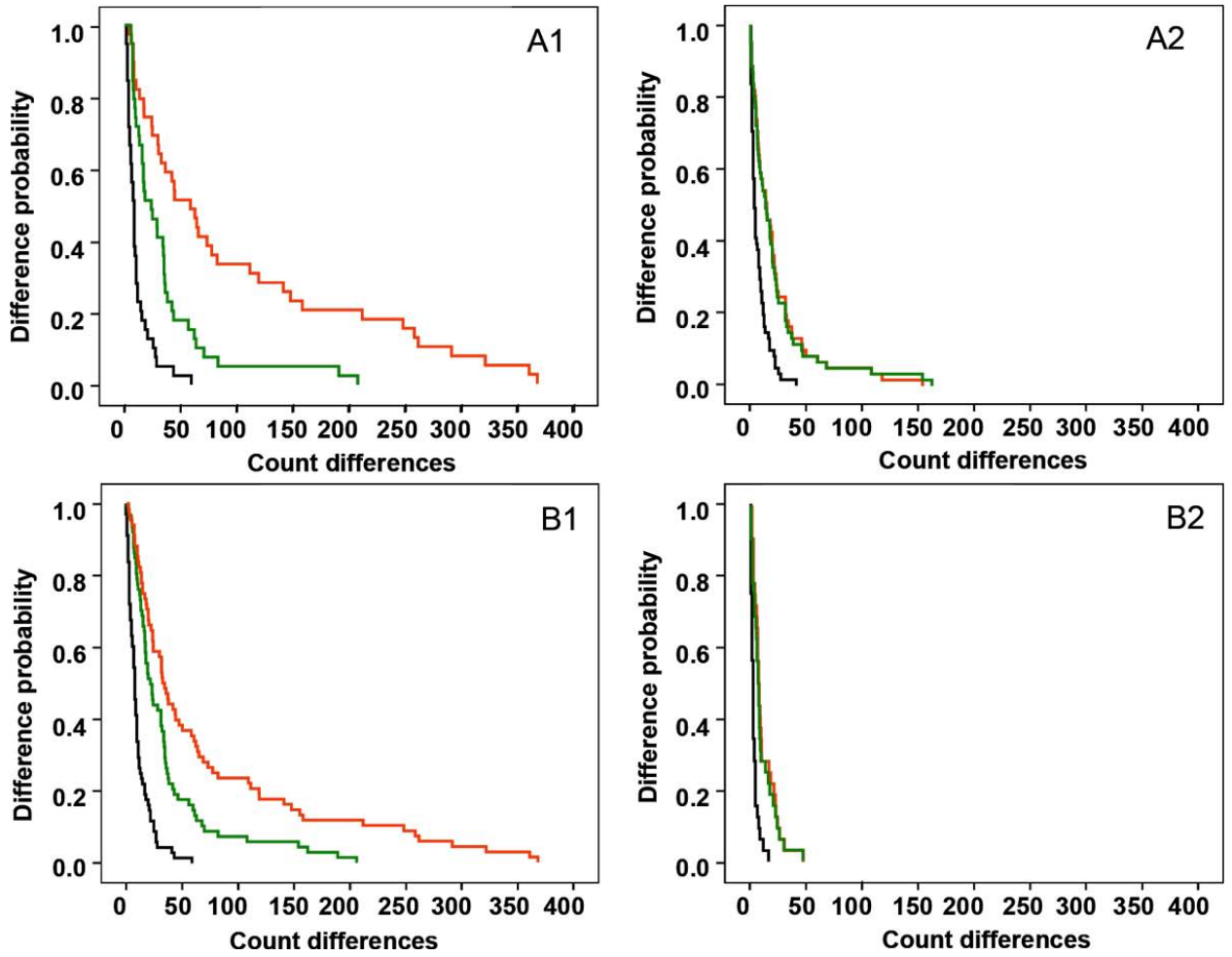



**Figure 5(abstract S13)** Superimposed Kaplan-Meier curves comparing the probability of difference between the manual reading 1 vs. manual reading 2 (in black), the mean of manual readings vs. old macro (in red) and mean of manual readings vs. new procedure (in green). The X axis represents count differences between the two methods or readings compared and the Y axis represents the probability of observing these differences. The upper row (A) represents the results obtained with DIs grouped in function of the background level: DIs with background (A1) versus DIs without background (A2). The lower row represents the results obtained with DIs grouped in function of their complexity: high-complexity DIs (B1) and low-complexity DIs (B2).

of the spots and the length of neighboring border segments. Further false curves can be eliminated based on nuclei segmentation: adjacent spots are merged which has common border segment over a nuclei. After successful subtraction of membrane and nuclei segmented mask images spots are disclosed which could not represent membranes based on their size. Average intensity of DAB image is measured along the border of spot locations to be used for classification (scoring) [6].

*MembraneQuant* detects all cells and counts individually its specific staining in a region of interest (Figure 1). These proportional and intensity data later combined to a Field Score according to the guidelines, but all other data can be extracted from the digital analysis (e.g. H-Score, Label area, mask area, number of detected objects).

**Semiquantitative-scoring of the slides:** Visual scoring of the digitized HER2 slides was performed blinded with regards of the original HER2 scores. A pathologist reviewed the digital TMA-cores using *Pannoramic Viewer* application (3DHISTECH) and provided a HER2 score and selected 1 to 4 annotations of tumorous tissue as regions of interest (ROI) on each core and scored them individually. Altogether 226 annotations were selected and analyzed using *MembraneQuant* and the scores of the ROI on one core were summarized into an overall core score, which was used for the data analysis.

FISH scoring and settings of *MembraneQuant* were equivalent to the international HER2- ASCO/CAP scoring guidelines [1]. Data analysis of the immunreaction was performed using the Statistica 9.1 (StatSoft Inc, Tulsa USA).

Agreement between the different scorings was calculated using Cohen's kappa. The strength of agreement was interpreted as proposed by Landis & Koch [7]. In order to test the clinical relevancy of the agreement, quadratic weighted kappa was calculated as well, by assigning the following weights: 1 for total agreement; 0.89 for 0 vs. 1+ or 2+ vs. 3+ or 1+ vs. 2+; 0.56 for 0 vs. 2+; 0 for 0 vs. 3+. the weight 0 for the most relevant disagreement (i.e. 0 for 2+ vs. 3+). The strength of the agreement was additionally assessed using the Spearman rank-correlation coefficient.

**Results and discussion:** Cohen's kappa revealed an almost perfect agreement ($\kappa$ = 0.857, 95%CI = 0.750-0.929) between the scores provided by the pathologists and those by *MembraneQuant*. While testing the agreement for clinical relevancy, this proved to be an almost perfect correlation, as showed by the high quadratic weighted kappa value ($\kappa$ = 0.962, 95%CI = 0.939-0.986). Spearman rank-correlation also provided a highly significant correlation between the results (Spearman's rho = 0.933, df = 99, $p < 0.0001$, 95% CI for rho 0.903-0.955).

In the 15 equivocal cases 9 were found FISH-, while 6 were FISH+. During digital processing of IHC-slides MembraenQuant calculates different values for each cell, which data were later analyzed in all the FISH tested cases in order to predict HER2 amplification status. There was a trend

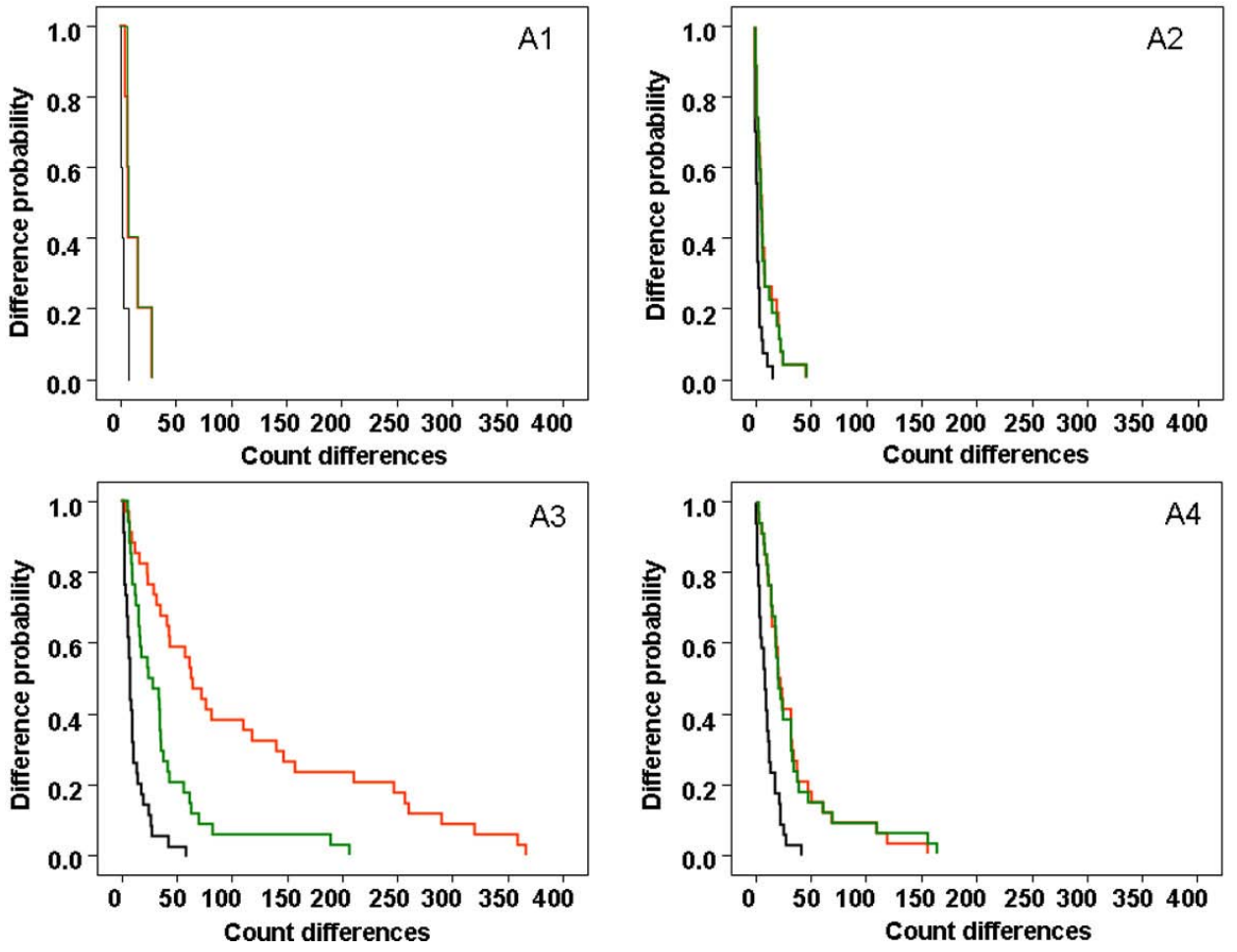



**Figure 6(abstract S13)** Kaplan-Meier curves comparing the probability of difference between the manual reading 1 vs. manual reading 2 (in black), the mean of manual readings vs. old macro (in red) and mean of manual readings vs. new procedure (in green). The X axis represents count differences between the two methods or readings compared and the Y axis represents the probability of observing these differences. DIs have been grouped according to their complexity and their background level: (A1) low-complexity DIs with background, (A2) low-complexity DIs without background, (A3) high-complexity DIs with background and (A4) high-complexity DIs without background.

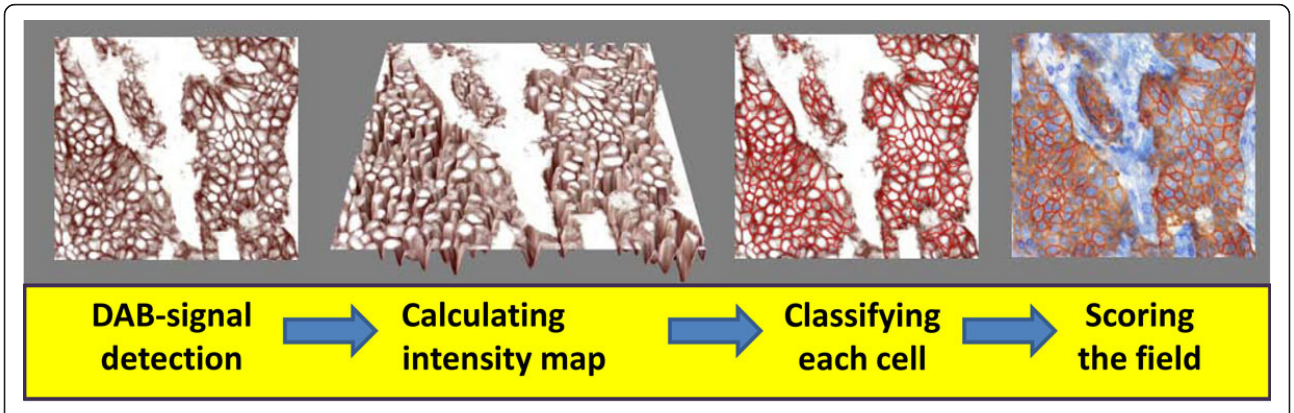



**Figure 1(abstract S14) Digital analysis of scanned slides** The DAB-imunostaining levels are detected on the brown-channel. MembraneQuant detects the nuclei on the blue channel to identify the cells of the ROI, than calculates an intensity topographic map to classify each cell to a distinct class, which is later on calculated to a field score (here the strong, complete membrane staining was of a 3+ score). On the result-slide the positivity-class of the cells are color-coded (blue: nucleus, yellow: 1+, orange: 2+, red: 3+).

**Table 1(abstract S14) Differences in membrane-stainings according to FISH-positivity**

| | FISH - | FISH + | FISH - | FISH + | FISH - | FISH + | FISH - | FISH + | FISH - | FISH + | FISH - | FISH + |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p= | *0.07097* | | 0.01782 | | *0.05574* | | 0.00031 | | 0.00166 | | 0.00486 | |
| | HER2 2+ % RATIO | HER2 2+ % RATIO | HER2 3+ % RATIO | HER2 3+ % RATIO | H-Score | H-Score | L0_CMA:1+ | L0_CMA:1+ | L0_CMA:2+ | L0_CMA:2+ | L0_CMA:3+ | L0_CMA:3+ |
| **Average** | **0.08386** | **0.152766** | **0.012607** | **0.04071** | **61.25344** | **88.80895** | **7315.882** | **25794.81** | **1317.471** | **12464.48** | **184.7127** | **2866.724** |
| SD | 0.107272 | 0.135455 | 0.024028 | 0.047413 | 42.95535 | 48.74765 | 4898.122 | 21455.78 | 1796.785 | 15429.72 | 365.2118 | 4210.592 |
| Min. | 0 | 0 | 0 | 0 | 3.007519 | 9.550562 | 207.6841 | 1472.817 | 0 | 0 | 0 | 0 |
| **Max.** | 0.28777 | 0.408517 | 0.098592 | 0.162461 | **136.6906** | **170.5047** | 19422.9 | 67658.91 | 5221.888 | 46501.73 | 1549.739 | 11813.45 |

*MembraneQuant* calculates intensity scores for each cell within an annotation and additional proportional score. Significant difference is highlighted underlining; italics indicates strong tendency. FISH+ cases had significantly higher HER2-IHC3+ proportional score, while IHER2-IHC2+ proportional score and H-score had a tendency of higher occurrence in the FISH+ cases. In our set, the most significant differences (underlined) were found between FISH-positive and negative cases in the label-area of different IHC-intensity classes.

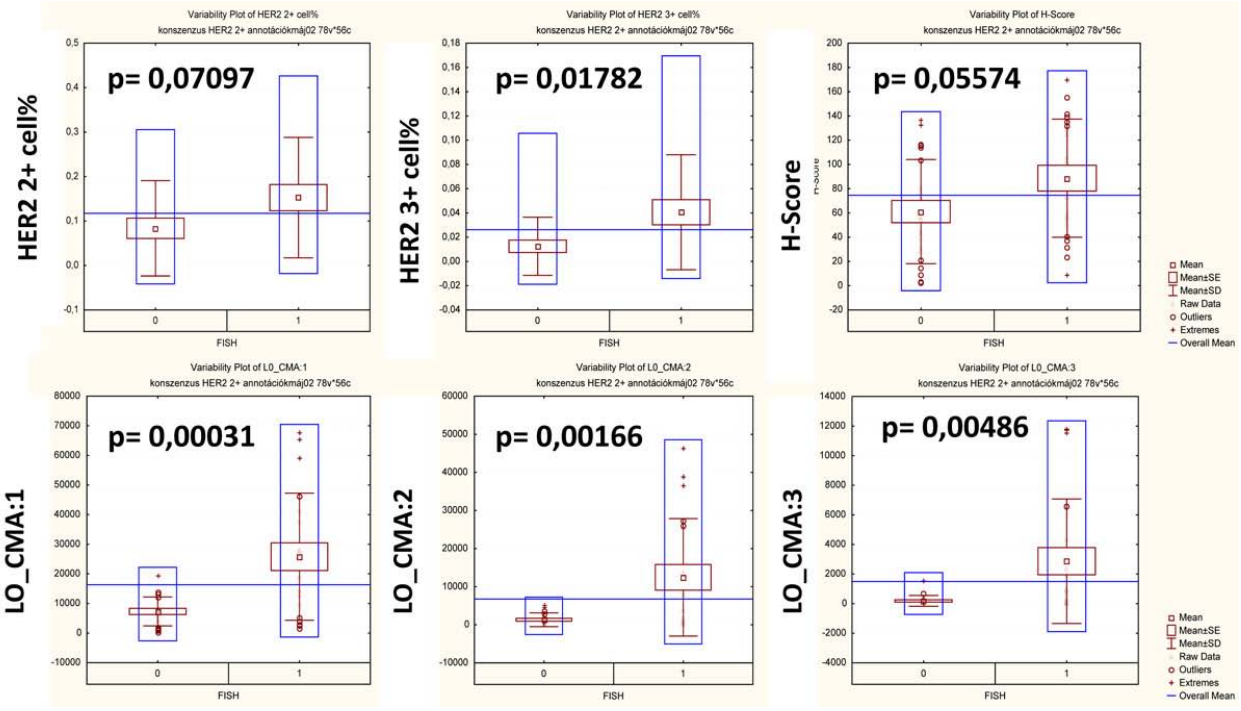



**Figure 2(abstract S14) Graphs of significant findings according to FISH-positivity.** Legend: According to data in Table 1. significant findings were visualized with variability plots and p-values are indicated in each graph. Most significant differences were found in the Label Area and within the HER2 3+ proportional score, while strong tendency was found in H-Score according to FISH-positivity. Only trend was found in HER2 2+ proportional score.

towards lower HER2-negative cell number and higher 2+ cell number in FISH-ve cases, while FISH+ve cases had significantly higher 3+ cell number. By multiplying the frequency of positive cells with the class of IHC-positivity given by the software we calculated the H-Score for all ROIs and found a borderline significant difference between the FISH+ve and FISH-ve cases with elevated H-score in the FISH+ cases. Among the data counted by the *MembraneQuant* we also found other significant differences and one of the most promising was within the 'mask area of objects in different classes' (LO-CMA) to differentiate between FISH+ve and FISH-ve cases (Table 1 and Figure 2).

The anti-HER2 targeted therapy is indicated upon strong (3+) HER2-IHC staining which correlates well with the HER gene amplification measured by FISH [6]. FISH, however, requires special infrastucture accessed by only limited number of laboratories and several studies have concluded, that protein expression level might correlate better with the efficacy of trastuzumab therapy [3]. The four-tiered classification of HER2 immunohistochemsitry may not be a sufficient thumbrule and there is an increasing amount of information suggesting that a continous HER2-score might correlate better to the response to trastuzumab therapy [8]. In our approach, the different datasets generated during the processing of digital slides might help to better differentiate markers of HER2-amplification, which may not be evaluated consistently by the pathologist's eyes during the semi-quantitative analysis. MembraneQuant application calculates the proportional score for each intensity class (0 to 3+) and also many other further derivatives on the annotated ROI areas related to our digital immunohistochemistry data, which may be used to validate HER2-IHC and its predictive value on HER2-amplification.

We found lower 0 proportional score, and higher 2+, 3+ proportional score and H-score in FISH-positive cases on a set of 15 equivocal cases further assesed with FISH. 3+ proportional score was significantly higher in FISH+ cases, and we also found more datasets (without known clinical meaning) derived by the algorythm, which showed highly significant differences in correlation with the FISH-findings. Althought, the negative and positive predictive values for reliable detection of FISH-amplification should be calculated on a higher sample number, our findings are promising since there are several other HER2 digital analysis platforms [8-10]. However, these platforms are based on more difficult and complex investigations and use more antibodies and fluorescent dyes. Our analysis relies on a standard quality IHC-reaction gained with a highly specific antibody (clone 4B5) and uses multiple factors to predict HER2-amplification.

**Conclusion:** We validated *MembraneQuant* application of *Pannoramic Viewer* platform by finding an almost perfect correlation between digital and semi-quantitative evaluation of HER2-IHC slides (quadratic $\kappa = 0.962$, Spearman's rho = 0.933). Furthermore, we found several significant differences in the staining-patterns of the equivocal cases, which could help to discriminate between FISH-positive and negative cases by combining the 4-tiered classification with other digitally-derived sample data. We strongly believe, that digital image analysis methods - digital immunohistochemistry - can improve the efficacy of anti-HER2 therapy by standardizing the evaluation protocols and finding discriminative patterns within digital data sets to detect HER2-amplification.

**List of abbreviations:** CCD: Charge Coupled Device; DAB: Diaminobenzidin; FISH: Fluorescens In Situ Hybridization; IHC: Immunohistochemistry; HER2: Human Epidermal growth factor Receptor 2; LO-CMA: Layer0 - Mask Area of objects in Class; TMA: Tissue Micro Array

**Competing interests:** Tamás Micsik and Tibor Krenács are working as consultants for 3DHistech for 5 years.
László Krecsák was formerly an employee for 3DHistech.

**Authors' contributions:** TM: Pathologist, performed immunohistochemical evaluation, wrote the majority of article,
GK: Made digitization, dealed with *MembraneQuant*
DS: optimized digitization and digital evaluation, protocol
LK: performed the majority of statistical analysis
KT: worked as consultant during work
BM: worked as consultant during work, managed company affairs

**Ackowledgements:** We are very thankful for Mrs. Renáta Kiss in helping digitization process and for Mrs Annamária Csizmadia for performing and helping to evaluate FISH-reactions.

**References**

1. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R,

Vance GH, van de Vijver M, Wheeler TM, Hayes DF: **American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.** *Arch Pathol Lab Med* 2007, **131(1)**:18-43.
2. Rhodes A, Jasani B, Anderson E, Dodson AR, Balaton AJ: **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* 2002, **118(3)**:408-17.
3. Nahta R, Shabaya S, Ozbay T, Rowe DL: **Personalizing HER2-targeted therapy in metastatic breast cancer beyond HER2 status: what we have learned from clinical specimens.** *Curr Pharmacogenomics Person Med* 2009, **7(4)**:263-274.
4. Dobson L, Conway C, Hanley A, Johnson A, Costello S, O'Grady A, Connolly Y, Magee H, O'Shea D, Jeffers M, Kay E: **Image analysis as an adjunct to manual HER-2 immunohistochemical review: a diagnostic tool to standardize interpretation.** *Histopathology* 2010, **57(1)**:27-38.
5. Krecsák L, Micsik T, Kiszler G, Krenács T, Szabó D, Jónás V, Császár G, Czuni L, Gurzó P, Ficsor L, Molnár B: **Technical note on the validation of a semi-automated image analysis software application for estrogen and progesterone receptor detection in breast cancer.**
6. Ruifrok AC, Johnston DA: **Quantification of histological staining by color deconvolution.** *Anal Quant Cytol Histol* 2001, **23**:291-299.
7. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**:159-74.
8. Giltnane JM, Molinaro A, Cheng H, Robinson A, Turbin D, Gelmon K, Huntsman D, Rimm DL: **Comparison of quantitative immunofluorescence with conventional methods for HER2/neu testing with respect to response to trastuzumab therapy in metastatic breast cancer.** *Arch Pathol Lab Med* 2008, **132(10)**:1635-47.
9. Turashvili G, Leung S, Turbin D, Montgomery K, Gilks B, West R, Carrier M, Huntsman D, Aparicio S: **Inter-observer reproducibility of HER2 immunohistochemical assessment and concordance with fluorescent in situ hybridization (FISH): pathologist assessment compared to quantitative image analysis.** *BMC Cancer* 2009, **9**:165.
10. Joshi AS, Sharangpani GM, Porter K, Keyhani S, Morrison C, Basu AS, Gholap GA, Gholap AS, Barsky SH: **Semi-automated imaging system to quantitate Her-2/neu membrane receptor immunoreactivity in human breast cancer.** *Cytometry A* 2007, **71(5)**:273-85.

## S15

### Digital immunohistochemistry: new horizons and practical solutions in breast cancer pathology

Arvydas Laurinavicius[1,2*], Justinas Besusparis[2], Justina Didziapetryte[2], Gedmante Radziuviene[1], Raimundas Meskauskas[1,2], Aida Laurinaviciene[1]
[1]National Center of Pathology, affiliate of Vilnius University Hospital Santariskiu Clinics P.Baublio 5, LT-08406 Vilnius, Lithuania; [2]Faculty of Medicine, Vilnius University, M.K.Ciurlionio 21, LT-03101 Vilnius, Lithuania
E-mail: lauar@vpc.lt
*Diagnostic Pathology* 2013, **8(Suppl 1):**S15

**Background:** Digital image analysis (DA) brings new opportunities to enhance breast cancer pathology testing by providing tools to read tissue-based visual data in a more precise, accurate, and high-throughput manner compared to traditional evaluation performed by a pathologist. The applications may vary from very practical computer-assisted diagnosis approaches to obtain more reproducible estimates of biomarker expression to more sophisticated efforts to retrieve and aggregate multi-modal and multi-dimensional data providing integrated metamarkers of the disease [1,2].
Immunohistochemistry (IHC) biomarkers are widely used for breast cancer categorization and provide basis for therapeutic decisions [3], therefore, standardization of IHC testing is of great importance. Applying DA tools for post-analytical IHC phase generates continuous data of broad dynamic range and can be used to analyse variance of the biomarker expression in more powerful statistical models and achieve better reproducibility. Also, analytical phase of IHC testing may also benefit from DA-based quality control systems. Another particular aspect of breast cancer diagnosis is related to HER2 gene and protein assays where better quantification systems are needed to resolve issues of tumours with equivocal and discordant HER2 status, potentially related to heterogeneity of tumour cell populations [4].
We hereby summarize our recent experiments to explore the advantages of DA for breast cancer pathology diagnosis and retrieval of new quality information based on IHC and FISH tests: DA-based quality monitor of tissue controls used for routine HER2 IHC testing, feasibility of DA to obtain automated cell-based HER2 FISH data, and the potential of factor analysis of the DA-generated IHC multi-marker expression data set to understand breast cancer immunoprofile variation.
**Materials and methods:** For DA-based quality monitor, IHC tissue controls represented by consecutive serial sections (n=91) of formalin-fixed paraffin-embedded multi-blocks containing 2 mm diameter cores from 4 different tumour samples (HER2 IHC score 0, 1+, 2+, 3+) were used for routine HER2 IHC staining (Ventana Benchmark XT, Ventana Medical Systems, Tucson, Arizona, USA) one control section per batch. The stained sections were scanned (Aperio ScanScope XT at 20x objective magnification) and submitted for routine quality review of the staff in charge. DA (Aperio Genie and Membrane algorithms) on each spot was performed on "per-batch" basis; the control sections were then reviewed by a pathologist to estimate potential HER2 staining intensity variation. Batch-to-batch variation of IHC average membrane staining intensity, number of tumour cells, and percentage of tumour cells with complete membranous staining was examined and compared to the visual evaluation results.
For automated HER2 FISH data, tissue microarray (TMA) 4 μm-thick sections were stained with Vysis HER2 FISH kit (as described previously [5]); 38 digital images from 19 patients with ductal breast carcinoma were obtained from TissueFAXS (TissueGnostics GmbH, Vienna, Austria) using 63x oil objective and extended focus option in 9 Z planes set at 0.45 μm interval between the planes; the Z stack images were then projected into 2D images. TissueQuest version 4 (TissueGnostics GmbH) DA with the DotFinder v.4 algorithm was used to detect the FISH signals. By using the gating feature of the TissueQuest software, analysis was restricted to cell populations with non-overlapping nuclei to obtain cell-based FISH results. Two observers performed conventional microscope FISH analysis (40 cells per patient) according to Food and Drug Administration scoring system. The same microscope and objective was used for the visual evaluation and digital image acquisition.
Factor analysis of the DA-generated IHC multi-marker expression data was performed on tissue microarrays (TMA) of ductal carcinoma samples from 109 patients stained for 10 IHC markers as recently reported [6].
Statistical analyses were performed with SAS 9.2 software.
**Results and discussion: Digital analysis-based HER2 IHC quality monitor:** DA performed on the IHC tissue multi-control sections provided continuous output data that could be efficiently summarized (Table 1) and visualized (Figure 1) to reveal variation in the serial sections of tissue multi-controls used in consecutive IHC batches. While linear plot of the mean values (for all 4 tissue samples in the multi-control) of the average membrane intensity and percentage of cells with complete membranous staining, indicated individual IHC batches with unexpected variation (Figure 1), the analyses performed on each multi-control sample uncovered even greater variation (not shown), potentially reflecting an impact of the variable content of the control tissue in the serial sections. The variation was not significantly related to neither the IHC staining time (morning/afternoon/overnight run), nor the weekday (by ANOVA, not shown). The pathologist's visual review of the controls mostly detected samples with lower than expected IHC HER2 mean intensity score (Table 1). Taking into account only 2+ and 3+ tissue controls given a lower than expected intensity score, only 3 IHC batches would have been considered as "under-stained". Interestingly, additional review of digital images of the 2+ serial sections arranged consecutively on computer monitor (Figure 2) revealed staining intensity variation, in particular, increased intensity that was missed by conventional microscope review but detected by the DA. To explore possible "long-term" drifts of the IHC sensitivity, we plotted intercepts of the parameters (Figure 3) along the consecutive tests: a mild decrease of the average membrane staining intensity was noted, further supported by corresponding weak correlation between the staining intensity and section number ($r=0.32$, $p<0.0001$). Nevertheless, the mean percentage of cells with complete membranous staining remained stable.
The DA-based HER2 IHC quality monitor provided a measurement of IHC staining quality useful in two aspects: as a quality control tool to alert on unexpected variation prospectively and as a quality improvement measure disclosing potential assay drifts based on retrospective analysis

**Table 1(abstract S15) Digital image analysis outputs and pathologist's IHC HER2 score on the IHC multi-block tissue controls**

| | Core #1 HER2 0 | Core #2 HER2 1+ | Core #3 HER2 2+ | Core #4 HER2 3+ | Cores #1-4 Mean |
|---|---|---|---|---|---|
| Mean membrane intensity* | 176 ± 16 | 190 ± 7 | 155 ± 11 | 86 ± 11 | 152 ± 8 |
| % of cells with complete membranous staining* | 1 ± 4 | 2 ± 5 | 17 ± 6 | 59 ± 6 | 19 ± 4 |
| Total number of cells* | 4277 ± 3205 | 6759 ± 3342 | 17536 ± 3382 | 19803 ± 4269 | 12094 ± 2778 |
| Pathologist's HER2 staining intensity score corresponding the expected category** | 87/88 | 73/89 | 88/91 | 84/87 | N/A |

* Mean ± standard deviation

** Data represented as a ratio of number of corresponding HER2 score/all cores evaluated. In all discordant cases, except in the category of Core#1 HER2 0, pathologist gave a lower intensity score than that expected for the category. Total number of IHC control slides evaluated was 91, except some tissue cores were considered inadequate by the pathologist (number of cores represented in the denominator).

of the data. Although the use of DA to improve analytical phase of the IHC test has been suggested [7], we were not able to find published data on practical solutions. In our study, the DA parameters analysed on "per batch" basis revealed a rather broad IHC staining intensity variation underestimated by microscope-based IHC quality control review, while only a few cases with potential quality issues were detected by additional retrospective pathologist's review.

The application of DA for IHC quality monitor is not as straightforward as it may appear. Several methodological issues have to be considered with caution. Although the same DA algorithm was used for the series, the impact of the tissue variation does not allow for a "pure" IHC tissue control system where, ideally, always an identical tissue section would be used as a reference. Although serial sections are expected to provide rather continuous changes of the tissue properties, cutting and other

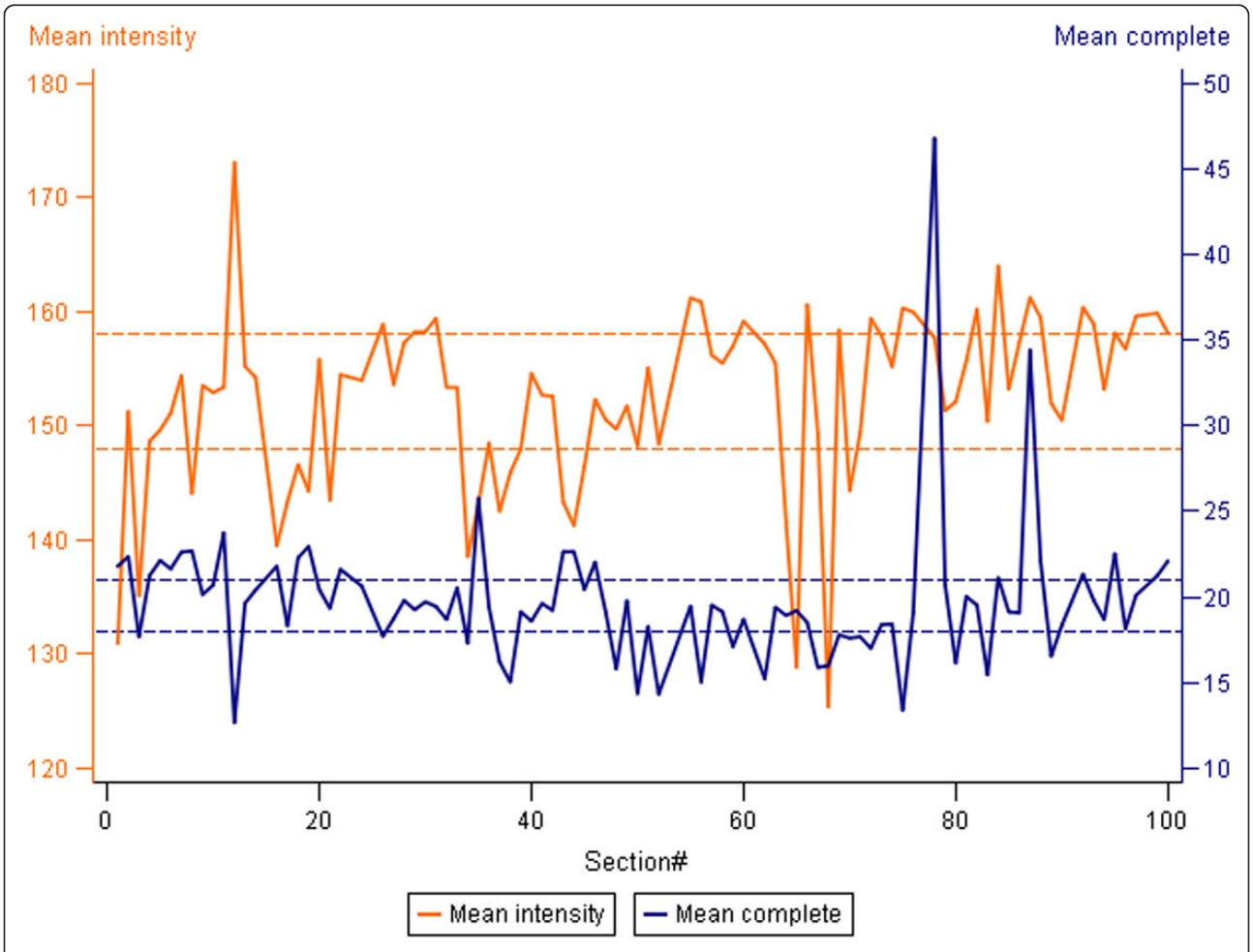



**Figure 1(abstract S15) Variation of the intensity of HER2 membranous staining and the percentage of cells with complete membranous staining in the consecutive tissue multi-control sections.** Mean intensity of the average HER2 membranous stain and Mean percentage of cells with complete membranous staining are represented by the orange and blue lines, respectively. The means are calculated from all 4 samples of the tissue multi-control. Dashed orange and blue reference lines delineate the corresponding upper and lower quartiles.

**Figure 2(abstract S15) Batch-to-batch variation of the mean staining intensity of HER2 can be noted visually on the consecutive tissue control sections arranged on computer monitor.** Nine consecutive HER2 IHC 2+ spot images are ordered from left to the right and down. Variable brown colour intensity can be noted.

artefacts may impact the DA results. Furthermore, in our experiment we used a sophisticated DA algorithm, based on automated tumour tissue recognition, which may be dependent on section thickness, hematoxylin staining variation, etc. For this matter, the use of IHC multi-controls containing 4 different tumour samples could be regarded as an "internal control" for tissue-related DA variation, however, further optimization of the approach is needed.

**Automated HER2 FISH quantification:** Strong correlation was observed between the numbers of microscope- and DA-estimated HER2 and CEP17 signals/cell ($r=0.83$, $p<0.0001$ and $r=0.68$, $p<0.002$, retrospectively), and HER2/CEP17 ratio ($r=0.71$, $p=0/0006$); the corresponding correlations between two microscope evaluations was close to perfect. However, ANOVA analysis revealed significant bias to lower values of the DA-counted HER2 signals when compared to the microscope evaluation (average difference 1.35, CI 0.4-2.3, $p<0.05$). Meanwhile, no significant difference was observed between the methods when comparing CEP17 signal counts and HER2/CEP17 ratio.

We present our initial data on the DotFinder algorithm to detect HER2 and CEP17 signals, which is followed by other TissueQuest functionalities to select populations of non-overlapping tumour cells in paraffin sections. Although the final aim of the analysis is to achieve an automated cell-based quantification of the FISH signals providing a robust and high-throughput tool to investigate tumour tissue heterogeneity, in this first stage of validation, we have explored the accuracy of the signal detection. One major obstacle to achieve full automation of the analysis is related to spatial clustering of HER2 signals in some amplified cases; other algorithms (e.g., AutoVysion, Metafer) switch to area-based estimate of the "clusters" [8,9]. This latter approach is clinically valid to detect amplified cases, however, to investigate intratumoral heterogeneity, a cell-based measurement providing actual number of FISH signals would be preferable. In our experiment, we did not include the "cluster" cases detected by conventional microscope evaluation, nevertheless, underestimation of the HER2 (but not CEP17) signal is likely to be related to spatial confluence of the signals in some HER2-amplified cells. Since the images for the analyses are produced by projecting 9 Z-stack images into one, some information allowing better signal discrimination is inevitably lost. We therefore suggest that with the increase in computing and storage capacity, 3-dimensional FISH-detection algorithms may be the basic way to progress further.

**Factor analysis of the DA-generated IHC multi-marker expression data:** We have recently published [6] a study on breast cancer TMA stained for 10 IHC markers proving a concept that important biological interdependencies can be detected at the level of tumour tissue immunophenotype based on the factor analysis of DA data. This "automated readout" of the IHC data in the TMA revealed independent biological processes standing behind the IHC profile variability in the disease entities. Integral characteristics (factor scores) of individual patients were associated with main conventional categories of the breast ductal carcinoma. In particular, we found that major factor of the IHC profile variation was characterized by a strong inverse relation between the expression of hormone (estrogen, progesteron, androgen) receptors along with anti-apoptotic marker BCL2, on one side, and Ki67 (proliferation) and HIF-1$\alpha$ (hypoxic stress, angiogenesis) on the other side. We named this factor the "i-Grade" since its pattern reflected the interdependent variance of the IHC markers known to represent the axis from aggressive (Ki67, HIF-1$\alpha$) to more indolent (hormone receptor-positive, BCL2) behaviour of the disease and was associated with the Nottingham histological grade. Also, we were able to test independent informative value of conventional and less explored IHC biomarkers and their combinations.

**Conclusions:** In summary, we report on three DA approaches expanding horizons of tissue-based breast cancer research and clinical practice.

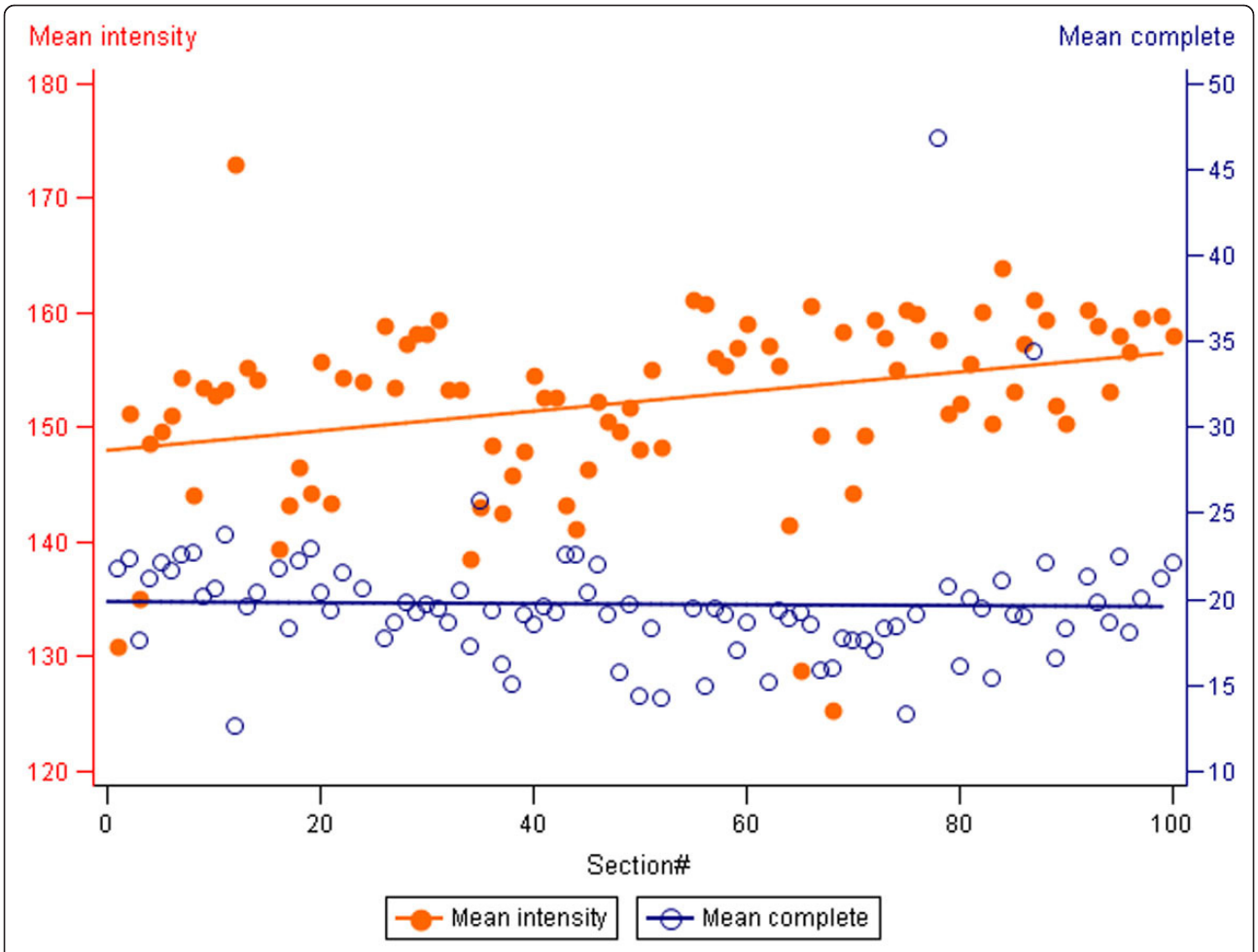



**Figure 3(abstract S15) Decreasing mean intensity of average HER2 membranous staining in the consecutive tissue multi-control sections.** Mean intensity of the average HER2 membranous stain and Mean percentage of cells with complete membranous staining are represented by the orange and blue interpolation lines and orange dots/blue circles, respectively. The values are calculated from all 4 samples of the tissue multi-control. Mean intensity of the membranous stain decreases (increasing pixel values) in the consecutive sections (r=0.32, p<0.0001).

We used standard IHC and FISH techniques and commercially available DA tools to retrieve new aspects of information that can be used to enhance quality assurance and understanding of breast cancer pathology.
**List of abbreviations:** IHC: immunohistochemistry; DA: digital image analysis; HER2: the human epidermal growth factor receptor 2; FISH: fluorescence *in situ* hybridization; TMA: Tissue microarrays.
**Competing interests:** The authors declare that they have no competing interests.
**Authors' contributions:** ArL, JB, JD, and AiL drafted the manuscript, ArL and JB performed statistical analysis, AiL and JB designed the TMA IHC digital analyses. JD scanned and performed visual and digital FISH image analyses. AiL and GR performed FISH microscope evaluation. RM performed microscope evaluation of the TMA IHC images. All authors participated in conception and design of the study, reviewing the analysis results, read and approved the final manuscript.
**Acknowledgement:** This work was partly supported by the Lithuanian Science Council Student Research Fellowship Award (JD). The authors thank Kestutis Mikalajunas for his excellent work with the Aperio Scanscope XT system and the TissueGnostics team for their professionalism and support.

**References**

1. Krenacs T, Zsakovics I, Diczhazi C, Ficsor L, Varga VS, Molnar B: **The potential of digital microscopy in breast pathology.** *Pathol Oncol Res* 2009, **15**:55-58.
2. Madabhushi A, Agner S, Basavanhally A, Doyle S, Lee G: **Computer-aided prognosis: predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data.** *Comput Med Imaging Graph* 2011, **35**:506-514.
3. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ: **Strategies for subtypes–dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011.** *Ann Oncol* 2011, **22**:1736-1747.
4. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, Kang E, Kim SW, Park SY: **Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance.** *Mod Pathol* 2012, 1-11.
5. Laurinaviciene A, Dasevicius D, Ostapenko V, Jarmalaite S, Lazutka J, Laurinavicius A: **Membrane connectivity estimated by digital image analysis of HER2 immunohistochemistry is concordant with visual scoring and fluorescence in situ hybridization results: algorithm evaluation on breast cancer tissue microarrays.** *Diagn Pathol* 2011, **6**:87.
6. Laurinavicius A, Laurinaviciene A, Ostapenko V, Dasevicius D, Jarmalaite S, Lazutka J: **Immunohistochemistry profiles of breast ductal carcinoma: factor analysis of digital image analysis data.** *Diagn Pathol* 2012, **7**:27.
7. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, *et al*: **American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and**

progesterone receptors in breast cancer. *Arch Pathol Lab Med* 2010, **134**:907-922.
8. Stevens R, Almanaseer I, Gonzalez M, Caglar D, Knudson RA, Ketterling RP, Schrock DS, Seemayer TA, Bridge JA: **Analysis of HER2 gene amplification using an automated fluorescence in situ hybridization signal enumeration system.** *J Mol Diagn* 2007, **9**:144-150.
9. Schunck C, Mohammad E: **Automated analysis of FISH-stained HER2/neu samples with Metafer.** *Methods Mol Biol* 2011, **724**:91-103.

## S16

### Automatic measurement of the evolutionary process dynamics of primary of biliary cirrhosis

Nicola Dioguardi[1*], Carlo Russo[1], Emanuela Morenghi[2], Barbara Franceschini[1], Sonia Di Biccari[1], Stefano Musardo[1], Guido Bosticco[3]
[1]Laboratory for metric method in Medicine, Istituto Clinico Humanitas IRCCS, Rozzano, Milan, Italy; [2]Biostatistic unit, Istituto Clinico Humanitas IRCCS, Rozzano, Milan, Italy; [3]Pavia University, Italy
*Diagnostic Pathology* 2013, **8(Suppl 1):**S16

**Background:** PBC is a chronic inflammatory autoimmune disease [1] that evolves into cirrhosis via four stages. These are determined by the successive implementation in hepatic tissue of the following events: inflammation, destruction and regeneration of biliary tissue CK7+ and fibrosis. This paper describes a method to study of the dynamics of the disease histologic behaviour [2].

**Materials and methods:** We studied 58 liver biopsy samples, from the archives of Departments of Pathological Anatomy of Istituto Clinico Humanitas IRCCS, Rozzano, Milan, Italy and the Department of Gastroenterology of Milan's Ospedale Policlinico IRCCS, of Caucasian subjects with PBC in various stages. Each procedure was performed in accordance with the guidelines of the Ethics Committees of the hospitals involved and diagnoses were defined by the expert pathologists of these same hospitals.

Three consecutive 2-3 μm thick sections were obtained from each sample: the first was used to identify inflammatory cells; the second was stained with Sirius red to visualize collagen deposits and the third was used to visualize the biliary tree with CK7 antibody.

The histological metrization method [3] was developed in the Laboratory for the Study of Metric Methods in Medicine of the Istituto Clinico Humanitas (Rozzano, Milan, Italy). The prototype of this computerized device called HM and its controlling software were entirely designed and developed in our laboratory. The Metrizer is a totally computer-driven machine that automates the focusing of the microscope and the exclusion of Glisson's capsule from the computation of fibrotic islets. It performs reproducible metric measurements from digitalized images of the entire histological section, giving results within a few minutes. This method measures all notable liver structures bared of any property not involved in the measurement.

**Statistical analysis:** The data are given as numbers and percentages or mean or median values and ranges where appropriate. All variables were log-transformed in order to approximate a Gaussian distribution and normalized to a (0,1) range to allow for comparison between variables. Thus the final data are expressed in percentage of the interval between 0 and 1. All of the analyses were performed with Stata10 [http://www.stata.com].

**Results and discussion:** ***Inflammation*** (Figure1A) was identified by defining the borders around clusters of mononuclear cells (lymphocytes) with Delaunay's triangulation method [4].

This method identifies a line that connects the centers of outermost cells that are distanced ≤20 μm. This line is arbitrarily considered the separator of the inflammatory cells within a cluster from the mononuclear cells dispersed in the surrounding hepatic tissue.

The *inflammatory basin*, which increases constantly due to the autoimmune process, is the set of areas in a section that are covered by inflammatory cell clusters. The area of the basin covered by cluster-resident cells is called the *pure inflammatory space*, which varies with the number (density) of cells present in the clusters.

The ***intra-hepatic biliary duct*** area (Figure 1B) was metrically measured. During the course of PBC, the biliary tree duct status is determined by two components: autoimmune destruction and regeneration of intrahepatic CK7+ ductular segment.

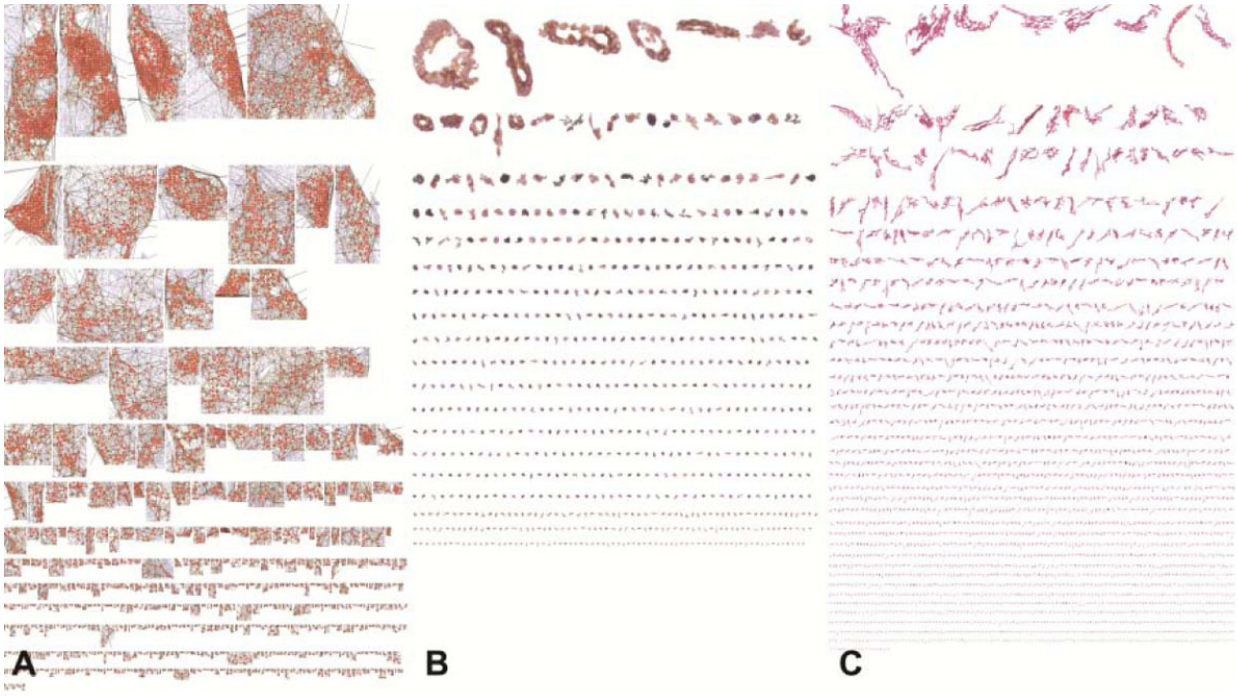



**Figure 1(abstract S16)** Catalogue of the multifarious fragments of liver structures measured by HM.; **A.** Inflammation cell clusters **B.** CK7; and **C.** fibrosis elements, The black image is a Glisson membrane fragment excluded by the Metrizer.

The ***collagen islets*** forming fibrosis (Figure 1-C) were measured in linear meters; the result was corrected by the fractal dimension [5] to include details of the irregularity of their shapes The fractal dimension was obtained by means of the box counting method because the objects to be measured were "truncated fractals" [6], the fractal dimension was used as a dilation factor rather than an exponent [7] Three classes of islet magnitude were arbitrarily identified: area from 10 and $10^3$ $\mu m^2$, from $10^3$ to $10^4$ $\mu m^2$ and over $10^4$ $\mu m^2$.

The ***tissue disorder*** quantitative assessment was performed with a Tectonic Index (TI), which describes the loss of tissue organization or any deviation from the natural order (a high TI indicates a high degree of tissue disorder). The TI defines the organization of liver tissues and is expressed with a range of values from 0 to 1. The TI was calculated as 1 – H, where H is Hurst's coefficient (range zero to 1) [8], defined as $D_\gamma$ + 1 – D, where D is the fractal dimension and $D_\gamma$ the Euclidean dimension of the observed object. So TI = 1 – H = D – $D_\gamma$. Numerical results of all of these parameters are summarized in Table 1.

In order to approach the PBC behavioural dynamics [9], the first key was to set into the interval (0,1) the logarithmic transformation and normalization of the measures of CK7+; tissue and fibrosis, taken as the most representative structural elements with which to metrically construct the liver state portrait that defines the grading and staging of the semiquantitative subjective methodologies.

A second key is the transformation of portrait scalars into a single vector to reduce the multiplicity of elements of the liver section into a dot-like geometrical figure. This translation into vectorial arithmetic in classic dynamics is crucial for the construction of the dot-like geometrical operator, called dynamic *particola* that represents the whole system.

In order to gain a method to study the system's behaviour dynamics [9], the immune CK7+ biliary ductules resulting from the destruction-regeneration and intrahepatic fibrosis, considered a set of phenomena that generates the irreversible tectonic disorder leading to cirrhosis. The values of vectors representing irreversible parameters, resulting from the HM measurement, were taken to create the *particola*. The vectorial transformation is obtained by plotting the value of CK7 area on the *y* and the value of fibrosis areas on the *x* orthogonal axes. The value of the modulus of this resultant sum in the orthogonal space is taken as the Newtonian dot-like dynamic *particola* that shows an instant of the PBC behaviour (Figure 2A). The set of points (each a *particola*) on the x-y orthogonal space, resembles a Gibbs cloud of points, each expressing in log scale the magnitude of the vectorial value of a *particola* with a scalar. These scalars ordered on a real number line, the simplest state space, will graphically show the trajectory that is the reference of the dynamic behavior of PBC process. (Figure 2B).

Furthermore, we plotted the value of each particula on the y axis (ranging from 0 to 1) versus the disease timing, identified with all particulae, on the x axis (Figure 2C). Next, we identified the tertiles on this curve that discriminate three phases. This term used in our new method is different from the 'stages' characterising semiquantitative static descriptions. As a result, the three phases are extrapolated by increasing amounts of particolae formed by fibrosis and CK7+ tissue. In particular, phases 1, 2, and 3 represent early, intermediate, and final disease to feature mean particula values of 0.2075, 0.4722, and 0.7386, respectively.

The inflammation was excluded from the constitutive components of the trajectory describing the course of PBC process, for its reversibility. This different behaviour is due to the entropy produced in the interior of the inflammation process that is not transferred across its boundaries into the surrounding environment [10].

**Conclusions:** The method we constructed, with its technology, strongly reduced the computational time and improved the liver tissue structure recognition and description. As these tools allowed this first study of PBC behaviour in terms of physics of dynamics, this paper supports the hypothesis that the long periods of cessation in the history of dynamic knowledge were due to the very higher rapid theoretic development toward facilitating computation devices [11,12]. The technology introduced into our methodology facilitated the study of PBC behaviour by the following points:

1) Exclusion of the inflammatory infiltrates from dynamic study, as they do not produce topic stable entropy deposits. Collagen interstitial deposition generating fibrosis is hepato-cellular necrosis dependent.

2) Standardization of a strictly objective evaluation producing scalars by metrizing the images of the histological structures of the liver section.

3) Metrical measurement also of the smallest dispersed islets of fibrosis, and tiny CK7+ biliary ducts normally undetected with optical microscope.

4) Transformation of scalars into vectors leading to vectorial PBC histological section portraits. A homogeneity was created with this operation that maintains the concepts of mixture and identity of the mixed elements in the sum of CK7+ biliary ducts and fibrosis vectors that define the *particola*, geometric figure which graphically will trace the process trajectory.

5) Description of PBC evolution by the cumulative curve of the *particolae*, each representing an actual state of the processes and taking this planar curve as the ideal finite trajectory $\alpha$–$\omega$) of entire PBC process.

6) Definition of past and future percentage of the disease course with its *particola* on the trajectory.

7) The score of PBC evolution into three phases on its trajectory and describe the ideal course of the process based on mathematical measurements.

**Table 1(abstract S16) Summary of all of the metric data obtained from the structural measurements and the data used for staging purposes. Minimum, median, and maximum values of tissue parameters are given in % of total histologic section area of biopsy specimen**

| Tissue parameters | Minimum | Median | Maximum |
|---|---|---|---|
| ***Area of the inflammatory basin*** | 0.026 | 1.005 | 6.444 |
| Area covered by resident cells | 0.002 | 0.636 | 5.295 |
| ***Area of intra-hepatic biliary ducts*** | 0.003 | 0.114 | 0.882 |
| Small areas (10-$10^3$ μm) | 0 | 0.009 | 0.105 |
| Medium-sized areas ($10^3$-$10^4$ μm) | 0.001 | 0.006 | 0.033 |
| Large areas (>$10^4$ μm) | 0 | 0 | 0 |
| ***Area of fibrosis*** | 0.30 | 1.456 | 24.4 |
| Small areas (10-$10^3$ μm) | 0.008 | 0.092 | 0.521 |
| Medium-sized areas ($10^3$-$10^4$ μm) | 0.001 | 0.026 | 0.203 |
| Large areas (>$10^4$ μm) | 0.01 | 0.075 | 1.005 |
| ***Neo-angiogenesis*** | 0.030 | 0.586 | 3.474 |
| ***Hemopoietic stem cells*** | 0 | 0.006 | 0? |
| ***Tissue disorder (TI)*** | 0.072 | 0.274 | 0.643 |

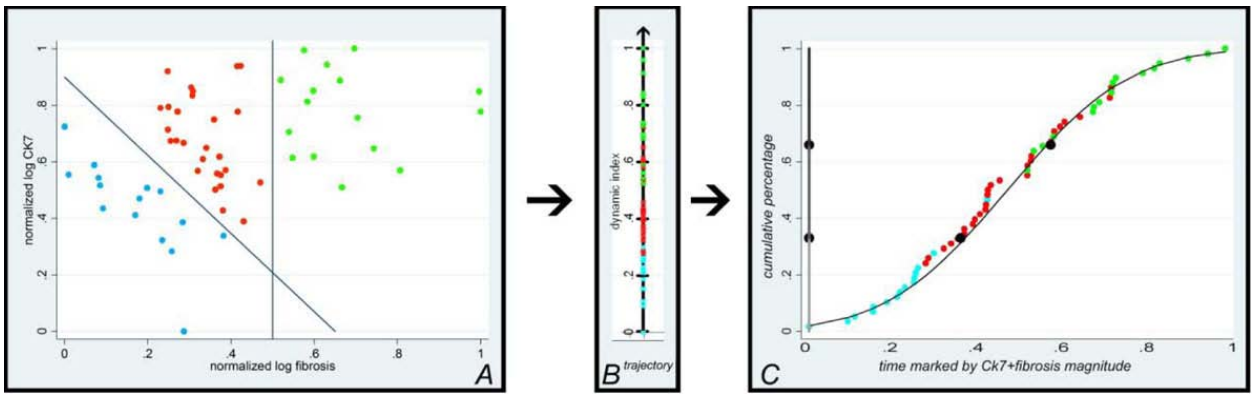



**Figure 2(abstract S16)** A): The set of dynamic *particulae* resembling a Gibbs cloud-point obtained expressing as vector magnitudes the normalised log value of CK7+ areas and the normalised log value of fibrosis in our patients. B) The points are transferred to the oriented line of real numbers representing the standardised trajectory of the process. C) An ogival cumulative curve orders the dynamic *particula* within the three clouds in panel A. This curve is sub-divided into three tertiles marked by blue points and represents the trajectory of the overall dynamic process of PBC from $\alpha$ to $\omega$.

To conclude let us say that any correlation is recognizable between the metrical data of our case-list, divided into three phases by the rules of dynamics and the semi-quantitative data, divided into four stages according to the rules of the method of Scheuer. (Table 2) For example the 20 specimen classified in our higher phase III included 9 patients (45%) classified by the semi-quantitative Scheuer classification at minor levels of staging. Where is the truth?

**List of abbreviations:** PBC: Primary biliary cirrhosis; IRCCS: Institute for Scientific Research Hospitalisation and Care; HM: Histologic Metrizer; TI: Tectonic Index

**Competing interests:** No conflicts of interest exists.

**Authors' contributions:** ND wrote the manuscript and ideated the theory around the machine and dynamics and the machine itself, CR ideated the software and constructed the machine, EM did the statistical analysis and contributed to revision of the text, BF, SDB and SM did the histological preparations, GB revised the text.

**Acknowledgements:** This study was completed with the financial support of the Fondazione Michele Rodriguez, Istituto Clinico Humanitas IRCCS, and Attilio and Livia Pagani.

The entire group is extremely grateful to Rosalind Roberts for her numerous translations of the manuscripts.

**References**

1. Rubin E, Schüffner F, Popper H: **Primary biliary cirrhosis. Chronic non-suppurative destructive cholangitis.** *Am J Pathol* 1965, **46**:387-407.
2. Dioguardi N: **hypothesis for a new method to measure the dynamic patterns of tissue injury.** *Med Hypotheses* 2011, **77(6)**:1022-7.
3. Dioguardi N, Grizzi F, Fiamengo B, Russo C: **Metrically measuring liver biopsy: a chronic hepatitis B and C computer-aided morphologic description.** *World J Gastroenterol* 2008, **14**:7335-7344.
4. Delaunay BN: **Sur la sphère vide.** Math and Nat Sci Division, Bulletin of the Academy of Sciences of USSR 1934, **6**:793.
5. Dioguardi N, Franceschini B, Aletti G, Russo C, Grizzi F: **Fractal dimension rectified meter for quantification of liver fibrosis and other irregular microscopic objects.** *Anal Quant Cytol Histol* 2003, **25**:312-320.
6. Rigaut JP, Schoëvaërt-Brossault D, Downs AM, Landini G: **Asymptotic fractals in the context of grey-scale images.** *J Microsc* 1998, **189**:57-63.
7. Dioguardi N, Grizzi F, Bossi P, Roncalli M: **Fractal and spectral dimension analysis of liver fibrosis in needle biopsy specimens.** *Anal. Quant. Cytol. Histol* 1999, **21**:262-266.
8. Hurst HE: **Long term storage capacity of reservoirs.** *Trans Amer Soc Civ Engrs* 1951, **116**:770-808.
9. Abraham R, Shaw R: **Dynamics. The geometry of behaviour. Part 2. Periodic behaviour.** Aerial Press, Inc Santa Cruz, California, third printing 1985.
10. **From being to becoming: Time and complexity in the physical sciences.** San Francisco, CA, WH freeman and Co: Prigogine I 1980.
11. Samburski S: **The physical world of the Greeks.** Routlerge and Kegan Paul London 1956.
12. Russel B: **Storia della filosophia occidentale.** TEA, Seventh essay 1910.

**Table 2(abstract S16) Distribution in the three stadia of the dynamic trajectory of the patients'list classified according to Scheuer's classification**

| | | Scheuer classification | | | |
|---|---|---|---|---|---|
| **Phase** | | **1** | **2** | **3** | **4** |
| **I** | **19 (33%)** | 11 (58%) | 7 (37%) | 1 (5%) | 0 |
| **II** | **19 (33%)** | 11 (58%) | 1 (5%) | 6 (32%) | 1 (5%) |
| **III** | **20 (34%)** | 9 (45%) | 2 (10%) | 5 (25%) | 4 (20%) |

## S17

### The Metrizer: an innovative device for achieving virtual hepatic biopsies

Carlo Russo[1], Barbara Franceschini[1], Sonia Di Biccari[1], Stefano Musardo[1], Guido Bosticco[2], Nicola Dioguardi[1*]
[1]Laboratory of Quantitative Medicine, Istituto Clinico Humanitas IRCCS, Rozzano, Milan, Italy; [2]University of Pavia, Pavia, Italy
E-mail: nicola.dioguardi@humanitas.it
*Diagnostic Pathology* 2013, **8(Suppl 1):**S17

**Background:** The last few years have brought rapid growth in the number of Virtual Microscopes (VM) and the promotion of digital histology. In this digital era, it is natural to assist in the transportation of histology imaging into computer technology support. We have also assisted in a quick and precise race towards improved technology capable of acquiring detailed digital images which realistically report the original slide for easy consultation [1,2].

The advanced technology herein proposed aims at being a sophisticated imaging archive. Digitally scanning slides however, does not give additional information to the vision of tissue structures, even with high resolution or improved colour and image precision. We can now say that high enlargement microscopic observation can only give same details to structures visible to the eye directly through the microscope. This means that when VMs are used in a correct manner they are capable of giving significant progress to highlighting, not quantifying, a critical field in applying technology specific to computerised measuring. The potential to improve precision and objectivity of measurements can be achieved with additional technical equipment. The aim of this study is to present a machine invented with a calculation potential to facilitate the work of the observer in a medical practice, not only in terms of easy retrieval of images but also as an instrument for the automatic analysis of digital histology.

The "Metrizer" aims at supplying precise and objective descriptions and measurements of the specimen under observation. These descriptions do not substitute the pathologist but they should assist him/her to assert with objectivity and safety the entity of the observed pathology [3-8].

**Material and methods:** The "Metrizer" (Figure 1) is an automatic, compact machine composed of a lens for microscopic observation, digital cameras, a movement device and a computer complete with software to control machine movement and image analysis. With these components the machine facilitates consultation of histological preparations (virtual microscope - slide scanner function) and supplies objective numeric data regarding the state of the tissue harbouring diverse diseases without causing fatigue or human error.

Using the Slide Scanner function it enables to automatically digitalise the entire histological slide in high definition and makes it easily accessible in the digital archive.

Furthermore the Metrizer gives: 1) custom image analysis algorithms, 2) objective data obtained from elements present in the tissue, and 3) what can be called an Automatic Diagnosis. This diagnosis, preset for each pathology, is generated by an expert system shaped on a series of data that has been preloaded once diagnosis is known. The expert system referral cut-off is then calculated from these data. The automatic diagnosis permits to completely abolish operator subjectivity, from image capture to the print-out of final report (Figure 2).

**Testing:** The validation process of the Metrizer was made by comparing measurements obtained with this machine on histological slides taken from a series of 120 patients previously achieved by the automatic reading software of the preliminary model of automatic analysis, consisting of a motorised microscope Leica DMLA [6].

**Model validation:** Complete histological specimen of hepatic biopsies of 235 patients affected by chronic inflammation were acquired, analyzed and compared with clinical data and semiquantitative evaluations. By arranging data and comparing them with the main methods of evaluation in use, cut-offs were located which enabled to opportunely delineate the expert system. From these retrospective cut-offs descriptions of a further 200 patients, making a total of 435 cases analysed with the so-called 'Model', were prospectively generated and verified [7].

**Validation of the Metrizer:** 120 cases coming from the validation series of the Model, which necessitated several months of work on the old system, were then examined in just a few days by the Metrizer and a high level of correlation was reached. Proof of the validity of this new, ultra rapid machine was had by comparing the groups of new data with those obtained by the already validated but much slower semiautomatic Model. In fact the Metrizer carries out all the steps in a few minutes without involving the operator, so that the method which was previously only used for research can be really applied to routine clinical practice in support of the evaluations of the pathologist.

**Results:** Comparison of data obtained with the Model with those of new data rapidly achieved by the Metrizer was made on the basis of percentage area values and tectonic index of the fibrosis, and of the inflammatory basin located in the biopsies of the 120 cases available. A double evaluation is thus obtained, from both the old and the new system. The catalogues of located objects in the two different systems were also compared.

The values are very sensitive to intense light and magnifying lens, however by correctly calibrating the machine it is possible to achieve reproducible data with very low margin of instrumental error. The comparison between

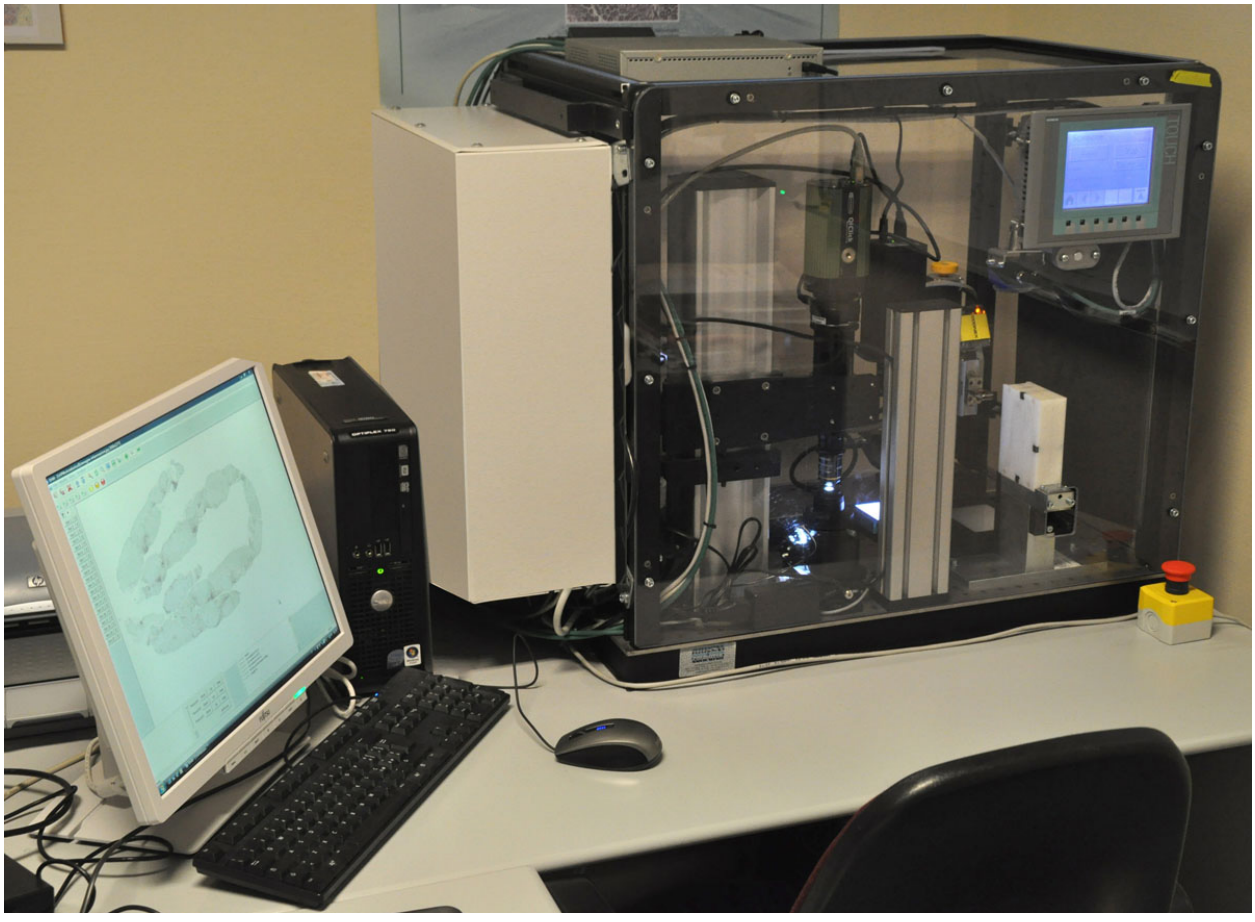
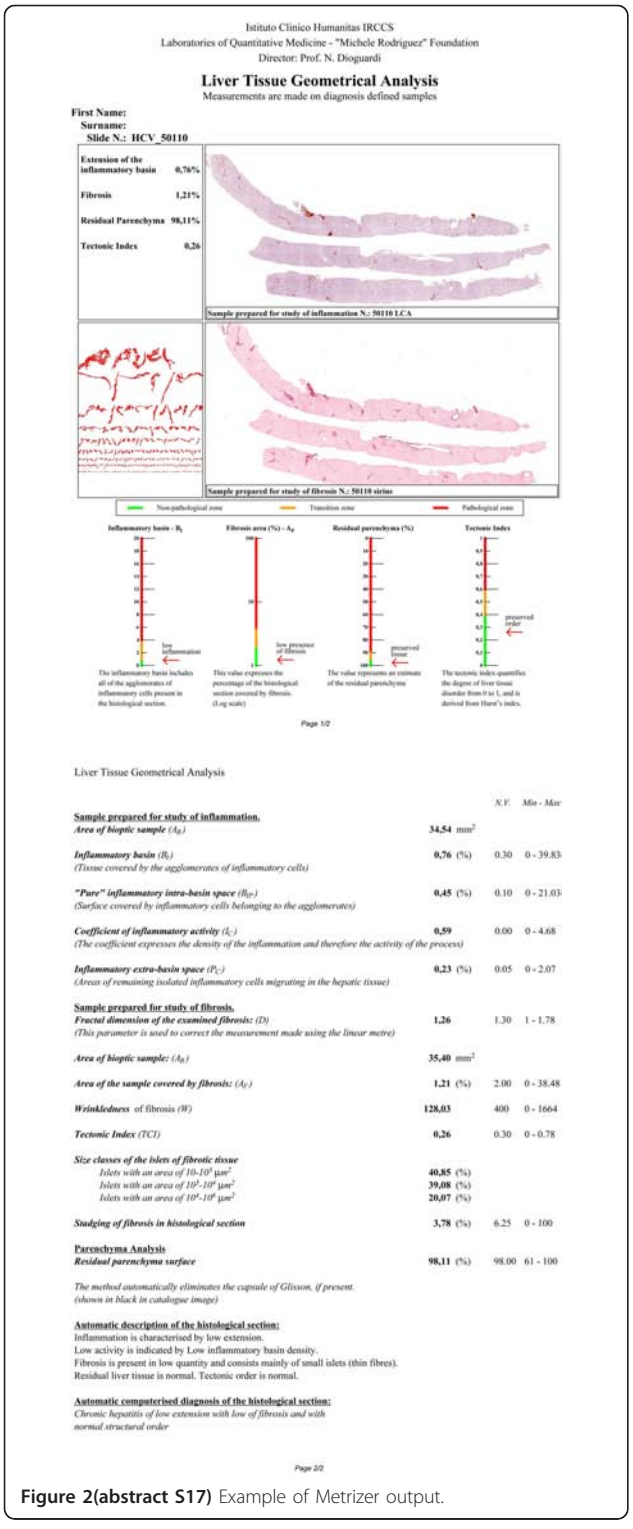


**Figure 1(abstract S17) The Metrizer** Prototypical version of the Metrizer, used in this work and for presentation purposes.



Istituto Clinico Humanitas IRCCS
Laboratories of Quantitative Medicine - "Michele Rodriguez" Foundation
Director: Prof. N. Dioguardi

**Liver Tissue Geometrical Analysis**
Measurements are made on diagnosis defined samples

First Name:
Surname:
Slide N.: HCV_50110

Estension of the inflammatory basin 0,76%
Fibrosis 1,21%
Residual Parenchyma 98,11%
Tectonic Index 0,26

Sample prepared for study of inflammation N.: 50110 LCA
Sample prepared for study of fibrosis N.: 50110 sirius

Non-pathological zone | Transition zone | Pathological zone

Inflammatory basin - $B_I$ | Fibrosis area (%) - $A_F$ | Residual parenchyma (%) | Tectonic Index

low inflammation | low presence of fibrosis | preserved tissue | preserved order

The inflammatory basin includes all of the agglomerates of inflammatory cells present in the histological section.
This value expresses the percentage of the histological section covered by fibrosis. (Log scale)
The value represents an estimate of the residual parenchyma
The tectonic index quantifies the degree of liver tissue disorder from 0 to 1, and is derived from Hurst's index.

Page 1/2

Liver Tissue Geometrical Analysis

| | | N.V. | Min - Max |
|---|---|---|---|
| **Sample prepared for study of inflammation.** | | | |
| ***Area of bioptic sample*** $(A_B)$ | **34,54** $mm^2$ | | |
| ***Inflammatory basin*** $(B_I)$ *(Tissue covered by the agglomerates of inflammatory cells)* | **0,76** (%) | 0.30 | 0 - 39.83 |
| ***"Pure" inflammatory intra-basin space*** $(B_{IP})$ *(Surface covered by inflammatory cells belonging to the agglomerates)* | **0,45** (%) | 0.10 | 0 - 21.03 |
| ***Coefficient of inflammatory activity*** $(I_C)$ *(The coefficient expresses the density of the inflammation and therefore the activity of the process)* | **0,59** | 0.00 | 0 - 4.68 |
| ***Inflammatory extra-basin space*** $(P_{IC})$ *(Areas of remaining isolated inflammatory cells migrating in the hepatic tissue)* | **0,23** (%) | 0.05 | 0 - 2.07 |
| **Sample prepared for study of fibrosis.** | | | |
| ***Fractal dimension of the examined fibrosis:*** *(D) (This parameter is used to correct the measurement made using the linear metre)* | **1,26** | 1.30 | 1 - 1.78 |
| ***Area of bioptic sample:*** $(A_B)$ | **35,40** $mm^2$ | | |
| ***Area of the sample covered by fibrosis:*** $(A_F)$ | **1,21** (%) | 2.00 | 0 - 38.48 |
| ***Wrinkledness*** of fibrosis *(W)* | **128,03** | 400 | 0 - 1664 |
| ***Tectonic Index*** *(TCI)* | **0,26** | 0.30 | 0 - 0.78 |
| ***Size classes of the Islets of fibrotic tissue*** | | | |
| *Islets with an area of 10-$10^3$ $\mu m^2$* | **40,85** (%) | | |
| *Islets with an area of $10^3$-$10^4$ $\mu m^2$* | **39,08** (%) | | |
| *Islets with an area of $10^4$-$10^6$ $\mu m^2$* | **20,07** (%) | | |
| ***Sludging of fibrosis in histological section*** | **3,78** (%) | 6.25 | 0 - 100 |
| **Parenchyma Analysis** | | | |
| ***Residual parenchyma surface*** | **98,11** (%) | 98.00 | 61 - 100 |

*The method automatically eliminates the capsule of Glisson, if present. (shown in black in catalogue image)*

**Automatic description of the histological section:**
Inflammation is characterised by low extension.
Low activity is indicated by Low inflammatory basin density.
Fibrosis is present in low quantity and consists mainly of small islets (thin fibres).
Residual liver tissue is normal. Tectonic order is normal.

**Automatic computerised diagnosis of the histological section:**
*Chronic hepatitis of low extension with low of fibrosis and with normal structural order*

Page 2/2

**Figure 2(abstract S17)** Example of Metrizer output.

the two machines can be seen in Table 1. Correlation coefficient between the two sets of data are R=0,94 in fibrosis quantification and R=0,96 in the case of inflammatory basin.

Histological Metrizer (shortened to Metrizer) described herein, is a machine capable of automatically measuring the state of opportunely stained histological specimens in a decidedly more accurate manner than all current methods of evaluation.

**Table 1(abstract S17) Model-Metrizer Comparison**

| | A% Fibrosis | | A% Inflammation | |
|---|---|---|---|---|
| **Series** | **Model** | **Metrizer** | **Model** | **Metrizer** |
| **Min** | 0,22% | 0,35% | 0,02% | 0,04% |
| **Max** | 8,32% | 14,44% | 11,77% | 11,48% |
| **Mean** | 2,23% | 2,99% | 1,87% | 1,88% |
| **St. Dev.** | 1,69% | 2,33% | 2,13% | 2,14% |
| **Relative Median Error (Log):** | 9,14% | | 7,23% | |

The table shows that the relative median difference of the calculated values was less than 10%. The highest error can be found comparing cases of fibrosis, which reaches 9,14%. This is due to the diverse composition of the CCD of the Model's cameras compared with those of the Metrizer, which loses some details in fibrosis. The video camera of the Metrizer have an higher colour definition so as to take even smaller details.

The application of quantitative methods in a medical setting faces numerous difficulties because these disciplines involve natural, complex and irregular objects and phenomena, many with characteristics that lead them to be classified as fractal objects [9,10]. Their evaluation describable with traditional linear mathematics and Euclidean geometry, can at times give results that are so distant from reality as to seem caricatures rather than metric descriptions of the object under examination.

However, these objects can be described and, with suitable non-Euclidean geometry, can also be measured with automated calculus technology for operational complexity that requires measurement of their complex structure and interpreted with basics of physics different from those of statics on which the greater part of medical disciplines is founded.

In order to obtain measurements as close as possible to reality, our Laboratory of Quantitative Medicine uses fractal geometry to correct the metric measurements of irregular objects.

**Conclusions:** The Metrizer changes the method of examining liver biopsies since it resolves fundamental problems to measure microscopic structures. For the first time it enables evaluation with a scalar number, in an absolutely objective and repeatable manner, the size of the structures that determine the state of the diseased liver. It must be stressed that histological diagnosis of hepatic disease is currently entrusted to the interpretive ability of the pathologist, but it also depends on his level of fatigue.

The automation of the Metrizer standardizes measurement and eliminates operator fatigue, it reads elements which are completely invisible to the human eye and has an unlimited calculation potential.

Furthermore, the Metrizer transforms clinical thought (reasoning) into technological terms and gives both a description and a computer diagnosis, which it obtains using new geometries such as fractals for measuring irregular shapes of scars (collagen islets), fragments of the biliary and the neoangiogenic microvascular network. Being able to measure the real determinants of chronic inflammatory disease dynamics of the liver, means understanding the amount of damage the liver has suffered and refines therapeutic choices, making a more rigorous assessment of its effects.

The totally computerized evaluation entrusted to the Metrizer, i.e. of the liver, completely excludes the subjective intervention of the operator when supplying a "metric hepatogram" where a description is given of the extent of inflammation and fibrosis (scarring), biliary or neoangiogenic regeneration and tissue disorder is quantitatively assessed with a tectonic index. All this is achieved in just a few minutes.

Everyone involved in biologic and medical research will, at some point, come across complex situations which escape thorough description.

All the methods elaborated in the medical world aim at offering adequate procedures with which to locate, define and correlate the largest number of aspects regarding changes in characteristics of the specimen under observation; rather like saying of lesions on organs and organisms determined by natural causes. One way to define tissutal lesions is to study the quality and quantity of the processes causing them and the descriptions of the differences between type, size and intensity of the lesions they cause.

The Metrizer was created to meet the needs of maturity gained in hepatology.

To have measurements which can be repeated anywhere using the same method, means the beginning of a world with less hypotheses and real information.

**List of abbreviations:** VM: Virtual Microscope

**Competing interests:** No conflict of interest exists.

**Authors' contributions:** ND write the manuscript and ideate the theory under the machine and the machine itself, CR revise the text, ideate the software and construct the machine, BF, SDB and SM contribute to revise the text and done the histological preparation, GB revised the text.

**Acknowledgements:** The study was completed with the financial support of the Fondazione Michele Rodriguez, Istituto Clinico Humanitas IRCCS, and Attilio e Livia Pagani.

The entire group is extremely grateful to Ms. Rosalind Roberts for her numerous translations of the manuscript.

**References**

1. Pantanowitz L, Valenstein PN, Evans AJ, Kaplan KJ, Pfeifer JD, Wilbur DC, Collins LC, Colgan TJ: **Review of the current state of whole slide imaging in pathology.** *J Pathol Inform* 2011, **2**:36.
2. Jara-Lazaro AR, Thamboo TP, Teh M, Tan PH: **Digital pathology: exploring its applications in diagnostic surgical pathology practice.** *Pathology* 2010, **42**:512-518.
3. Kayser K, Görtler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, Kayser G: **AI (artificial intelligence) in histopathology–from image analysis to automated diagnosis.** *Folia Histochem Cytobiol* 2009, **47**:355-361, Review.
4. Dioguardi N, Franceschini B, Russo C, Grizzi F: **Computer-aided morphometry of liver inflammation in needle biopsies.** *World J Gastroenterol* 2005, **11**:6995-7000.
5. Grizzi F, Russo C, Franceschini B, Di Rocco M, Torri V, Morenghi E, Fassati LR, Dioguardi N: **Sampling variability of computer-aided fractal-corrected measures of liver fibrosis in needle biopsy specimens.** *World J Gastroenterol* 2006, **12**:7660-7665.
6. Dioguardi N, Grizzi F, Franceschini B, Bossi P, Russo C: **Liver fibrosis and tissue architectural change measurement using fractal-rectified metrics and Hurst's exponent.** *World J Gastroenterol* 2006, **12**:2187-2194.
7. Dioguardi N, Grizzi F, Fiamengo B, Russo C: **Metrically measuring liver biopsy: a chronic hepatitis B and C computer-aided morphologic description.** *World J Gastroenterol* 2008, **14**:7335-7344.
8. Dioguardi N: **Hypothesis for a new method to measure the dynamic patterns of tissue injury.** *Med Hypotheses* 2011, **77**:1022-1027.
9. Bassingthwaighte JB, Liebovitch LS, West BJ: **Fractal Phisiology.** *New York: Oxford University Press* 1994, 11-44.
10. Hastings HM, Sugihara G: **Fractals. A user's guide for the natural sciences.** *Oxford Science Pubblications* 1993, 36-55.

## S18

### Diagnostic reproducibility on a digital evaluation system slide in cytology and histology in oncologic screening

Stefania Lega[1], Paola Crucitti[1], Paola Pierotti[1], Roberta Rapezzi[1], Priscilla Sassoli de' Bianchi[2], Carlo Naldoni[2], Arrigo Bondi[1*]
[1]Anatomia Patologica Ospedale Maggiore, Azienda USL di Bologna, Italy; [2]Direzione Generale Sanità e Politiche Sociali – Regione Emilia-Romagna, Italy

*Diagnostic Pathology* 2013, **8(Suppl 1):**S18

**Background:** Diagnostic reproducibility and accuracy in cytology and histology are the main issues in Oncologic Screening of cervix, breast and colorectal cancer: it can be achieved by programs for quality assurance (QA). The slides set standard represents the most used method to compare diagnostic proficiency, the chance of interpreting microscopic digital photographs provided an interesting alternative to read conventional microscope slides.

The whole digital slide observed in a computer screen is a third, interesting, option to reach the goal. In fact all the information on conventional samples are transferred into a file, easily archived, catalogued, duplicated or advice for quality control, but is especially available at distance and from multiple locations simultaneously with drastic reduction of time needed to achieve proficiency test reproducibility [1].

The production of digital slides with modern scanners is relatively simple and quick. All suppliers offer services into private or public networks server in the literature [2] and software able to track scanned cases stored in comprehensive database to build large casistic archives online [3] Tools are already available for a teleconference discussion of cases with vision of cytological preparations on line [4,5], educational programs with integrated digital slides are poorly developed or proficiency tests for continuing education and professional updating are easily accessible.

A project on Virtual Microscopy and Digital Pathology has been conducted in Emilia-Romagna Region (Italy) with the goal to promote quality in diagnostic cytology and histology in Screening programs by testing a different system involving pathologists and cytologists using digital slides, with a faster and reproducible program easier to manage than standard diagnostic sets and by distance for retraining of patrhologists with a final consensus meeting.

The aim has been reached with the realization of a management system for cytological and histological whole-slides digital images and related clinical data and the building of a picture archive and communication system (PACS) among pathologists of our (and probably other) region. This must be backed by software for the realization of network slide seminars to perform periodic diagnostic reproducibility and proficiency test. The cases, collected and properly catalogued in an online, easily accessible and systemic digital archive of slides, with diagnoses discussed in clinical and pathologic audit meetings and validated by experts, can be used as diagnostic reference tools (case registry online). The cataloguing and indexing is performed with NAP codes derived from SNOMED [6], which contains terms and definitions in Italian and English and encompasses extensive synonyms and complex researches.

**Material and methods:** The cancer screening group of the Emilia-Romagna Region (Italy) set up a picture, archive and communication system (PACS) devoted to pathologists for cooperative diagnosis, teaching and training, teleconsulting, documentation of rare cases and pilot experiences; furthermore selected cases are catalogued in the PACS with the aim to check the diagnostic concordance in regional oncologic screening (cervix, breast and colon). The PACS system is made by two Aperio scanner and an adequate internet server where the described programs perform, (see Figure 1) [7].

The slides have been digitalized using an Aperio scanner, 20x for histology and 40x for cytology and an internet server was used to store the files, arranged into a Spectrum database (Aperio). An e-learning platform (Docebo) [8] has been used to built interface for the applicants: cases and slides were considered "teaching instruments" for the educational software (slide seminars) and appropriate questioning forms have been designed with the diagnostic occurrences of Bethesda System 2011 for cytology and CIN options for histology for the cervical cancer and of International guidelines for breast and colorectal cancer.

Nowadays the diagnostic reproducibility has been performed in colorectal and cervical cancer screening (Bologna October 2010, Bologna June 2011), and for breast cancer is ongoing (Bologna June 2012). In all three slide seminars a number of cases have been selected by a committee of pathologists working in regional anatomo-pathologic units.

Colorectal cancer screening was the first retraining slide seminar for pathologists performed with these features in our region and probably in Italy. Three regional units (Bologna, Cesena and Ferrara) were involved by sending representative histological cases of all main diagnostic occurrences to test the diagnostic reproducibility; 28 histological cases were collected. A definite and limited time interval was indicated to study slides, then a consensus conference was organized in the same day.

The second Seminar of QA to test the diagnostic reproducibility was performed in cervical cancer screening program; 30 cytological and 30 histological cases have been selected by a committee of pathologists among the cases proposed by the regional units. All main diagnostic occurrences were represented, basic clinical information and relevant follow-up information were available; the cases have been completely anonymized.

A 30 days interval was indicated to study the slides, then a consensus conference has been programmed at the end of the evaluation to present the results and discuss cases. Before the meeting each participant received a report with the gold standard diagnosis performed buy the committee and her/his diagnosis and concordance result.

**Results and discussion:** 15 pathologists of regional units attended the **colon-rectal QA** and the diagnostic reproducibility have been evaluated matching their results with the final gold standard diagnosis reached during the consensus conference. The observed agreement was 69% and the overall performance of the participating pathologists was assessed

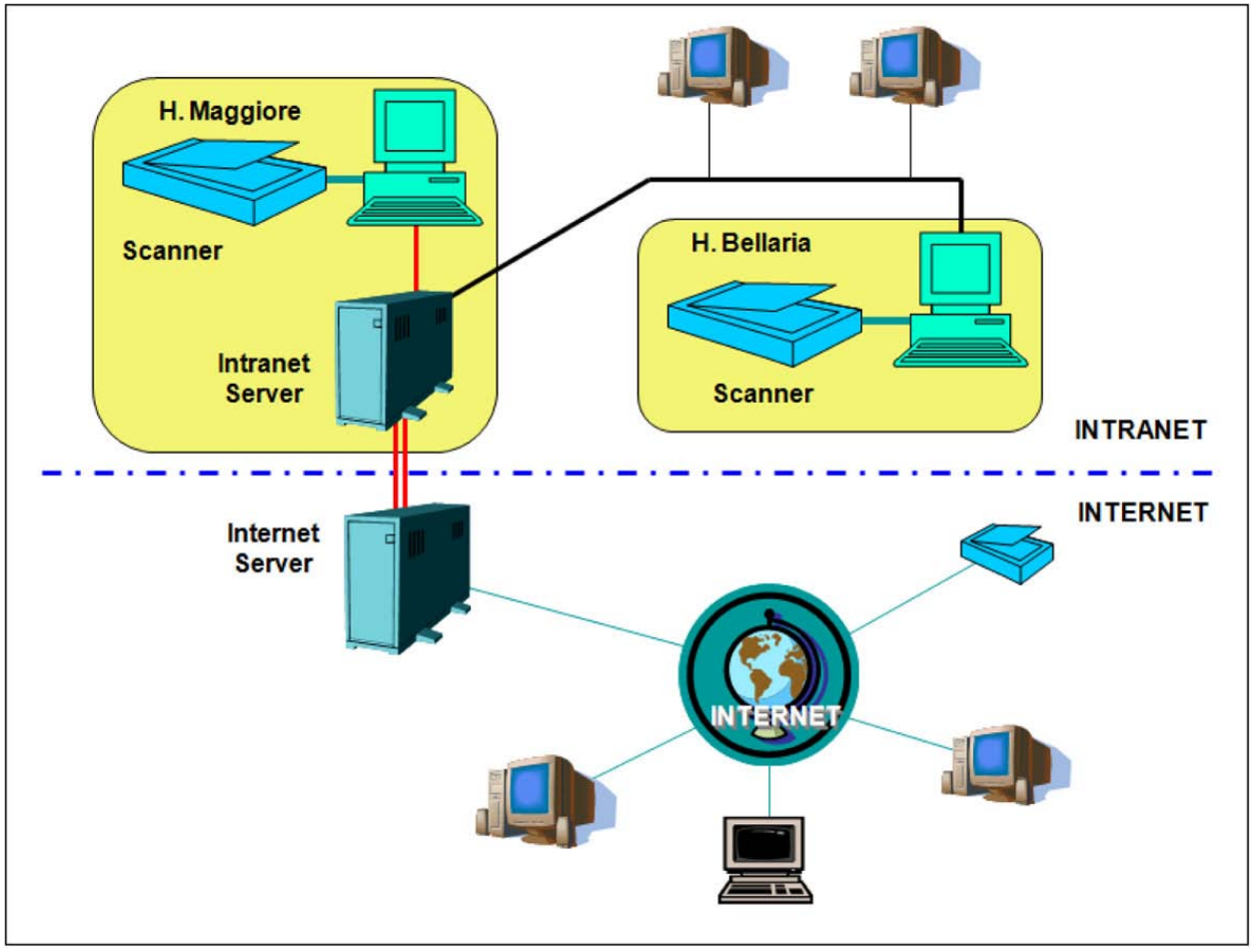



**Figure 1(abstract S18)** Network scheme: two scanner are in an intranet environment with an local disk server, connected to an internet server where the public products are stored.

with a statistical analysis using Cohen's kappa: the average value was 0.64 (substantial).
95 cytologists and 32 histopathologists have been involved in the **cervical cancer screening QA**.
The diagnostic reproducibility has been evaluated using the final diagnosis reached in the consensus conference: in 2 out of 30 cytological cases the diagnosis was different from the diagnosis of the committee, while all histologic diagnoses were in agreement. The overall performance of the participating readers is reported in table 1.
**Conclusion:** Whole digital slide is suitable for proficiency tests and the internet e-learning platform allows to share cases and to get the answers from participants, in a easier way than by of a set of conventional slides. The quality of whole slides is very good, approaching optical microscopic resolution.

**Table 1(abstract S18) Screening PAP test - distribution of agreement in diagnosis**

| | | Agreement in homogeneous groups: final diagnoses | | | |
|---|---|---|---|---|---|
| | | **negative** | **ASCUS L SIL** | **ASC-H H SIL / SqCC** | **AGC AIS / ADC** |
| **Reader's diagnosis** | **negative** | **85.1** | 15 | 6.6 | 2.1 |
| | **ASCUS/ L SIL** | 5.1 | **72.4** | 20.2 | 3.6 |
| | **ASC-H/ H SIL/ SqCC** | 2.9 | 11.9 | **67.7** | 14 |
| | **AGC / AIS/ ADC** | 2.9 | 0.3 | 5.2 | **79.2** |
| | **other tumours** | 3.5 | 0.2 | 0.2 | 0.7 |
| | **no answer** | 0.5 | 0.2 | 0.1 | 0.4 |
| | | **100%** | **100%** | **100%** | **100%** |
| | Overall observed agreement | | 73% | | |
| | Cohen's Kappa | | 0.63 | (substantial) | |

In a cytological environment the bias is to get a perfect focus in all parts of the slide; the wider area of slide to examine and the higher number of diagnostic classes may justify a worse agreement of the pathologists and a poorer performance (lower Cohen's kappa) than histology.
We have produced an integrated environment that includes many of the modern aspects of digital pathology that can be shared with the PACS system in many laboratories of the region, including quality promotion and control of image interpretation in cytology and histology applied to cancer screening.

**References**

1. Demichelis F, Della Mea V, Forti S, Dalla Palma P, Beltrami CA: **Digital storage of glass slides for quality assurance in histopathology and cytopathology.** *J Telemed Telecare* 2002, **8**:138-142.
2. Wilbur DC, Madi K, Colvin RB, Duncan LM, Faquin WC, Ferry JA, *et al*: **Whole-slide imaging digital pathology as a platform for teleconsultation: a pilot study using paired subspecialist correlations.** *Arch Pathol Lab Med* 2009, **133**:1949-1953.
3. Huisman A, Looijen A, van den Brink SM, van Diest PJ: **Creation of a fully digital pathology slide archive by high-volume tissue slide scanning.** *Hum Pathol* 2010, **41**:751-757.
4. Saysell E, Routley C: **Telemedicine in community-based palliative care: evaluation of a videolink teleconference project.** *Int J Palliat Nurs* 2003, **9**:489-495.
5. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, *et al*: **Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future.** *Hum Pathol* 2009, **40**:1057-1069.
6. College of American Pathologists: **SNOMED - Systematized nomenclature of medicine.** Skokie ,Ill., USA, College of American Pathologists, 2 1979.
7. Bondi A, Pierotti P, Crucitti P, Lega S: **The virtual slide in the promotion of cytologic and hystologic quality in oncologic screenings.** *Ann Ist Super Sanita* 2010, **46**:144-150.
8. Docebo: **e-learning open source platform.** [http://www.docebo.org].

## S19

### Hosting and managing large sets of virtual microscopy slides on the internet for E-learning and for reference

Hans-Peter Sinn
Department of Pathology, University of Heidelberg, 69120 Heidelberg, Germany
E-mail: peter.sinn@med.uni-heidelberg.de
*Diagnostic Pathology* 2013, **8(Suppl 1):**S19

**Background:** While the imaging technology for virtual microscopy can be considered mature, the techniques used for the organization and presentation of the digital slides are more complex because of the diversity of user requirements for every project and the technical challenges regarding the structuring and management of the slides. Generally, the presentation of virtual microscopy on the intra- and internet requires the availability of tools that act as a bridge between the digital slide repository and a content management system (CMS) for the web presentation. This involves specific programming of the CMS in order to display the virtual slides in a consistent and specific manner and provide for easy expandability. This is true for E-Learning systems, but also for general digital slide repositories on the web. We have created internet platforms for undergraduate and postgraduate pathology teaching and reference and will present our approach for an effective management and presentation of large numbers of digital slides.
**Materials and methods:** Educational slides were digitized using an Aperio CS scanner and stored in the Aperio svs file format, using JPEG and JPEG2000 compression. The Aperio Spectrum platform was used as the primary image and metadata repository, but because not used to publicly display its content because of its technical limitations and licensing issues. Therefore, all images and metadata information were exported to specialized E-Learning and pathology reference platforms. The implementation of digital pathology slides on these platforms included the following common important aspects:
1. Load balancing: The simultaneous access of up to 70 students in the classroom to few digital slides at one time, or dozens of users to many slides on an open internet platform requires effective load balancing for the image server. This was achieved by setting up a reverse proxy load balancer on Apache that distributes the load to several image server processes which are running on subordinated machines.
2. For E-Learning, the Moodle open source E-Learning platform was chosen, making use of its database module and customized javascripts. With this approach, a generic format for the presentation of the slides was created that allows for adding structured metadata, side-by side display of normal and pathological slides, and multislide images of the same entity in various special stains. Thumbnail images that are provided by the image server software are embedded dynamically.
3. For the postgraduate digital slide pathology repository, a general purpose content management system (CMS) was chosen (Textpattern), that builds upon the open source LAMP (Linux, Apache, MySQL, PHP) architecture. All metadata, and including the links to the virtual slides could be seamlessly transferred from the Spectrum database into Textpattern.
4. The Textpattern CMS was programmed to provide for a hierarchical structure that reflects the pathological disease taxonomy in a hierarchical fashion. It uses a tree structure with unlimited depth, and is easily expandable. Digital slides are assigned to one or more categories in this system.
**Results and discussion:** The user of a pathology reference site expects a systematic layout of the internet site, allowing for easy access of its content based on the medical disease categories, and to be able to browse related entities more easily than in a medical textbook. Because of the many synonyms and related terms of medical terminology, it is not sufficient to rely on a built-in search engine to access the digital slides in question. For an effective approach, a nosological hierarchy of the digital slides must be built that is governed by the pathological diagnosis and organ sites in focus. This requires the building of hierarchical and associative relationships for the diagnostic terminology as well as the use of dictionaries and thesauri. In this fashion, the content management system is able to focus on the pathological slides searched for, and at the same time provide links to related entities in the database. An associative and disease-oriented browsing becomes possible.
With undergraduate and postgraduate pathology teaching, the focus is on the display of cases series that illustrate a pathologic process or the topic of a seminar or a diagnostic entity. Here, the impact is on the linking of the pathology slides with metadata and the display of related slides, such as special stains, differential diagnoses, or normal counterparts. Therefore, a format must be chosen that allows for the consistent building and management of cases series with structured metadata, such as clinical or disease related information, related slides, and other data, such as PDF files. E-Learning systems such as Moodle that provide modular plugins for multimedia data can be used for this purpose. The linking the the digital slides with the content on the E-Learning system, and the handling of different browser requirements can be achieved by custom Javascript programming in the E-Learning system. Both pathology reference sites and E-Learning sites which present digital slides must have sufficient server capacity to be able to deal with multiple concurrent users who are accessing the same or different digital slides simultaneously. A typical example is a class of more than 50 medical students that use virtual microscopy in the classroom. The user experience of quickly navigating the digital slides is dependent on an effective load balancing solution and an image server technology that supports this. Today's advanced servers with multiprocessor and multicore architectures allow for the upscaling of digital pathology proportional to the number of cores in one machine, and thus limit the need of running multiple servers in parallel for the handling of the image data.
**Conclusion:** The building and management of a digital pathology internet site for the purpose of pathology teaching and reference requires special attention to the database design, the handling of metadata, the user interface, and load balancing.
**Competing interests:** The author declares no competing interests.
**Authors' contributions:** H-PS designed the study, contributed the experimental data, and wrote the manuscript.

## S20

### A morphometric tool applied to angiogenesis research based on vessel segmentation

Maria-Milagro Fernández-Carrobles[1†], Irene Tadeo[2†], Rosa Noguera[3], Marcial García-Rojo[4], Oscar Déniz[1], Jesús Salido[1], Gloria Bueno[1*]
[1]VISILAB, Universidad de Castilla-La Mancha, Spain; [2]Fundación Investigación Clínico de Valencia-Instituto de Investigación Sanitaria INCLIVA, Spain;

[3]Laboratorio de Patología Molecular, Dpto. de Patología, Fac. de Medicina y Odontología, Universidad de Valencia, Spain; [4]Dpto. Anatomía Patológica, Hospital General Universitario de Ciudad Real, Spain
E-mail: gloria.bueno@uclm.es
*Diagnostic Pathology* 2013, **8(Suppl 1):**S20

**Background:** It is now widely accepted that tumor growth and metastasis are angiogenesis and lymphangiogenesis-dependent providing novel therapeutic targets in malignant disease [1-3]. A common feature of the tumor vessels studies is that the investigators focus on microvessel density overlooking other parameters that might be significant, such as the size and shape of the microvessels [4]. In many aspects, tumor vessels are different from normal vessels [5,6]. Studies have revealed the importance of the size and shape of blood vessels in, for example laryngeal tumors [7].

Our aim is to develop a morphometric tool able to perform an easy segmentation of blood and lymphatic vessels to study vascularization following the hypothesis that tumor prognosis may not only be influenced by microvascular density but also by the shape and size of the vessels. To this end a segmentation algorithm based on two complementary methodologies have been developed to segment close and open vessels.

**Material and methods:** A dataset of six core images from the Department of Pathology, University of Valencia (DP. UV) was used. They were extracted from TMA scanned slides (Aperio ScanScope XT) at 40x; five images stained with IHC technique against D2-40 and one image from a TMA previously stained with anti-CD34 antibody. Histologic sections comprise two types of vessels: vessels with unquestionable lumen and vessels without evident lumen. Those with defined vascular lumen can be closed or opened, depending on the continuity of the stained endothelial cells in the perimeter.

The images selected to develop the tool contain from low to high number of vessels of both types (from 2 to 600 vessels/image). In the case of closed vessels with and without lumen we can readily calculate their morphometric measurements. The challenge appears when the vascular lumen are not closed or are vaguely stained. In this case a radial algorithm is required.

The algorithm developed by the authors for the segmentation of the vessels, called AngioPath, consists of two parts: a) based on colour segmentation and b) based on the radial distribution of the vessel contour pixels. This algorithm has been compared with other free available algorithm call Caiman [8].

The methods applied for the development of AngioPath is described as follows:

**a) Segmentation based on HSV colour model:** Closed vessels with and/or without obvious lumen can be detected through their continuous brown colour. The algorithm proceeds as follows:

1) Conversion of the RGB TMA image to HSV colour model. This permits the segmentation of the brown colour enabling the extraction of the vessel contours.

2) Extraction of the *S* and *H* channels from HSV image. S channel contains most brown shades and ground staining although it is not sufficient. For this reason H channel is also used.

3) Application of a binary inverted thresholding to *S* channel image and of a binary thresholding to H channel image with thresholds of 30 and 20 respectively.

4) Application of a logical *OR* operator to both images. This operation segments brown colour and erases the rest of the colours.

5) Application of a logical *NOT* operator in order to invert the image.

6) Elimination of small artifacts of the input image and joining nearby structures. Erode and a dilate operations of 2 and 4 iterations respectively are performed in the image.

7) Application of a contour-finding operator. This is used to detect the vessel contour pixels that divide each segment of the image, allowing storage through sequences and individual manipulation. This is applied to images created by a binary thresholding or a Canny operator.

8) Discard false positives such as macrophages, which could be detected as small vessels. Only the contours larger than 6 pixels overall size or with a width and height larger than 20 pixels are selected. These valid contours are the valid vessels.

9) Extraction of morphometric measurements (table 1).

10) Storage of all data measurements in an Excel format file. The final image is saved with the vessel contours in a TIFF format file.

**Table 1(abstract S20) Morphometric Measures**

| Measure Name | Meaning | Units |
|---|---|---|
| **Position** | x-y coordinates | pixels |
| **Area** | vessel contour area | physical units (μm) |
| **Size** | vessel width and height | physical units (μm) |
| **Vascular Density** | $\frac{number\ of\ vessels}{mm^2}$ | |
| **Aspect** | $\frac{\max\ axis}{\min\ axis}(*)$ | |
| **Roundness** | $\frac{perimeter^2}{4\pi\ area}$ | |
| **Perimeter Ratio** | $\frac{convex\ perimeter}{perimeter}$ | |

(*) Axes correspond to the best fitting ellipsoid of the vessel contour.

**b) Radial distribution of the vessel contour pixels:** This algorithm finds the vascular lumen and their brown endothelial surrounding cells. The unconnected parts are then joined together. Once the open vessels are closed, morphometric measurements are calculated. The algorithm proceeds as follows:

1) Extraction of the green channel from the RGB image. It helps to better distinguish the different vascular lumens. The use of a single channel reduces the computational time and the RAM memory used to process images.

2) Application of a binary thresholding to extract vascular lumens from the image. The threshold value was set to 236.

3) Application of a combination of erode (3 iterations) and dilate (2 iterations) operations to join and remove large and small image structures from the previous step.

4) Filling of those closed contours having internal holes smaller than a minimum size (400 pixels).

5) Performing a second erosion of one iteration which increases the space between the vascular lumens and the vessel membrane.

6) Computing the normal direction of each point of the border of a potential vascular lumen using that direction to check if there are some vessel contour pixels near that point. Vessels are considered as valid depending on the ratio of checked points that are actually membrane vessel points. The ratio is also adjusted depending on the size of the vascular lumen contour.

7) Application of a contour-finding operator to find the vessel contours of the radial algorithm.

8), 9) and 10) As with the other algorithm, removal of small artifacts, extraction of measurements for each valid vessel and storage of data measurements as well as the final image with the vessel contours.

The algorithm developed is illustrated in figure 1. The morphometric measurements calculated for each vessel are indicated in table 1. The parameter position enables location of each vessel and thus their elimination in cases of wrong segmentation. These parameters are calculated for the external and internal contour of the vessels.

**Results and discussion:** To our knowledge, only two applications provide vessel closing when the whole perimeter of the vessels is not completely stained which could be a basic feature in translational research. Aperio's application for angiogenesis analysis performs an adequate segmentation of the previously closed vessels but does not provide information about shape [9]. The second algorithm, Caiman, measures shape parameters after closing the vessels but is restricted in the size of the files analyzed (2MB). It shows 96.3% of the contour pixels correctly detected [8]. AngioPath closes vessels, measures shape parameters, supports any image size and shows an average of 95% sensitivity and 98% specificity (ROC analysis). Among the parameters measured, the shape factors roundness and aspect are calculated in both

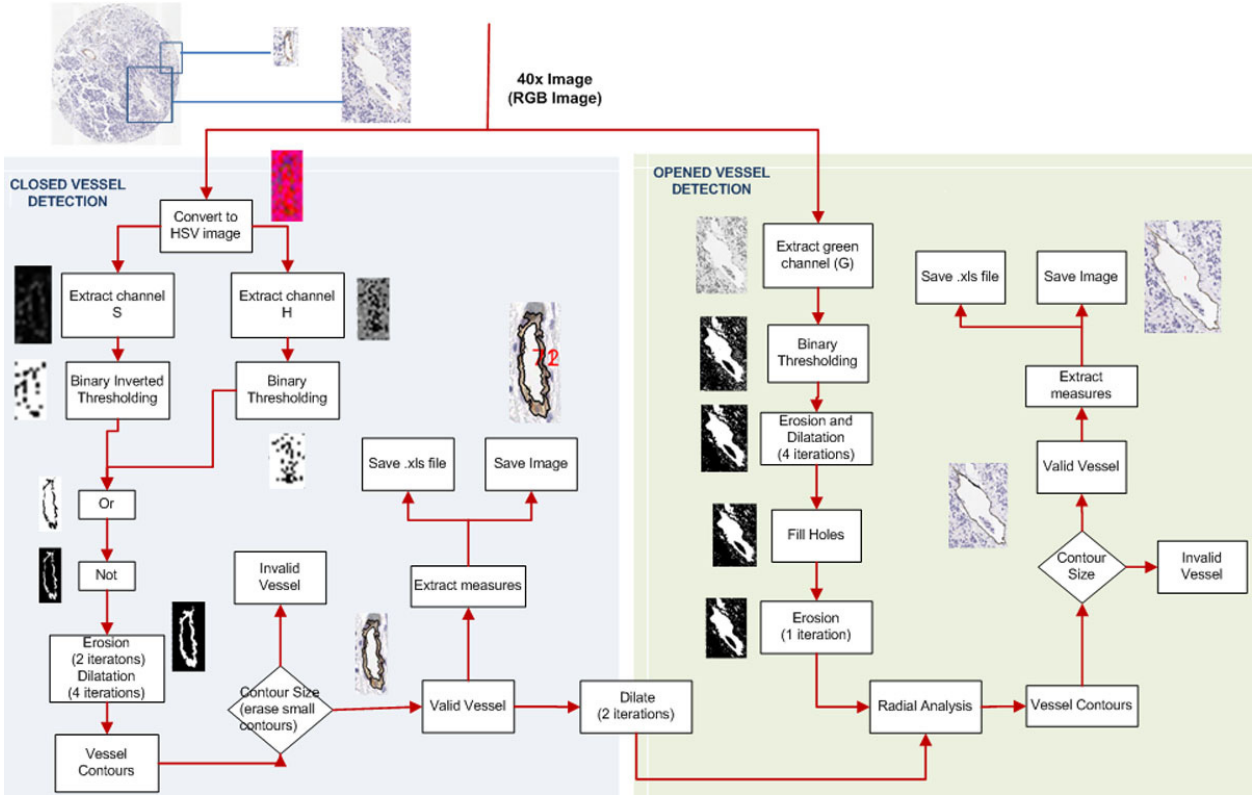



**Figure 1(abstract S20) Vessel segmentation process.** Right panel) segmentation based on HSV colour model and left panel) radial algorithm for detecting open vessels.

Caiman and AngioPath algorithms but AngioPath includes perimeter-ratio which indicates the regularity of the contour of the vessels and which we found to be related to clinical-biological features in, at least, neuroblastic tumors (data not shown). AngioPath allows correlation of a measurement with a given vessel and its elimination it if the automatic analysis makes a wrong measurement, thus solving the problem. Caiman lasts an average of 230s for 1,5Mb or 1200x960 pixel images and could not support the image size of our dataset. The average image size of our dataset is 6300x6300 pixels, that is 120MB. AngioPath, takes between 10 to 180 seconds (s) for images with 2 and 600 vessels respectively.

A limited and preliminary analysis of two images from our database and one example image from Caiman was carried out to assess the quality of the segmentation (see figure 2). The study showed a higher fidelity in the segmentation of the vessels for each tool when using its own images. Nevertheless, AngioPath segmentation correlates with the real vascular structures of the Caiman image better than Caiman does with the images from our database. The differences could probably be related to a specific and differently designed brown colour spectrum provided by the stain or the digital image quality.

**Conclusions:** The developed tool, AngioPath, is based on the essential property of vessel closing to properly count the number and characterize the size and the shape of blood and lymphatic vessels. In the same way, the set of endothelial cells forming a vessel are then considered together as a single object, making vascular density measurement more accurate.

Although this tool has shown good results in the database tested, it may be improved applying invariant colour analysis techniques to properly segment vessels with different stains as well as improve robustness for segmenting poor quality digital images. This tool it is expected to enable wide studies to be carried out to test if shape and size measurements are important for prognosis.

**List of abbreviations:** IHC: Immunohistochemistry; TMA: Tissue microarray

**Competing interests:** The authors declare that they have no competing interests

**Authors' contributions:** MMF, OD, JS and GB have developed the morphometric tool. IT and RN designed the morphometric tool. MG scanned the images. All authors contributed equally in writing the manuscript.

**Acknowledgements:** We would like to thank the Pathology services of the Hospital Clínico de Valencia and of the Hospital General Universitario de Ciudad Real. This work was carried out with the support of the research projects ISCIII (FIS PI10/15), DPI2008-06071, ISCIII (RD 06/0020/0102) of the Spanish Research Ministry.

**References**

1. Folkman J, Merler E, Abernathy C, Williams G: **Isolation of a tumor factor responsible for angiogenesis.** *J Exp Med* 1971, **133**:275-288.
2. Ichihara E, Kiura K, Tanimoto M: **Targeting angiogenesis in cancer therapy.** *Acta Med Okayama* 2011, **65**:353-362.
3. McAllaster JD, Cohen MS: **Role of the lymphatics in cancer metastasis and chemotherapy applications.** *Adv Drug Deliv Rev* 2011, **63**:867-875.
4. Korkolopoulou P, Patsouris E, Kavantzas N, Konstantinidou AE, Christodoulou P, Thomas-Tsagli E, Pananikolaou A, Eftychiadis C, Pavlopoulos PM, Angelidakis D, *et al*: **Prognostic implications of microvessel morphometry in diffuse astrocytic neoplasms.** *Neuropathol Appl Neurobiol* 2002, **28**:57-66.
5. Bergers G, Benjamin LE: **Tumorigenesis and the angiogenic switch.** *Nat Rev Cancer* 2003, **3**:401-410.
6. Carmeliet P, Jain RK: **Angiogenesis in cancer and other diseases.** *Nature* 2000, **407**:249-257.
7. Laitakari J, Nayha V, Stenback F: **Size, shape, structure, and direction of angiogenesis in laryngeal tumour development.** *J Clin Pathol* 2004, **57**:394-401.
8. Reyes-Aldasoro CC, Williams LJ, Akerman S, Kanthou C, Tozer GM: **An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections.** *J Microsc* 2011, **242**:262-278.
9. Aperio: **Angiogenesis applications.** [http://tmalab.jhmi.edu/aperiou/userguides/Microvessel.pdf].

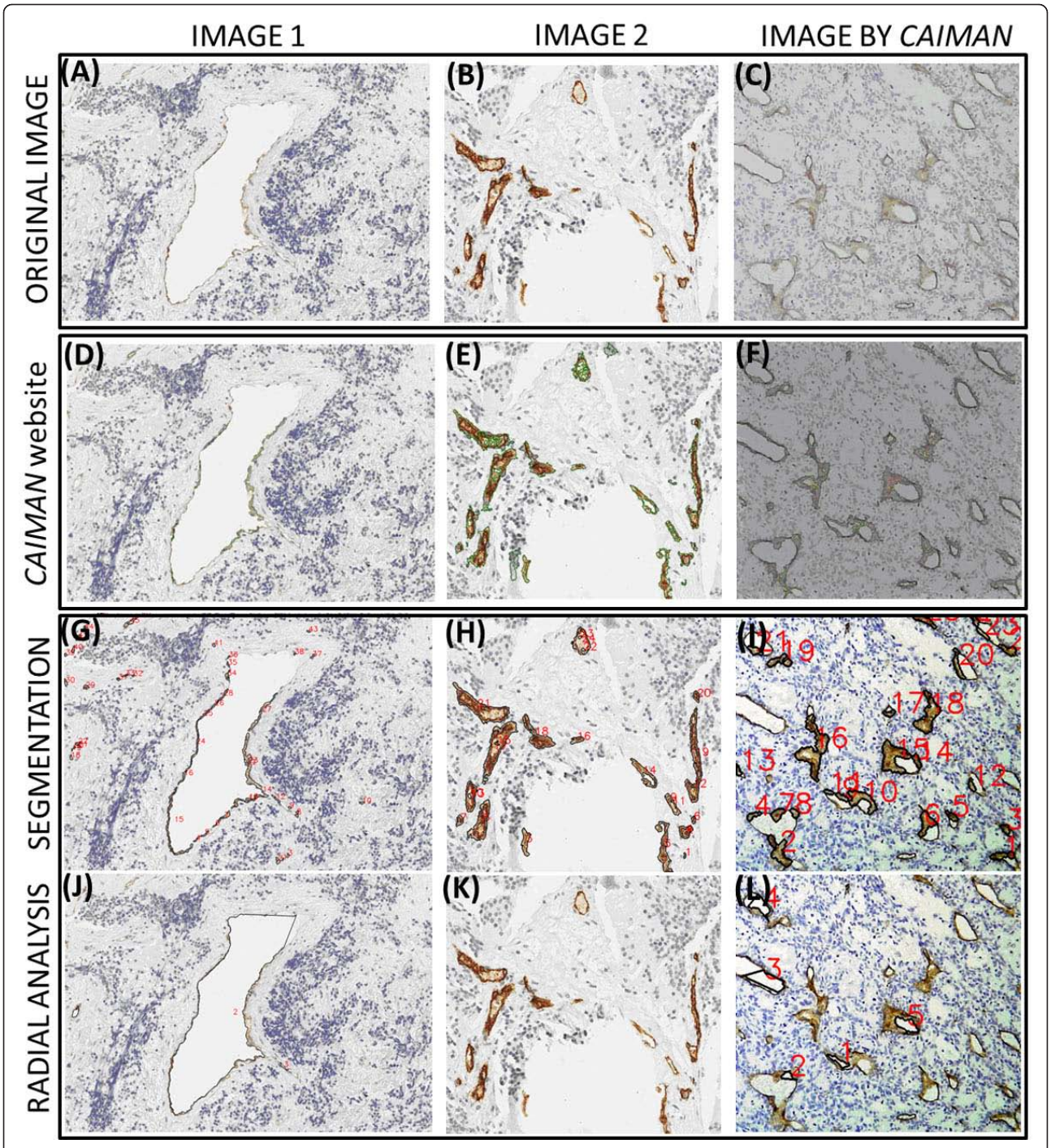



**Figure 2(abstract S20) Comparison of the segmentation quality between the developed tool and CAIMAN application.** 1st row) original images: A and B are images from our database and C is an example image provided by CAIMAN with a less well-lit caption. 2nd row) segmentation performed by CAIMAN (green line) which does not detect small or very large vessels (D) and sometimes divides large vessels into small ones or considers blue areas as vessels (E). The image is well segmented (F). 3rdrow) result of the HSV segmentation (black line).In our images the morphometric tool makes a specific and good detection of the brown staining (G and H). The fine contour lines of the CAIMAN image have not been detected (I). 4throw) a result of the radial analysis. The opened contour vessels have been properly closed, including the large vessel (J). The tool detected the absence of opened vessels in the second image (K). Finally, as it does not detect the fine contours in the CAIMAN image, it does not detect the opened vessels (L).

## S21

### Semi-automatic FISH quantification on digital slides

Gábor Kiszler[1*], László Krecsák[3], Annamária Csizmadia[1], Tamás Micsik[2], Dániel Szabó[1], Viktor Jónás[1], Viktória Prémusz[3], Tibor Krenács[2], Béla Molnár[1]
[1]Department of Image Analysis, 3DHISTECH Ltd., H-1121 Budapest, Konkoly-Thege Miklós u. 29-33, Hungary; [2]I[st] Department of Pathology and Experimental Cancer Research, Semmelweis University, H-1085 Budapest, Üllöi u. 26, Hungary; [3]H-1063 Budapest, Podmaniczky u. 63, Hungary
E-mail: gabor.kiszler@3dhistech.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S21

**Background:** HER2 is a transmembrane glycoprotein, a member of epidermal growth factor receptor family. It was documented that the amplification and the over expression of this gene plays an important role in the pathogenesis and in the progression of breast cancer [1]. Nowadays this is one of the most important biomarker and target for breast cancer therapy. 10-30% of invasive breast carcinomas are HER2 positive, the gene over expression occurring in invasive ductal adenocarcinomas and in invasive lobular carcinomas as well. Therefore the assessment of the HER2 receptor status of the formalin-fixed paraffin embedded cancer specimens has a key importance for specifying the appropriate therapy. For prognostic and predictive testing the immuncytochemical and FISH stains are routinely used in the clinical diagnostic. According to the guideline of the American Society of Clinical Oncology/College of American Pathologists [2] when the immunoquantification is not clear FISH stain should be apply to support the diagnosis. By using dual or multicolor probes, the chromosomal aberration and gene sequence modification can be detected in parallel.

Several methods are available for the evaluation of gene status in FISH stained samples, such as the analysis of the image histogram [3]. Netten et al. [4] used image processing tools to develop their own algorithm, which automatically detects the region of interest, counts the cell nuclei and finds the spots in each nucleus. Furthermore, there is an algorithm which quantifies 2D and 3D FISH images as well [5]. All of these methods work on images which were recorded by microscope. The conventional microscopic diagnosis process allows to make quantification on the recorded sub-regions of the sample. With the expansion of the digital pathology, we intended to develop a FISH quantification platform, which combines the innovative fluorescence whole slide scanning technology with image processing applications.

**Material and methods: Samples and scoring:** Two TMA samples with 20-19 representative HER2 cases were used in this study. Samples were selected from the archive of the 1st Department of Pathology and Experimental Cancer Research of the Semmelweis University, Budapest, Hungary. The survey was performed with the permission of the Institutional Review Board and Regional Ethics Committee of the Semmelweis University (RKEB) (permit no. 7/2006). Sample selection was based on the original IHC HER2 (c-erbB-2/HER-2/neu Ab-12, Thermo Fisher Scientific Inc. USA) scoring in order to use cases that cover all positivity ranges. Thus 24 HER2 positive and 16 HER2 negative cases were included in the TMAs. The formalin-fixed paraffin embedded specimens were stained using HER2 FISH pharmDx™ kit (K5331, Dako, Denmark). Scanned was performed using Pannoramic 250 FLASH (3DHISTECH Ltd., Hungary), that utilizes Plan-Apochromat 40x magnification, 0.95 numerical aperture objectives (Zeiss, Germany) and a PCO.edge 5.5 megapixel, scientific Complementary Metal-Oxide Semiconductor camera (Kelheim, Germany) for fluorescent image acquisition. One part of the image optimalization processes was integrated in the Pannoramic 250 FLASH scanner, which can set up the white balance and can make a shedding correction and a special fluorescent compensation as well. By applying the extended focus scanning method, the different recorded focus lanes (Z=5, step size=0,8 µm) were represented in one summarized virtual lane, which formed the basis of the image processing. The three fluorescent channels were recorded separately and represented the inputs of our FISH detection algorithm.

**Integrated development environment:** We used Visual C++ for developing the image analysis software application. The IPP (Intel Performance Primitives) library was used for image processing.

**Test process:** After the whole slide scanning, all TMA cores were scored by an expert pathologist, using Pannoramic Viewer 1.15.2 ver. software (3DHISTECH Ltd., Hungary). Maximum 40 cell nuclei were quantified, as detailed in the HER2 FISH pharmDx™ kit scoring guideline.

Comparisons were made between the semi-automated assessments, the scorings performed using the software platform and the original ICH scores.

As the data obtained show a non-normal distribution, non-parametric tests (Cohen's kappa and Spearman rank-correlation) had to be applied, using the MedCalc for Windows v. 11.2.1.0 (MedCalc Software, Mariakerke, Belgium) software application. The strength of agreement was interpreted as proposed by Landis & Koch [7].

**Results and discussion: Image analysis:** First we analyzed the blue (DAPI) channel, which represented all cell nuclei. The intensity dynamics and the geometrical morphology of the epithelial cell nuclei were investigated (Figure 1A). After a locally adaptive contrast enhancement the intensity profile of the nuclei followed smooth shape, where the most frequent pixel intensities were from the central chromatin of the cellnuclei.

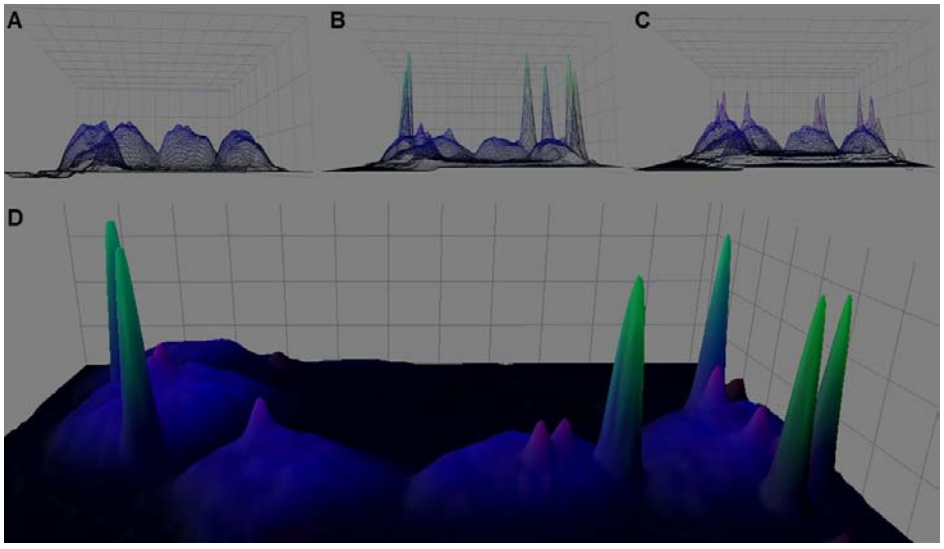



**Figure 1(abstract S21)** The results of the intensity analysis of the digital slide. The DAPI (A), FITC (B) and the rhodamine (C) channels after intensity amplifying. Signals are represented with different intensity in the summarized image (D), therefore an optimal threshold had to be defined in each channel.

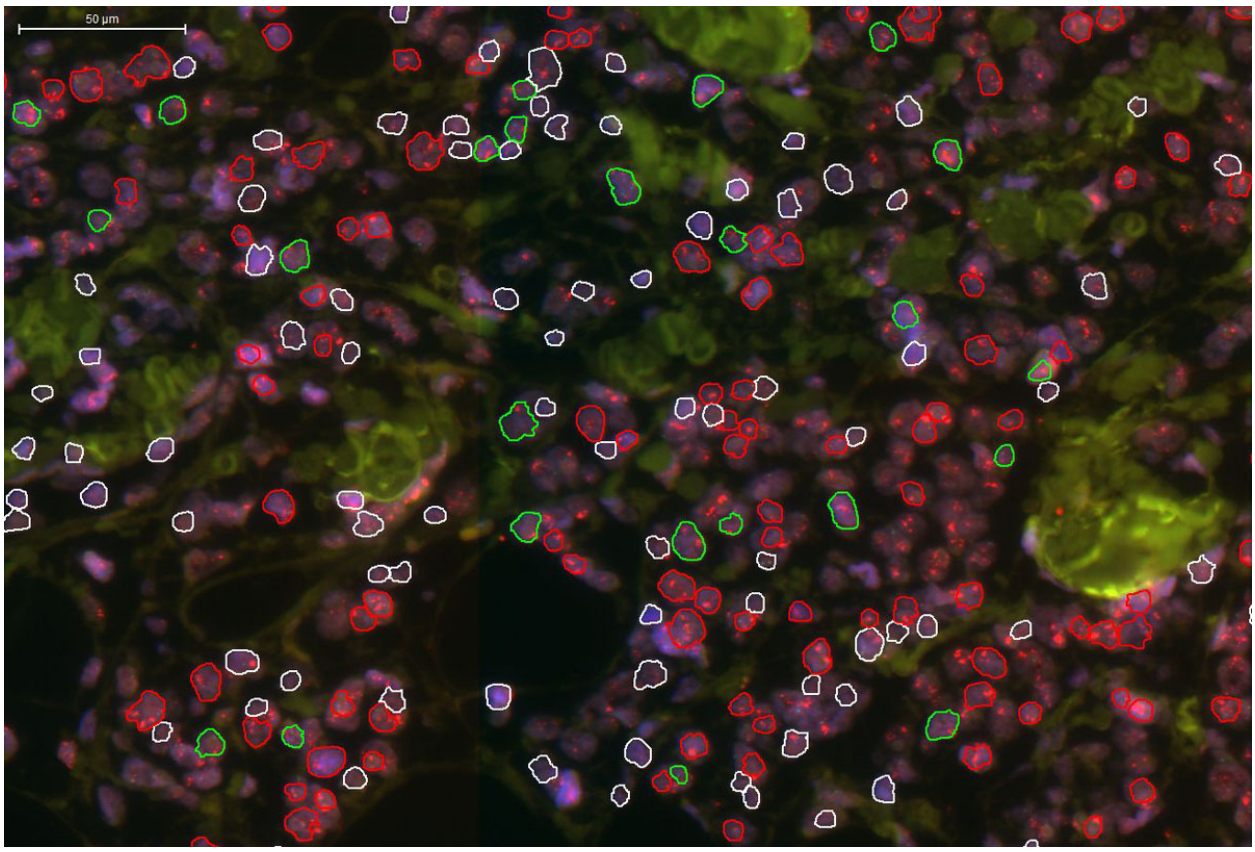



**Figure 2(abstract S21)** The images results after the cell nuclei quantification. Different outlines show the classified nuclei: red=amplification, green=normal, white=non classified, artifact. The cell nuclei were filtered based on their shape, thus only the rounded cell nuclei were classified.

A specific slope graph could be defined at the border of the nuclei in the region of the marginal chromatin and the nuclear membrane. Using these findings, the intensity and the geometrical features of an ideal cell nucleus could be described. This "ideal nucleus" was used hereafter as a nuclei filter in the detection algorithm.

After extended focus recording, all FISH spots were represented in focus plane. Thus the geometrical and intensity parameters of the signals were similar in one channel (Figure 1B, 1C, 1D). This allowed us to apply a general spot detection algorithm for the whole slide. Our algorithm used a local intensity amplifying mechanism, followed by a threshold process, the latter dependent on the quality of the digital slide.

The actual spot numbers were identified for each of the detected cell nucleus and based on this the cell nuclei were classified using the usual Her2 classification rule. Three groups were defined; normal (red signal and green signal were represented equally in the cell nucleus), amplification (the ratio of the red signal and green signal was higher than one per nucleus) and artifact (others).

**Preliminary validation:** A Microscopical Image Segmentation Profile (.misp file) was defined based on the morphological parameters of the nuclei (i.e. radius, area and circularity), the contrast and the intensity of the slide. This profile was used to make investigation on 39 TMA spots. The results of the measurement were saved into an Extensible Markup Language (XML) file and were compared with the results of the Her2 immunostain.

Results of the TMA cores (Figure 2) could be used in the statistical analysis, in diverse extents. The number of nuclei identified and scored using the software platform within each TMA core higly exceeded (mean 279.84 cells, min. 26-max. 939 cells) the number of cells suitable for enumeration of borderline cases (i.e 40 cells) as defines in the FISH pharmDx™ kit scoring guideline. Therefore the ratio of normal and amplified cells has been standardized to 40 cells/TMA core prior to conducting the statistical analysis.

An almost perfect agreement has been found between the semi-automated scoring and the results provided by the software platform (Table 1). Apposite to this result, both the semi-automated investigation and the automated scoring showed only substantial agreements with the initial ICH HER2 score.

**Conclusions:** The basis of our algorithm was the description of the intensity and morphometric features of an "ideal nucleus". This was used to identify cell nuclei with ideal diagnostic potential.

We tested our image analysis application under routine conditions. The selected TMA samples were produced by the histopathology labor and scored manually by an expert pathologist based on the IHC stains and

**Table 1(abstract S21) Agreement between the different scorings**

| Variables | Cohen's $\kappa$ (95% CI) | Spearman's rho (df, 95% CI, p) |
|---|---|---|
| Semi-automated scoring vs Software platform | 0.911 (0.750-1.0) | 0.992 (25, 0.981-0.996, <0.0001) |
| Semi-automated scoring vs Initial ICH score | 0.606 (0.322-0.890) | 0.659 (25, 0.357-0.836, =0.0003) |
| Software platform vs Initial ICH score | 0.703 (0.498-0.909) | 0.762 (39, 0.588-0.869, <0.0001) |

the virtual FISH slide. The results were compared to the measurements of the software. The selection of a high number of cells for FISH scoring did not alter the final results, as showed by the almost perfect agreement found beetween the semi-automated scoring and the one provided using the software application. The use of automated application allow users to select and investigate any sample to a greater extent and with more precison as they would do by using the conventional, manual microscopic, assessment.

The FISH quantification is usually run according to the investigation of fluorescent samples. The quantification process is visually harmful and cumbersome because of the dark room and the bright mercury lump light illumination. The combination of the whole slide digitalization with image segmentation makes the microscopic investigations more accurate and reproducible. Moreover this technic make possible to detect and quantify high numbers of cell nuclei on the whole territory of the recorded samples and analyze fluorescent signals without bleaching.

**List of abbreviations:** FISH: Fluorescens In Situ Hybridization; IHC: Immunohistochemistry; Her2: Human Epidermal growth factor Receptor 2; TMA: Tissue Micro Array; DAPI: 4',6-diamidino-2-phenylindole, a fluorescent cell nuclei stain; FITC: Fluorescein isothiocyanate, a fluorescent stain; XML: Extensible Markup Language

**Competing interests:** Tamás Micsik MD and Tibor Krenács MD are working as a consultant for 3DHISTECH Ltd. for 5 years.

**Author contributions:** GK: made digitization and wrote the majority of the manuscript

TM: pathologist, performed immunohistochemical evaluation,

AC: performed and evaluated FISH-reactions and analyzed the samples with FISHQuant

DS and VJ: optimized the and developed image segmentation

LK: performed the statistical analysis and wrote a part of the manuscript

TK: pathologist, worked as consultant during this work

BM: worked as consultant during this work, managed company affairs

**References**

1. Tan M, Yu D: **Molecular mechanisms of erbB2-mediated breast cancer chemoresistance.** *Adv Exp Med Biol* 2007, **608**:119-29.
2. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF: **American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.** *Arch Pathol Lab Med* 2007, **131(1)**:18-43.
3. Fernandez JL, Goyanes V, Fernandez CL, Buno I, Gosalvez J: **Quantification of C-ERB-B2 gene amplification in breast cancer using fluorescence in situ hybridisation and digital image analysis.** *Cancer Genet Cytogenet* 1996, **86**:18-21.
4. Netten H, Young IT, van VlietLJ, Tanke HJ, Vroljik H, Sloos WCR: **FISH and Chips: automation of fluorescent dot counting in interphase cell nuclei.** *Cytometry* 1997, **28**:1-10.
5. Kozubek M, Koözubek S, Lukasova E, Mareckova A, Bartova E, Skalnikova M, Jergova A: **High-resolution cytometry of FISH dots in interphase nucleus.** *Cytometry* 1999, **36**:279-293.
6. Krecsák L, Micsik T, Kiszler G, Krenács T, Szabó D, Jónás V, Császár G, Czuni L, Gurzó P, Ficsor L, Molnár B: **Technical note on the validation of a semi-automated image analysis software application for estrogen and progesterone receptor detection in breast cancer.** *Diagn Pathol* 2011, **6**:16.
7. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**:159-74.

## S22

### Impact of terminologies for tumor pathology structured reports

G Haroske[1*], T Schrader[2]
[1]Institute of Pathology, Dresden-Friedrichstadt General Hospital, Dresden, Germany; [2]Department Informatics and Media, University of Applied Sciences Brandenburg, Germany

*Diagnostic Pathology* 2013, **8(Suppl 1):**S22

**Background:** For information exchange and data mining structured reports in tumor pathology have to be based on controlled vocabulary as to get a model of meaning. SNOMED CT, originated from a pathology terminology in the U.S. and Great Britain will probably become a global reference health terminology. With SNOMED CT or UMLS very comprehensive medical terminologies exist, which are reference terminologies for tumor pathology applications, too.

However, their direct use for structured reporting is limited so far, because many pathologists are rather bound to data entry systems agreed regionally or nationally and based on pre-digital recommendations and templates, published by scientific boards. As to fill the gap between well-known, often inconsistent and poorly agreed documentation tools and the comprehensive reference terminologies, interface terminologies are being developed [1,2].

National and international initiatives are necessary to reach a growing agreement on particular aspects and needs towards it. Interface terminologies might be a tool for drawing existing separate terminology systems to a finally global standard.

**Methods:** Controlled vocabularies in guidelines of German pathologists for colorectal, breast and prostate cancer [3], in the basic tumor documentation of German cancer registries [4], and in the German HL7 Diagnoses Implementation Guide (http://wiki.hl7.de/index.php/IG:HL7_diagnosis), all standing for German terminologies in the field of tumors, have been mapped to PathLex, an interface terminology developed by IHE [5]. For all German terms without successful PathLex mapping a mapping to LOINC and SNOMED CT was performed. The mapping was done by comparing the German concepts term by term with the PathLex terms and the UMLS concepts using the terminology server http://terminology.vetmed.vt.edu/sct/menu.cfm for SNOMED CT and the current LOINC database http://search.loinc.org/.

**Results and discussion:** On average a pathology guideline describes 40 to 50 terms which have to be registered as to fulfill the minimum documentation requirements. PathLex provides between 30 to 40 terms per tumor entity, less than 80% of them exactly match the German guideline vocabulary (Figure 1). Only a few German terms could be split in components, then matching PathLex terms. The coincidence of PathLex with HL7 Germany vocabulary or the basic data set of cancer registries is still lower, because these are rather classifications than terminologies and are focused on TNM and ICD. Furthermore the German guideline vocabulary, although organ oriented, does not differentiate between generic and organ specific information. This leads to imperfect mapping, too.

The remaining gaps could also not completely be filled with LOINC codes or SNOMED CT terms. There still exist very special terms used in the national scope only.

So far there is no standard on structured reporting in tumor pathology in Germany. The appropriate guidelines having been agreed during the last five years are based on a mixture of national and international terminologies and classifications, as those are used in the daily routine of pathologists, registrars and clinicians. With the upcoming use of HL7 CDA pathologists, too, are interested to use checklists and templates for structured reporting.

The mapping of national terminologies (which are interface terminologies, too) to an internationally agreed interface terminology and to reference terminologies clearly showed that some essentials of terminology building have to be given: clearly and uniquely defined medical concepts, defined terminology structure, relationships and codes. All of them are more or less lacking in one or the other terminology studied. Furthermore, terminologies, nomenclatures and classifications should more clearly divided from each other [6]. Mixing up classifications with terminology is causing some problems with life cycles of those terminologies. Life cycle management of interface terminology should be agreed as soon as possible.

**Conclusions:** Although based on internationally agreed understanding, sharing the same concepts of tumor pathology, the terminology differences among the different sources are quite obvious. Those differences have to be overcome as to ascertain a reliable information exchange between different actors in the care of tumor patients. Increased attention should be paid to the focus of terminology building: basic observable entities for anatomic pathology reports, such as staining characteristics (intensity, pattern, structures stained) evaluation of immunohistochemical reactions, molecular tests, etc. should be included in the scope of activities.

Terminology mapping is one solution, but not the optimal one. Due to the ontological properties of reference terminologies an ontological approach seems to be successful and should therefore be taken into consideration.

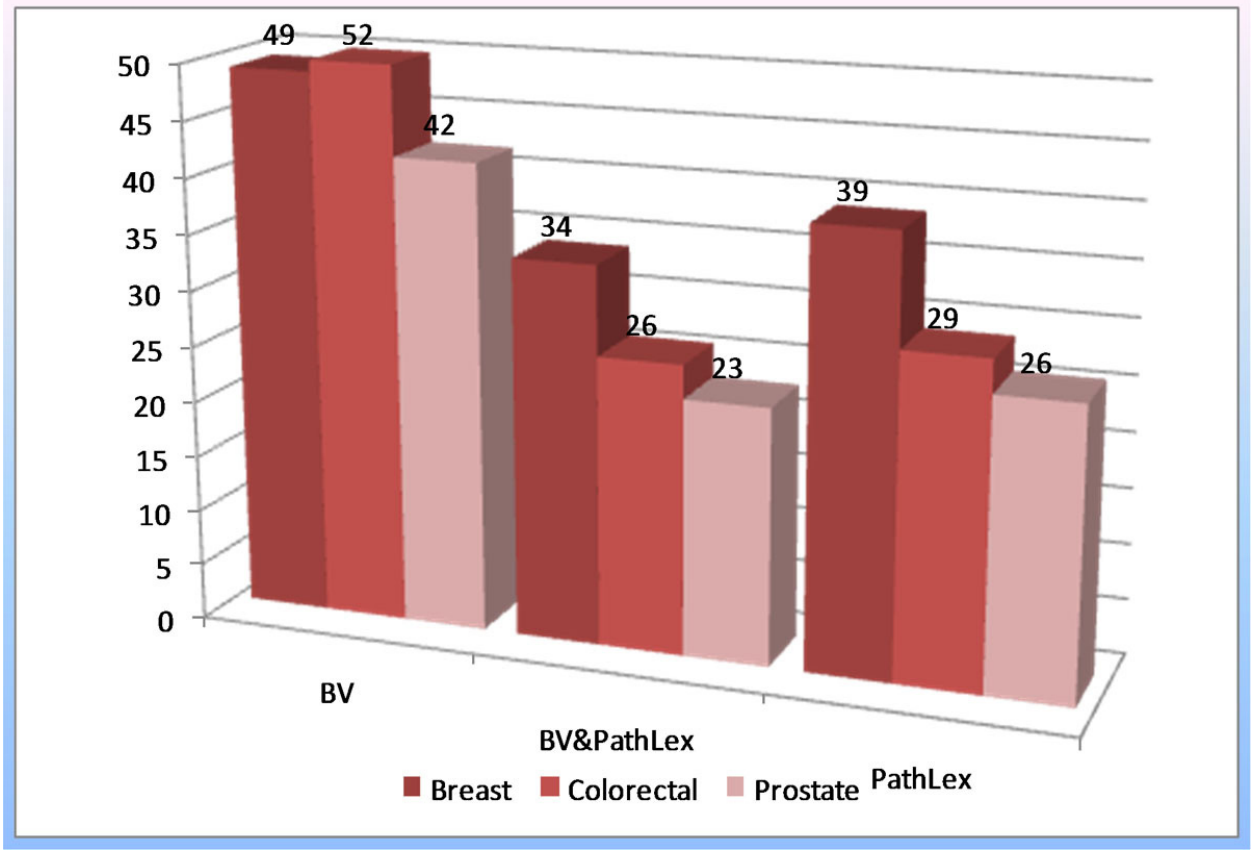



**Figure 1(abstract S22)** Number of terms in German Pathology Guidelines (BV), PathLex and commonly occurring in BV and PathLex for organspecific structured reports.

The toolbox for terminology development has to be improved. In all steps of mapping a link to the concepts underlying the terms studied must be available. A closer collaboration with international terminology bodies as well as a sharpened realization of the impact of terminology in home made guidelines and beginning with ontology construction would contribute to an advanced progress of structured reporting in routine.

**List of abbreviations:** SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms; HL7: Health Level 7; IHE: Integrating the Healthcare Enterprises; LOINC: Logical Observation Identifiers Names and Codes; UMLS: Unified Medical Language System

**Competing interests:** The authors declare that they have no competing interests.

**References**

1. Rosenbloom ST, Brown SH, Froehling D, Bauer BA, Wahner-Roedler DL, Gregg WM, Elkin PL: **Using SNOMED CT to represent two interface terminologies.** *J Am Med Inform Assoc* 2009, **16**:81-88.
2. Daniel C, Macary F, Rojo MG, Klossa J, Laurinavičius A, Beckwith BA, Della Mea V: **Recent advances in standards for Collaborative Digital Anatomic Pathology.** *Diagnostic Pathology* 2011, **6(Suppl 1)**:S17.
3. Wittekind C: **Leitlinien. Vorwort.** *Berufsverband Deutscher Pathologen e.V., Deutsche Gesellschaft für Pathologie e.V., Vorwort* 2006 [http://www.bv-pathologie.de/mitgliederbereich/files/leitlinien_vorwort.pdf].
4. Altmann U, Katz FR, Dudeck J: **A reference model for clinical tumour documentation.** *Studies in Health Technology and Informatics* 2006, **124**:139-144.
5. Daniel C, Macary F: **Anatomic Pathology Structured Reports (APSR) Trial Implementation.** *IHE Anatomic Pathology Technical Framework Supplement* 2011 [http://www.ihe.net/Technical_Framework/upload/IHE_PAT_Suppl_APSR_Rev1-1_TI_2011_03_31.pdf].
6. Bodesinsky P: **SNOMED CT.** *LV:Seminar mit Bakk-Arbeit WS06* Wien, Medizinische Universität 2006 [http://www.meduniwien.ac.at/msi/mias/studarbeiten/2006-BA-Bodesinsky.pdf].

## S23

### Stack or trash? Quality assessment of virtual slides

David Ameisen[1,2,3*], Christophe Deroulers[4], Valérie Perrier[5], Jean-Baptiste Yunès[6], Fatiha Bouhidel[1,2,3], Maxime Battistella[1,2,3], Luc Legrès[1,2,3], Anne Janin[1,2,3], Philippe Bertheau[1,2,3]

[1]Institut Universitaire d'Hématologie, Université Paris-Diderot, F-75010 Paris, France; [2]Service de Pathologie, Hôpital Saint-Louis, APHP, F-75010 Paris, France; [3]Laboratoire de Pathologie, Inserm UMR_S728 / Université Paris-Diderot, F-75010 Paris, France; [4]IMNC - UMR 8165 CNRS / Université Paris-Diderot, Université Paris-Sud, F-91405 Orsay, France; [5]Laboratoire Jean-Kunztmann, Université de Grenoble / CNRS, UMR 5224, 38041 Grenoble Cedex 9, France; [6]Laboratoire LIAFA, Université Paris Diderot, Sorbonne Paris Cité, 75205 Paris Cedex 13, France

E-mail: david.ameisen@univ-paris-diderot.fr

*Diagnostic Pathology* 2013, **8(Suppl 1):**S23

**Background:** Since microscopic slides can now be automatically digitized and integrated in the clinical workflow, quality assessment of these Whole Slide Images (WSI) has become a crucial issue. At this time, the quality of a WSI is verified a posteriori by a technician or by a pathologist. There is however a significant amount of WSI that are too insufficient in quality (blurred, bad colors, poor contrast.) to be used for diagnoses. These slides have then to be scanned again with delay thus slowing down the diagnostic workflow.

To address this problem, we chose to design a method of quality assessment followed by reacquisition, as opposed to a process of enhancement or restoration [1,2]. Such process indeed too frequently results in the degradation of image quality, a key factor in medical diagnosis.

The quality of a flat image can be defined by several quantifiable parameters such as color, brightness, and contrast. One of the most

important parameters, yet difficult to assess, is the focus sharpness (i.e. the level of focus blur) [3]).

Quality assessment of WSI is much more complex than that of flat images because of their intrinsic structure made of multiple magnification levels (pyramidal structure) and resolutions above the gigapixel. One study [4] has shown the possibility of comparing the tiles' contrast and entropy in two WSI obtained with two different scanners digitizing the same slide. Another work [5] assessed the focus sharpness of the tiles of a WSI with the generation of a focus assessment map of the WSI at a given magnification level. However, both these methods still require a human eye to assess if the WSI must be accepted or discarded after the scan.

We describe here a fast method to automatically assess quality and to accept or discard WSI at the time of acquisition.

**Material and methods: Material and software:** For the computations that follow, we used a machine at the University Paris Diderot

Paris 7, with the following configuration: 2 Quad-Core Xeon X5450 3.0GHz/2x6MB, 8GB 667MHz FBD RAM.

The program implementing the new quality assessment method has been developed in Java Web with the NetBeans 6.1 Integrated Development Environment, the Tomcat 6 application server and the database server MySQL 5.

The web survey was developed in PHP5 and MySQL 5.

The tiles of each magnification level of the WSI need to be accessible to perform the analysis. Many open-source programs [6,7] as well as proprietary ones [8] can be used to extract WSI files from different formats (3dHistech, Aperio, Hamamatsu, Olympus) into series of tiles at different magnification levels.

**Methods:** Once the tiles are extracted, the saturation of each of them is computed. In every system, many "blank tiles" are stored because they contain visual artifacts detected as regions of interest but do not contain any specimen. As these blank tiles have saturation values close to zero, our system discards them from the set of images to analyze, saving from 5 to 40 percents of the time required to complete an analysis of a virtual slide at maximum magnification.

The remaining tiles are then analyzed with different tests such as blurriness, contrast, brightness and color. More tests can be integrated as plug-ins in the program. For the blurriness assessment we used a fast reference-free method designed to compute accurately the amount of blur in a single tile based on an edge brightness ratio [9]. Other tests such as contrast, brightness and color assessment are a result of computations made on the tile's pixels vales, compared with their respective thresholds. For instance, one test could be to check if more than 90% of the pixels color values inside a tile were contained in three ranges of color.

Each tile receives quantitative and qualitative scores for each of the analyzed parameters and are compared to their respective thresholds. Note that the tiles can be virtually split to add granularity and refine the final assessment. For instance, at a 2x magnification, if more than 90% of the tiles are considered sharp, the complete 2x layer of the WSI is considered as sharp. If more than 70% of the 10x magnification is considered sharp, the 10x layer of the WSI is considered as sharp.

The analysis can be limited to the lower magnification levels of a WSI for a quicker result or extended to the highest magnification level for a more comprehensive quality assessment.

Once the tile analysis is done, if the WSI passed the quality assessment tests at each processed layer of magnification the WSI is suitable for further use.

In order to test and validate the method, we analyzed a series of 100 WSI made of a mix of WSI with optimal focus and of WSI with various blurred areas, some of them being obviously totally blurred. We compared the computer assessment of these WSI to the human assessment in two settings:

- We first presented the 100 WSI in a random order to two observers from our research team.
- We then conducted a web survey [10] among 22 trained pathologists, asking them whether the overall quality of each WSI seemed sufficient for a clinical use. The human assessment was distributed among three possible answers: Poor; Fair; Good. The computer assessment represented the computed highest acceptable magnification for a WSI, higher magnifications being therefore considered by the computer as of insufficient quality for diagnosis.

**Results and discussion:** In the following, we use the blur assessment method described in the method section as an example to describe any other quantifiable criterion in an image, to be used a fortiori to assess the quality of WSI.

The complete quality assessment method is a logical intersection of independent tests, marking a WSI as of insufficient quality if at least one of the tests fails.

We applied the quality analysis routine with the blur assessment parameter on hundreds of WSI. An example of automatic blur assessment is shown in Figure 1.

On a collection of 100 WSI, two observers could easily assess the overall level of quality they observed and they visually verified that the thresholds we set were highly predictive of the global sharpness or blurriness of the WSI.

For the web survey, the results [10] obtained after the visual analysis on 100 WSI by 22 pathologists are shown in Figure 2. The results found by our algorithms are fully consistent with the pathologists' answers to the survey: the mean computer assessment is 1.25X with a standard deviation of 2.37X in the "poor" human assessment category, increasing to 2.90X with a standard deviation of 2.51X in the "fair" category and to 6.35X with a standard deviation of 5.57X in the "good" category.

However, the survey showed that the human assessment do not entirely correspond to the computer assessment, due to the fact that some diagnoses do not need high magnification for human eyes to be done. Indeed, a high computer quality at low magnification was sometimes enough to give a correct diagnosis (blue disks on the lower right part of Figure 2), but a high-level computer assessment (computed high quality at high magnification) always corresponded to a high level human assessment (blue disks on the upper right part of Figure 2).

As further improvements of our method, we will contextualize the assessment by refining the thresholds depending on staining and lesion.

In terms of computing speed, Zerbe *et al*. [5] showed a distributed computing model to assess the focus sharpness of a WSI, generating a focus assessment map of the WSI at a given magnification level in around 6 minutes per gigapixel per computer. We analyzed on our machine (see Material and software sub-section) 8 complete 1.73 gigapixel digital slides in 400 seconds as eight distinct threads, equivalent to 34 Megapixels per second or 2 gigapixels per minute, per computer. Already 12 times faster than the previous method, we are currently optimizing the program into a multi-thread, multi-node parallel processing system using C++ with OpenMP and OpenMPI libraries to scale it up to match demanding industry requirements. A plug-in support and an API are also being integrated in this optimization to facilitate further integration.

**Conclusions:** As quality assurance is crucial in a context of daily use in diagnostic pathology, we have developed a fast and reliable reference-free tool for quality assessment of WSI.

Our method can be used upstream, as a calibration and quality control tool for the WSI acquisition systems, or as a tool to reacquire tiles while the WSI is being scanned. It can also be used downstream to reacquire the complete slides that are below the quality threshold for surgical pathology analysis.

We are currently optimizing the program to improve its speed and refining its threshold, according to the magnification levels, the staining of the slides, and the type of acquisition devices used.

Such quality assessment scores could be integrated as metadata in WSI shared in clinical, research or teaching contexts, for a more efficient medical informatics workflow.

**Competing interests: Financial competing interests:** David Ameisen was a recipient of a doctoral fellowship grant from Aurora Interactive (2008-2011); Olympus provided him travel reimbursements for scientific meeting presentations. None of these organizations are financing this manuscript. No other author has financial competing interests.

David Ameisen and Philippe Bertheau are currently applying for two patents relating to the content of this manuscript. They are receiving salaries from Université Paris Diderot that has applied for these patents.

**Non-financial competing interests:** None.

**Authors' contributions:** DA participated in the design of the study and drafted the manuscript, CD carried out the Hamamatsu tiles extraction and participated in the design of the study, VP participated in the design of the study, JBY participated in the design of the study, FB participated in the statistical analysis, MB participated in the statistical analysis, LL participated in the statistical analysis, AJ participated in the design of the study, PB participated in the design of the study, and drafted the manuscript. All authors read and approved the final manuscript.

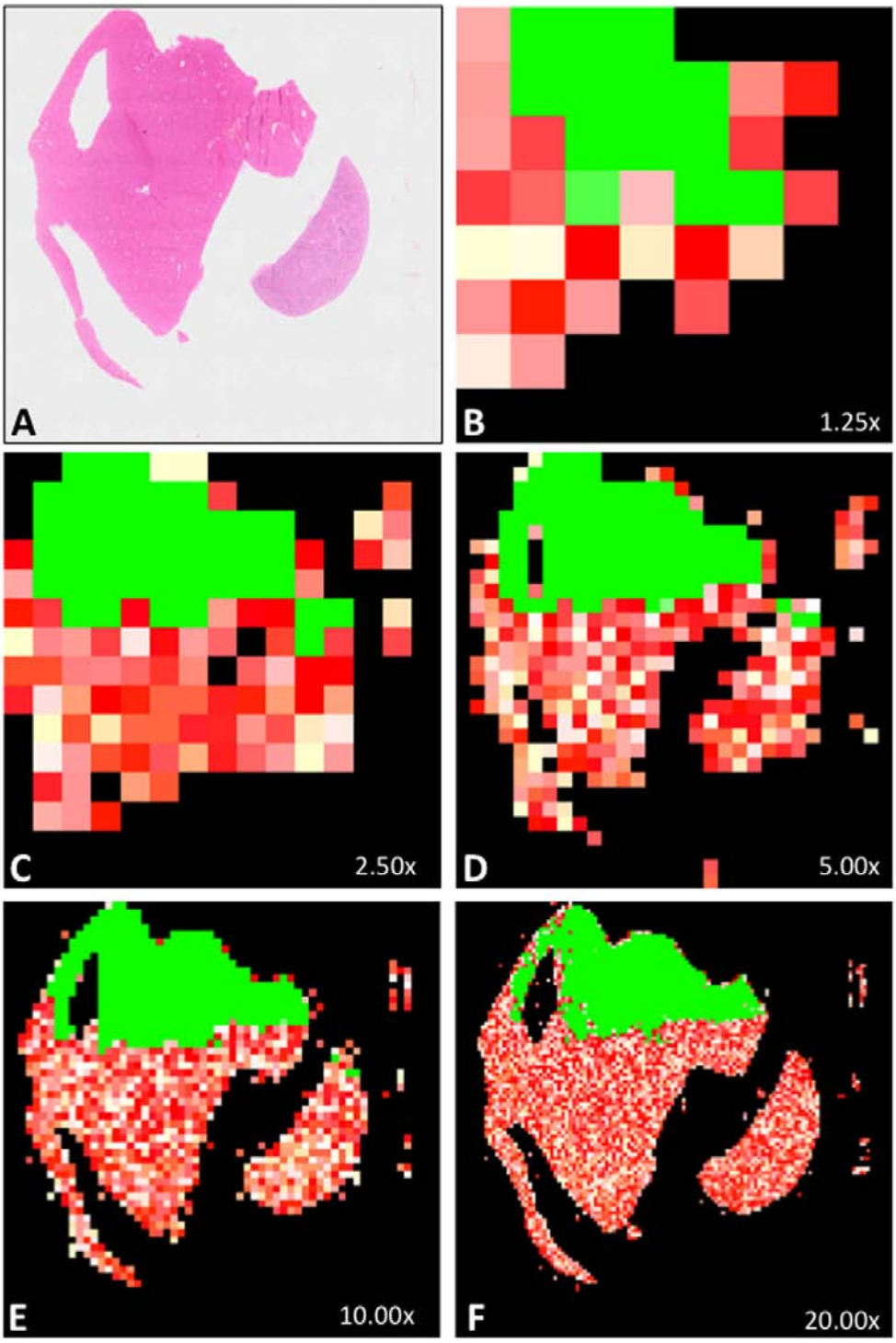



**Figure 1(abstract S23) Automatic quality analysis of a virtual slide (parameter used: blur)** *A* represents the thumbnail of a whole slide image (H&E staining) whose upper third part is in focus and lower two thirds part is totally out of focus. Each thumbnail *B* to *F* shows sharp tiles in green and blurry tiles going from white (a little blurry) to red (the most blurry). Out of 43 tiles at 1.25x (*B*), 83% were detected as non-blank, and 36% were detected as sharp. For *C*, *D*, *E* and *F*, the respective values were (146 tiles, 2.5x, 86% non-blank, 34% sharp), (493 tiles, 5.0x, 83% non-blank, 33% sharp), (1751 tiles, 10.0x, 77% non-blank, 31% sharp), (6589 tiles, 20.0x, 76% non-blank, 25% sharp). The WSI is thus considered as of insufficient quality in terms of blurriness, for all its magnification levels being under their respective blur assessment thresholds.

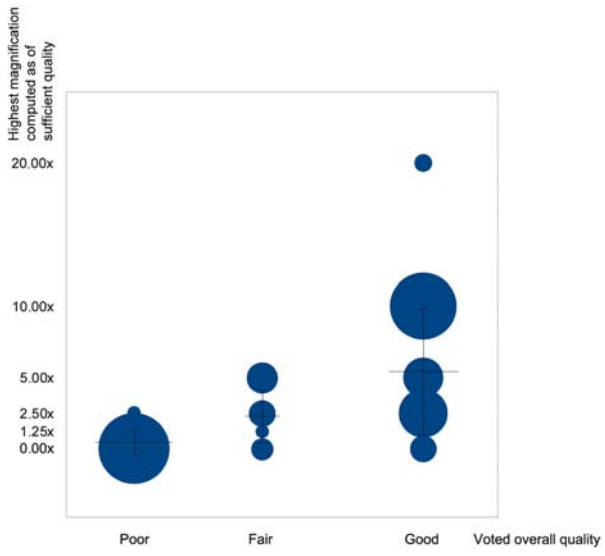



**Figure 2(abstract S23) Comparison between voted overall quality and best detected magnification** Distribution of the (human assessment; computer assessment) pairs for 100 WSI with various blurred levels. Human assessment is distributed in three categories: poor/fair/good quality for diagnosis. Computer assessment is distributed in five different magnifications (from 1.25X to 20X): it shows the highest acceptable magnification for a WSI, i.e. the magnification for which the WSI computed quality is sufficient, implying that higher magnifications of this WSI are of insufficient quality. The surface of the disk is proportional to the number of identical pairs. The horizontal bars represent the mean of the highest acceptable magnifications of the computer assessment at each category of human assessment, with vertical bars as their respective standard deviation.

**Acknowledgements:** The authors are grateful for financial support from the French Agence Nationale de la Recherche (ANR). We also like to thank our colleagues in the Pathology Department of the Saint-Louis Hospital and the Hematology University Institute who made all these experiments easier, Tzuchien Tho for his careful reading of the manuscript and helpful suggestions and Julie Auger-Kantor for her immense help and support throughout this work.

**References**

1. Kim SK, Paik JK: **Out-of-focus blur estimation and restoration for digital auto-focusing system.** *Electronics Letters* 1998, **34(12)**:1217-1219.
2. Lam EY, Goodman JW: **Iterative statistical approach to blind image deconvolution.** *J. Opt. Soc. Am. A* 2000, **17(7)**:1177-84.
3. Ferzli R, Karam LJ: **A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB).** *Image Processing, IEEE Transactions on* 2009, **18(4)**:717-728.
4. Walkowski S, Szymas J: **Quality evaluation of virtual slides using methods based on comparing common image areas.** *Diagn Pathol* 2011, **6(Suppl 1)**:S14.
5. Zerbe N, Hufnagl P, Schlüns K: **Distributed computing in image analysis using open source frameworks and application to image sharpness assessment of histological whole slide images.** *Diagn Pathol* 2011, **6(Suppl 1)**:S16.
6. Triola MM, Holloway WJ: **Enhanced virtual microscopy for collaborative education.** *BMC Medical Education* 2011, **11(1)**:4.
7. Deroulers C: **NDPITools. [Internet].** 2011, [cited 2012 Apr 30]. Available from: http://www.imnc.in2p3.fr/pagesperso/deroulers/software/ndpitools/.
8. Aurora Interactive: **mScope medical communication solution. [Internet].** 2002, [cited 2012 Apr 30]. Available from: http://www.aurorainteractive.com/.
9. Ameisen D, Bertheau P: **Method of blur detection and assessment in a digital image.** *PCT/FR2011/052951* [http://patentscope.wipo.int/search/en/WO2012080643], priority date 2010 Dec 13.
10. Ameisen D, Bertheau P: **Results of the quality assessment in virtual slides survey [Internet].** 2011, [cited 2012 Apr 30]. Available from: http://virtual-slides.univ-paris7.fr/flou/results.php.

## S24

### An entropy-based automated approach to prostate biopsy ROI segmentation

Gloria Bueno[1*], Maria-Milagro Fernández-Carrobles[1], Oscar Déniz[1], Jesús Salido[1], Noelia Vállez[1], Marcial García-Rojo[2]
[1]VISILAB, E.T.S.Ingenieros Industriales, Universidad de Castilla-La Mancha, Spain; [2]Dpto. Anatomía Patológica, Hospital General Universitario de Ciudad Real, Spain
E-mail: gloria.bueno@uclm.es
*Diagnostic Pathology* 2013, **8(Suppl 1):**S24

**Background:** Despite significant improvements in computer vision and image processing techniques, there are few software tools that are able to analyze prostate biopsy images in a fully automated way in order to find ROIs in those images. In order to develop a useful system, user interaction should be minimized, and the system should also be capable of dealing with images acquired at least at 10x magnification, since images of lower resolution do not provide enough information for cancer diagnosis.

The segmentation of the ROI mentioned above is a complex task that includes several challenges. The suppression of user interaction means that the system should be robust enough to deal with image irregularities by itself. Such irregularities may include stain intensity variations, tissue cuts, and even dust over the slide when it is digitized. The physical size and memory requirements of the images also limit the processing algorithms that may be used, since we want our system to be used in personal computers (i.e. not clusters).

In recent years, there have been several studies that focus on Hematoxilin and Eosin (HE) prostate biopsy image processing and histological image analysis. Most of them are focused on the segmentation of only one ROI, usually the nucleus and glands, as well as the extraction of descriptors for classification purposes. A thorough review of the research related to classification may be found in Bueno *et al.* [1]. Statistical information techniques, region growing algorithms [2], fuzzy c-means [3] active contour models, including level set methods [4,5], filtering and morphological analysis [2,5,6] have been also used for ROI detection. The main problem with these methods is that they are not designed to process large amounts of data, which is the case when working with whole digital slides in pathology. Besides, many of these methods yield limited results because they focus mainly on a single structure or type of tissue.

None of the previous techniques use complete mosaics or WSI but rather fragments or magnifications lower than 5x, with the exception of Doyle *et al.* [7] and Vidal *et al.* [5]. Doyle *et al.* used 40x images ranging between 1-2 GB and Vidal *et al.* worked with images acquired at low magnification (5x, 10x) and up to 1.19 GB in size. One of the problems with level set methods is that they are not suitable for parallel processing. Moreover, level set methods have not been used in a general way, only applied to one type of histological images HE biopsies). In a recent work by the authors [1], a general solution is described for parallelizing in an efficient way a set of heterogeneous low and high-level image processing algorithms to be applied to high resolution histopathological WSI. The imaging tools implemented in [1] are general for all types of histological images with different stain, acquired from different anatomical parts and digitized at different magnifications. These tools deal with contrast analysis, ROI detection and classification applied to high resolution images that range from 300 MB to 30 GB.

The aim of this system that will be described herein is the segmentation of ROIs from these images in a way that mimics the method used by doctors, that is, identifying at low magnification the regions with high concentration of cells or where the architectural distribution between lumen and cells is relevant. In this way this work differs from those previously mentioned. The method may be applied to different histological images and are suitable for parallel processing, which differs also from the method presented by Vidal *et al.* [5].

**Material and methods:** A dataset of 200 biopsies stained with HE and provided by the Department of Pathology, Hospital General Universitario de Ciudad Real (HGUCR) was used. The images were digitized using an ALIAS II microscope, from LifeSpan Biosciences Inc. This system acquires tiles with a size of 2000 x 2000 pixels and 24 bits per pixel (RGB). Each tile requires 11.4 MB. The ALIAS microscope is equipped with five

different objectives, whose magnifications are 2.5x, 5x, 10x, 20x, and 40x. We have focused on samples digitized with 10x magnification, although our work could be easily adapted to be magnification-independent. The images at 10x magnification have memory requirements ranging from 8.83 MB (1899 x 1626 pixels) to 220 MB (9755 x 7884 pixels).

The pathologists at HGUCR have also specified the most relevant features that should be considered when analyzing prostate biopsies at these magnifications. HE stained prostate biopsies have three types of well-differentiated structures of interest: lumen, cytoplasm, and cells. For pathological purposes, the most important structures are cells [1]. Their morphology, distribution between them, and relationship with lumen and cytoplasm are the most relevant features that pathologists consider to elaborate a diagnosis. Groups of cells are especially important, and they are what we consider a true ROI. These groups may appear either surrounding a lumen area, or packed very closely. Typically, in both cases some cytoplasm will appear between the cells. Thus, ROI are complex areas, where the three types of structures of interest appear in an unpredictable fashion. Although lumen and cell areas can be individually segmented without much effort, the segmentation of ROIs where the structures are grouped requires advanced techniques. It is desirable that all the three types of structures share a common feature (or a manageable set of them), so they all can be separated from the rest of the image using that feature.

If an RGB image of a prostate biopsy is converted to the YIQ colour model, and then the I channel is extracted from the image and equalized, the result is an image where the regions of interest are clearly highlighted respect to the rest of the image. It is important to apply the equalization only to the region where the tissue lies, since its results vary if it is applied to a region where there is no tissue present.

Once the I channel is properly equalized, regions of interest clearly appear darker than other regions, so a binarization could be used to separate ROI from the rest of the tissue. This binarization sets as foreground all the dark regions in the image (without the non-tissue region), and sets everything else as background. However, this basic technique alone does not produce good results on most images. Since images tend to vary greatly in the edges of the three types of the structures of interest, especially in the outer border of the cells, entropy turns out to be a great feature to determine where the ROIs are located. It has been observed that entropy calculations produce better results when applied to the green channel of the RGB image, because it is the one that features higher contrast between the structures of interest.

Given a pixel in an image, $P_i = (x_i,y_i)$, a circular neighborhood of radius $R$ around it is composed by all the pixels $P_j = (x_j,y_j)$ with an Euclidean distance to $P_i$ is lower than $R$. For each pixel $P_i$ in the green channel, a circular neighborhood around it is defined. Then, the histogram of the neighborhood is calculated as well as its entropy. Suppose that $f_i$ is the relative frequency of pixels with intensity value $i$, then the entropy at $P_i$ is calculated as: $\sum_{j=0}^{255}[-f_i \cdot \log_2(f_i)]$.

The results of entropy calculation depend on the radius of the neighborhood considered. Obviously, the computational footprint of entropy calculation increases with the size of the neighborhood. The chosen radius for entropy calculation is 27. The result of entropy calculation is also equalized. It should be remarked that entropy calculation, as well as equalization, is only applied to the tissue in the image, and therefore the background does not affect these calculations.

The entropy image may be binarized in order to separate regions of high entropy from the rest of the image. Since big lumen areas present low entropy, connected component analysis is performed to fill in the holes that are smaller than a predetermined size. Then, it is combined with the binarized I channel using the logic operator AND. Finally, morphologic operators are used to remove noise and smooth the results. These morphologic operations involve a first dilation with a circular kernel with radius 3, followed by an erosion (circular kernel, radius 5) and a final dilation (circular kernel, radius 2). Figure 1 illustrates the full segmentation process.

**Results and discussion:** Some selected fragments that exemplify the algorithm are shown in Figure 2. Although the images used to test the algorithm were large, computational times were not deemed excessive. The algorithm takes from 9 seconds to 5 minutes for 9MB (1899x1626 pixels) and 220MB (9755x7884 pixels) images respectively. The test machine was equipped with an Intel Pentium 4 640 (3.2 GHz) and 2GB RAM (DDR2-533MHz).

A quantitative validation based on ROC analysis was carried out with our set of 100 different tissue samples of WSI. The samples were both benign and malign samples of prostate biopsy. The results of the algorithm were compared to the manual selection of ROIs done by pathologists from the HGUCR. Thus, the rates of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) detections were calculated. On average, 15% of detections were FN, 2% were FP, 85% were TP, and 98% were TN. The type I error, that is, the FP for the 100 images, occurs in those samples stained with weak HE dye. The final results show an average sensitivity of 83% with specificity above 99%.

**Conclusions:** In this paper, an approach to ROI segmentation in whole slide images of prostate biopsies has been described. The method proposed is based on texture and colour based analysis.

The novelty of the technique lies in the ability to detect complete ROIs, where a ROI is composed by the conjunction of three different structures, that is, lumen, cytoplasm, and cells with a high density of cells and the architectural distribution between lumen and cells. The method is capable of dealing with full biopsies digitized at different magnification. The proposed algorithm is also original because it works on large images acquired with low magnification, thus being different from other algorithms that require higher magnification and have been tested only on small samples. In this way, the method tries to mimic the manual procedure of expert clinicians. Moreover, the method is suitable for parallelization and may be applied to different tissue samples.

The proposed system is also useful because it can be used for different purposes. It could be integrated into a slide visualization environment to highlight the ROIs for the pathologists, either for slide analysis or even with teaching purposes. The system could also be used as a previous step in classification applications, since it could reduce the amount of information to be processed, and probably speed up the whole classification process.

Segmentation accuracy is high for HE stained samples. Furthermore, the algorithm is also reasonably fast. We are currently working to improve the robustness and speed of the algorithm, making it less sensitive to disturbing factors such as different illumination conditions, tissue thickness, and stain amount, as well as parallelizing some parts of the pipeline in order to speed it up. We are also working in developing a set of features of interest that should be segmented and analyzed in order to provide further information to the doctors.

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** All authors from VISILAB have developed and tested the algorithm. MG-R from HGUCR has provided the tissue samples and qualitatively validated the results. All authors contributed equally in writing the manuscript.

**Acknowledgements:** We would like to thank the Pathology services of the Hospital General Universitario de Ciudad Real. This work was carried out with the support of the research project ISCIII DPI2008-06071 of the Spanish Research Ministry.

**References**

1. Bueno G, González R, Déniz o, García-Rojo M, González-García J, Fernández-Carrobles MM, Vállez N, Salido J: **A parallel solution for high resolution histological image análisis.** *Computer Methods and Programs in Biomedicine* 2012.
2. Belkacem-Boussaid K, Samsi S, Lozanski G, Gurcan MN: **Automatic detection of follicular regions in H&E images using iterative shape index.** *Computerized Medical Imaging and Graphics* 2011, **35(7-8)**:592-602.
3. Peng Y, Jiang Y, Eisengart L, Healy M, Straus FH, Yang X: **Computer-aided identification of prostatic adenocarcinoma: segmentation of glandular structures.** *Journal of Pathology Informatics* 2011, **2(33)**:1-10, doi: 10.4103/2153-3539.83193.
4. Xu J, Madabhushi A, Janowczyk A, Chandran S: **A weighted mean shift, normalized cuts initialized color gradient based geodesic active contour model: Applications to histopathology image segmentation.** *Proceedings of SPIE* 2010, **7623.**
5. Vidal J, Bueno G, Galeotti J, García-Rojo M, Relea F, Déniz O: **A fully automated approach to prostate biopsy segmentation based on level-set and mean filtering.** *Journal of Pathology Informatics* 2011, **2(5)**:1-11, doi: 10.4103/2153-3539.92032.

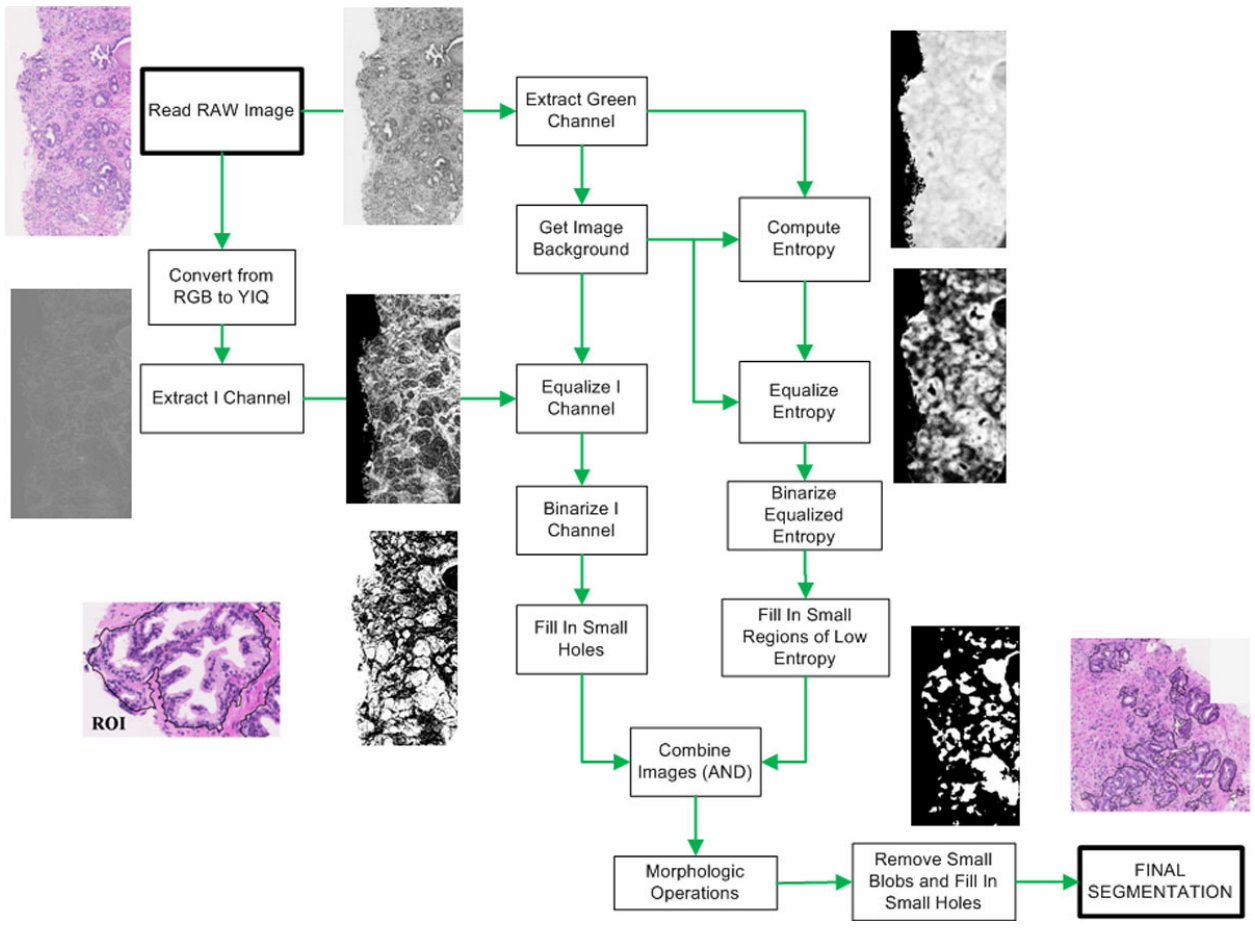



**Figure 1(abstract S24)** Flowchart of the segmentation process.

6. Sertel O, Lozanski G, Shana'ah A, Gurcan MN: **Computer-aided detection of centroblasts for follicular lymphoma grading using adaptive likelihood-based cell segmentation.** *IEEE Transactions on Biomedical Engineering* 2010, **57(10)**:2613-2616.
7. Doyle S, Feldman M, Tomaszewski J, Madabhushi A: **A boosted Bayesian multi-resolution classifier for prostate cancer detection from digitized needle biopsies.** *IEEE Transactions on Biomedical Engineering* 2010, 1-14.

## S25

### Teachers' impact on dental students' exam scores in teaching pathology of the oral cavity using WSI

Janusz Szymas[1*], Mikael Lundin[2], Johan Lundin[2]
[1]Department of Clinical Pathology, Poznan University of Medical Sciences. Przybyszewski Str. 49. 60-355 Poznan, Poland; [2]Institute for Molecular Medicine Finland FIMM, P.O. Box 20, FN-00014, University of Helsinki, Finland
E-mail: jszymas@ampat.amu.edu.pl
*Diagnostic Pathology* 2013, **8(Suppl 1):**S25

**Background:** There is an overall increasing tendency at medical universities around the world to digitize microscope histopathological slides from teaching collections for web-based studies [1]. For the last seven academic years (2005 - 2012), we have developed and evaluated a user-friendly on-line interactive teaching and examination system to facilitate student access to Pathology of the Oral Cavity course material, including histological laboratory specimens. It made microscope laboratory studies in oral cavity pathology more efficient and teaching resources have become available for interactive use from anywhere at any time, independent of class schedules. The WebMicroscope allows students to independently explore entire histological slides. Students control the content and the rhythm of delivery of the content. Digital slides produced using whole slide imaging (WSI) can be visualized at any magnification and moved in the x-y axis, which perfectly imitates using a traditional microscope and a glass slide [2].

Examination which applies WSI technology on-line requires a sophisticated test management software. In the past, we used an examination system that was restricted to using standard static images. This was technically less challenging and followed a routine which was similar to that of conventional paper-based examinations [3,4]. WSI is more demanding as it requires the use of innovative image serving and viewing software and, because of the enormous size of the images, a storage facility capable of storing hundreds of gigabytes. However, it reflects much better the usual laboratory activities [5]. Using WSI allows asking questions about the overall diagnosis without specifically highlighting the key diagnostic features as it must be done in the field of view of a static image. Even if a question refers to a specific field of view, the student can zoom out and see that field in the overall context of the slide. This contextual information easily available in virtual microscopy is essential in teaching students both how to find the key diagnostic fields within a whole slide and to perceive how a diagnostic field fits into the surrounding non-diagnostic slide context. Our first seven-year experience in developing and using an examination system which applies WSI indicates that configuring an on-line practical exam in Pathology of the Oral Cavity requires flexibility over the question types that can be asked.

Our practical examination in oral pathology is scheduled at the end of the 3rd year dental curriculum of students in the Department of Clinical Pathology at the University of Medical Sciences in Poznan, Poland. Because the number of new students enrolling every academic year is

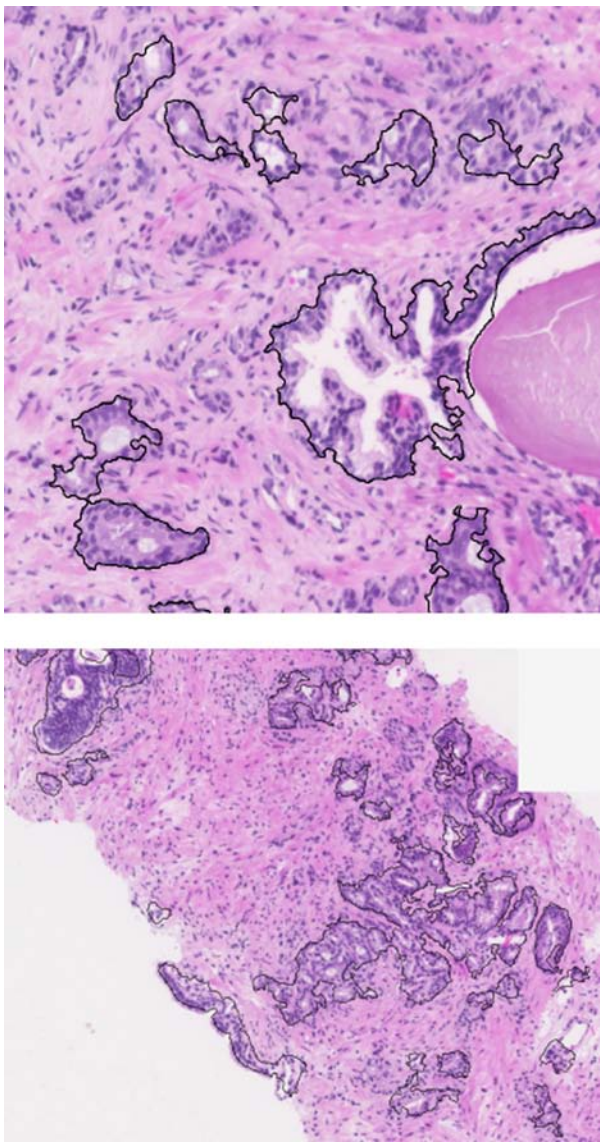


**Figure 2(abstract S24) Examples of segmented region of interests.** Results of the segmentation algorithm applied to prostate biopsy samples digitalized at 10x.

high, students work in groups which are supervised by different teaching assistants with heterogeneous teaching experience. The aim of this study was identifying if in such a context the teacher has still an impact on the students' exam scores.

**Material and methods:** Each practical exam consisted of 50 multiple-choice questions per student. Every question was displayed together with the adequate whole slide image and students had to provide an answer based on the interpretation of this slide. The scope of whole slide images were broad, designed to include all slides used during the laboratory practicals in the Pathology of the Oral Cavity. Questions and whole slide images were presented in a different order to every student. For the examination in the academic years 2008 - 2012, we used 50% of randomly rotated whole slide images to prevent "fossilization" of the students. Students were shown how to use the software at the beginning of the practical exam and were given unlimited time to complete the exam. There was no time limit per slide either. Students could also view slides or answer questions in any order. All sessions of the practical exam using WSI took place in a secure location, the computer laboratory at the Department of Pathology, and were proctored all the time by teaching assistants. All examinations were graded using the same scale. Participants were all dental students enrolled to the Pathology of the Oral Cavity course between 2005 and 2012.Students were also required to complete a short anonymous survey at the time of the exam.

**Server architecture:** The WebMicroscope was used in the current examination system to host, manage, and deliver WSI-based digital slides to the students during the practical examination. In this context, a dedicated Web viewer interacts with an image server, which serve out appropriate regions of the slide image. In this way, students can view whole slides in real time without the need to download entire images. For the practical examination in the Pathology of the Oral Cavity, multiple image servers operated at least in tandem (Poznan, Helsinki) allowing to process multiple requests for images per second.

**Client module software:** The client module is responsible for presenting data to the student. It is the only part of the system visible to the user and consists mainly of an independent application which communicates with the other modules of the system through well-defined interfaces. The messages of the front-end virtual-slide viewer were written based on a native, downloadable browser plug-in, or java script based dynamic HTML. In the latter case, it does not require any specific downloads to the client work station.

**Statistical analysis:** The results from all students were collected and analysed automatically by the Examination module and the Statistic module of the examination system software. Already at the end of the examination session, the administrator could download results for single participants, groups of students, or the whole cohort. In parallel the software automatically compiled the data on chosen answers to multiple-choice questions and performed statistical analysis on them. Here we provide descriptive statistics and t-test analysis using the analysis dataset from Pathology of the Oral Cavity examinations.

**Results:** Students were surveyed regarding their preferences for the use of virtual microscopy in the curriculum based on their first exposure to this technology. The survey results indicate that an overwhelming majority of students found the WebMicroscope system relatively easy to use and helpful. Their answers to the question about the usefulness of the WSI-based examination system (rating from 0 to 10) were constant throughout the seven years of this study. The lowest mean value (8.4 ± 1.4) was recorded in 2006 and the highest (9.1 ± 1.7) in 2009 and later, after some improvements.

In the first examination, in 2006, a cohort of 8 students scored with a 98% ± 2.7%. Similar result was obtained one year later with a cohort of 84 students ( 97,2% ± 4.7%). At this time, 50 whole slide images were used in 2006 and 64 whole slide images in 2007 during laboratory practices and during the practical examination. With a growing number of whole slide images used in the subsequent years (81 whole slide images used in 2008 and 125 in 2011), the rate of correct students' answers to the exam questions was only a little bit lower - circa 94% (Table 1). In the last four academic years (2008 – 2012), we used 50% of randomly rotated whole slide images. The rotation did not influence the performance of the students.

When analyzing the students' scores by student groups supervised by different teaching assistants, we found that the mean exam results were different between the groups. There were some student groups which scored high and some groups with lower grades (Table 2). Differences between these student groups were statistically significant (P-value < 0.050) in all academic years except for 2009-10. Unfortunately, the rotation of teaching assistants in our Department is high but during the period of this study we could identify at least three teachers whose groups have permanently scored high or the highest during the exam. We have also noted a couple of unsuccessful teaching assistants.

**Discussion:** Testing large numbers of students in the exam context with many virtual slides is a challenging operation with many potential bottle necks. To overcome some of them, whole slide images were stored on an independent server. Serving these slides as part of the practical examination is bandwidth and server intense, particularly if the image latency is to be kept to a minimum with many students accessing the images simultaneously. A successful examination system requires multiple image servers which are carefully load balanced which implies that, when one server gets busy due to user activity, another server takes over and provides the services that are being requested. For this practical

**Table 1(abstract S25) Computer-administered practical exam performance using WSI**

| Year | | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|---|
| Number of students | | 8 | 84 | 92 | 96 | 101 | 83 |
| Number of questions | | 50 | 64 | 50 | 50 | 50 | 50 |
| Number of whole slide images | | 50 | 64 | 81 | 82 | 103 | 125 |
| Correct students' answers | AVG | 98% | 97.2% | 92.1% | 93.8% | 93.4% | 93.9% |
| | STDEV | 2.7% | 4.7% | 4.4% | 5.6% | 8% | 5.2% |
| | RANGE | 88% - 100% | 73% - 100% | 76% - 100% | 78% - 100% | 44% - 100% | 74% - 100% |

**Table 2(abstract S25) Students' scores by students groups**

| Year | Students' group | Number of students | Correct students' answers | | | | Number of students with all corrected answers |
|---|---|---|---|---|---|---|---|
| | | | AVG | STDEV | MIN | MAX | |
| 2009 | I | 15 | 44.60 | 4.64 | 33 | 49 | - |
| | II | 16 | 47.18 | 1.97 | 43 | 50 | 1 |
| | III | 16 | 47.50 | 2.19 | 44 | 50 | 4 |
| | IV | 17 | 48.41 | 2.12 | 44 | 50 | 6 |
| | V | 16 | 46.19 | 2.53 | 41 | 50 | 2 |
| | VI | 16 | 47.31 | 1.49 | 45 | 50 | 3 |
| 2010 | I | 18 | 47.44 | 2.93 | 39 | 50 | 2 |
| | II | 14 | 47.67 | 2.29 | 42 | 50 | 4 |
| | III | 15 | 45.27 | 4.54 | 37 | 50 | 2 |
| | IV | 17 | 47.23 | 3.17 | 38 | 50 | 3 |
| | V | 20 | 45.55 | 6.37 | 22 | 50 | 3 |
| | VI | 17 | 46.70 | 2.52 | 40 | 50 | 1 |
| 2011 | I | 14 | 45.14 | 3.57 | 37 | 49 | - |
| | II | 13 | 48.38 | 1.26 | 45 | 50 | 2 |
| | III | 12 | 47.58 | 1.44 | 46 | 50 | 2 |
| | IV | 15 | 48.20 | 2.18 | 42 | 50 | 4 |
| | V | 15 | 46.13 | 3.00 | 38 | 50 | 1 |
| | VI | 14 | 46.14 | 2.98 | 39 | 50 | 2 |

examination in the Pathology of the Oral Cavity, multiple image servers operated at least in tandem (Poznan, Helsinki) allowing processing multiple requests for images per second. The fine focus feature, the small overview slide [6] and the possibility to use the WebMicroscope to view the whole slide images were all perceived favorable, and most students did not notice any delays during examination. More than 90% of responders found the WSI-based laboratory practicals and the practical exam useful and helpful to improve their understanding of Pathology of the Oral Cavity. At practical examinations, the students gave from 94% to 98% correct answers. So there is clear evidence of learning benefits from using WSI. Despite widely available self-study possibilities for the students, we were able to demonstrate differences between student groups. High and low scoring in student groups was associated with particular teaching assistants.

**Conclusions:** Despite all technical problems, the WebMicroscope system is a promising tool for practical student examination using WSI during laboratory practicals in the Pathology of the Oral Cavity. With all the people involved working hard to solve these technical problems and preparing every year exam, we are confident that testing using WSI is going to play an essential role in the future of the university education and examinations. Nevertheless, despite widely available self-study possibilities, good teachers still create a substantial value. Exam scores are helpful in identifying such teachers. This study also shows evidence that existing measures are informative about the teachers' impact on student scores.

**Abbreviations used:** WSI: whole slide imaging.

**Competing interests:** The authors have no competing interests.

**Authors contributions:** JS organised and managed the study and drafted the manuscript. ML was in charge of data management, managed WebMicroscope server and software and assisted in the manuscript, JL provided advice and expertise, and assisted in the manuscript.

**References**

1. Dee FR: **Virtual microscopy in pathology education.** *Hum Pathol* 2009, **40**:1112-1121.
2. Szymas J, Lundin M: **Five years of experience teaching pathology to dental students using WebMicroscope.** *Diagnostic Pathology* 2011, **6(Suppl 1)**:S13, doi:10.1186/1746-1596-6-S1-S13.
3. Szymas J, Gawronski M: **Multimedial data base and management system for self-education and testing the students' knowledge on pathomorphology.** *Patol Pol* 1993, **44**:183-187.
4. Wolynska B, Kaczalski M, Szymas J: **Computerized evaluation of students, knowledge in a course of pathology.** *Electronic Journal Pathol* 2000, **6(2)**.
5. Lundin M, Szymas J, Linder E, Beck H, de Wilde P, van Krieken H, Garcia Rojo M, Moreno I, Ariza A, Tuzlali S, Dervisoglu S, Helin H, Lehto VP, Lundin J: **A European Network for Virtual Microscopy - Design, Implementation and Evaluation of Performance.** *Virchows Arch* 2009, **454**:421-9.
6. Lundin M, Lundin J, Isola J: **Virtual microscopy.** *J Clin Pathol* 2004, **57**:1250-1251.

## S26

### A rich internet application for remote visualization and collaborative annotation of digital slides in histology and cytology

Raphaël Marée[1,2,3*], Benjamin Stévens[1,2], Loïc Rollus[1,2], Natacha Rocks[4], Xavier Moles Lopez[5,6], Isabelle Salmon[6,7], Didier Cataldo[4], Louis Wehenkel[1,2]
[1]GIGA-Systems Biology & Chemical Biology, GIGA-R, University of Liège, Belgium; [2]Systems and Modeling, Montefiore Institute, University of Liège, Belgium; [3]GIGA Bioinformatics Core Facility, GIGA, University of Liège, Belgium; [4]Laboratory of Tumor & Development Biology, GIGA-Cancer, University of Liège, Belgium; [5]Laboratory of Image Synthesis and Analysis (LISA), Faculty of Applied Sciences, ULB, Brussels, Belgium; [6]DIAPATH, Center for Microscopy and Molecular Imaging (CMMI), Gosselies, Belgium; [7]Department of Pathology, Erasme University Hospital, Université Libre de Bruxelles, Belgium
E-mail: raphael.maree@ulg.ac.be
*Diagnostic Pathology* 2013, **8(Suppl 1):**S26

**Background:** In the field of digital pathology and biomedical research, there is a strong need for efficient tools to build pathology atlases and to foster collaboration between researchers, pathologists (e.g. for inter-observer concordance studies) and computer scientists (e.g. for development and extensive validation of novel computer vision algorithms). Although many efforts have been made in virtual microscopy and telepathology in the recent years [1-4], many of the resulting frameworks are not fully web-based therefore limiting collaboration, or they are vendor-dependant therefore limited in terms of supported image formats, or they use proprietary modules that prevent cross-browser compatibility and seamless execution on mobile devices, or they have restricted functionalities (e.g. images can only be annotated manually with image-level tags or fixed markers), or their design limits their application domain (e.g. education only, or disease-specific). In this paper, we present a general-purpose, rich internet application using recent web technologies and integrating various open-source tools, standards and generic algorithms for remote visualization and collaborative annotation of digital slides.
**Material and methods: General architecture design:** Our application follows a representational state transfer (REST) architecture style that structures database resources and that standardizes communication interfaces. In such a setting, each resource can be referenced by a uniform resource locator (URL) and they can be located at different physical sites and updated/deleted if necessary. By following these programming guidelines, we defined a RESTful JSON application programming interface (API) to allow communication between servers and clients.
On the server-side, our underlying data model allows to create multiple projects, where each project corresponds to a specific study or experiment. A project is described by a list of authenticated users with permission rights, a list of digital slide images, an ontology definition with domain-specific, user-defined, vocabulary terms, and annotations (regions of interest) associated to digital slides and drawn by users. All project data are stored in a spatial, relational database (PostgreSQL with PostGIS extension). The core of our application uses the Grails framework based on Spring, with Groovy dynamic programming language for Java, and Hibernate framework with its spatial extension for object/relational mapping.
On the client-side (i.e. the Web client), the source code is based on model-view-controller design patterns and it communicates directly through the API to visualize and edit resources. Data can also be retrieved or updated by third-party computer programs through the API.
**Visualization tools:** In order to visualize whole-slide images at multiple resolutions in traditional web clients, we implemented the web interface with fully Javascript interfaces and libraries (including JQuery, Backbone.js and Twitter Bootstrap components). Visible parts of high-resolution images are delivered through distributed image tile servers (using IIPImage system) that supports TIFF and JPEG2000 image formats, and it was combined with the OpenSlide library to further support various digital slide image formats (Aperio SVS, Hamamatsu VMS, 3DHistech Mirax, ...). Additional caching mechanisms are implemented in-between the image servers and clients (using Varnish library) to speed up the delivery of the most frequent data.
**Annotation and collaboration tools:** In addition to remote visualization capabilities, each user of a given project can create and edit his own layer of annotation geometries (e.g. polygons, ellipses, rectangles, or freehand drawings) drawn on top of digital slide images, and visualize annotations created by others associated to the current project, using the OpenLayers library, as illustrated by Figure 1. Each region of interest can be associated to one or multiple term(s) from a structured vocabulary defined on-line by the users of each project. Within a given project, relational queries are used to filter annotations based on image names, user names, and/or ontology terms, so that annotation galleries (with cropped image regions) and statistics can be easily gathered and visualized, hence facilitating the shaping of pathology atlases. Furthermore, to ease collaboration between pathologists, an e-mailing mechanism allows sharing and discussing annotations, and a communication mechanism allows one user to follow another user's observation paths in real-time through the Internet.
**Image processing and retrieval algorithms:** We implemented image processing routines and a recent content-based image retrieval (CBIR) algorithm to speed up the exploration and annotation of digital slides. The image processing routines are based on ImageJ/FIJI plugins and include various image filtering operations (e.g. binarization, splitting color channels, and color deconvolution) that can be applied on-the-fly on image tiles to ease image inspection, and adaptive thresholding operations that can be used to semi-automatically draw annotation geometries around objects of interest. The CBIR algorithm uses random subwindow extraction and vectors of random tests on raw pixel values [5]. It is used to search visually similar annotations and automatically suggests

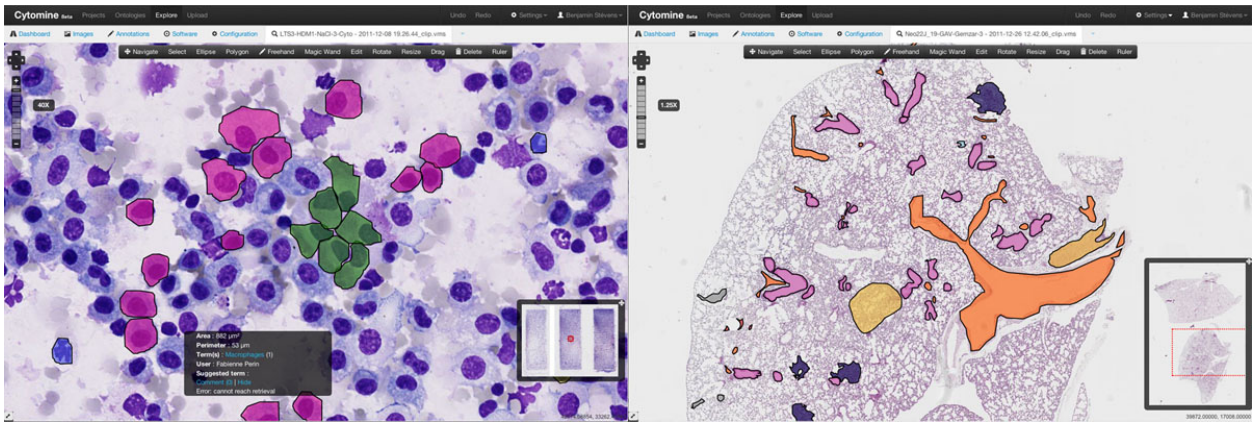



**Figure 1(abstract S26) Viewing and annotating whole-slide images** Exploration and annotation of whole-slide cytology (left) and histology (right) images from experimental mice lungs. Colors correspond to different ontology terms associated to annotations.

ontology terms through an average voting scheme based on computed image similarities with cropped images of previously indexed annotations. We implemented the CBIR algorithm using an efficient key-value store based on hash tables (using Kyoto Cabinet or Redis NoSQL databases).

**Results and discussion: Results:** Our application runs in any popular web browsers and on mobile devices without the need for proprietary browser add-ons. It has been used for one year by our collaborators from two geographically distant locations through the Internet. About one thousand whole-slide images of lung cancer studies (corresponding to roughly 1TBytes of data) acquired by two slide scanners (Olympus VS100 with 20X magnification and Hamamatsu Nanozoomer 2.0 with 40X magnification) have been uploaded. These include Hematoxylin&Eosin (H&E) stained histology images of experimental mice, and bronchoalveolar lavage (BAL) cytology images. Three ontologies describing various tissue types (e.g. bronchus, blood vessel, cartilage, adenocarcinoma, nodular lymphoid hyperplasia,. ..) and various cell types (e.g. squamous epithelial cells, macrophages, eosinophils, neutrophils, mucosecreting cells, ciliated bronchial cells,...) were defined and used by seven users (pathologists, pneumologists, and technicians) to annotate more than five thousand regions of interest.

Once a user has drawn an annotation geometry on an image, our CBIR algorithm automatically suggests ontology terms in roughly 500 ms by searching visually similar regions of interest in the database of all users' annotations, as illustrated by Figure 2. Automatic term suggestions match the ground-truth (manually associated vocabulary terms) in roughly 70% for lung H&E histology images (n=1509), and 89% for lung BAL cytology images (n=675), using the same algorithm parameters for both image types as in [5] (ie. 1000 random subwindows encoded in HSV colorspace, 10 vectors of random tests, and 30 binary tests per vector). Inspection of confusion matrices reveals that per class recognition rates vary from 0% (for mucosecreting cells often misclassified as ciliated bronchial cells) to 95% (for macrophages and polynuclear neutrophils) in our cytology images, and from roughly 40% (for nodular lymphoid hyperplasia often misclassified as adenocarcinoma, and blood vessels often misclassified as bronchus) to 95% (for bronchus and red-blood cells) in our histology images.

**Discussion:** Although the amount of data our software is already dealing with is rather large, it is expected that the wider adoption of digital acquisition equipments will generate much larger datasets. The design of our software allows its scaling to larger sets of images as most of the components (e.g. image servers, and image retrieval algorithm) can be distributed on multiple machines. It is also important to note that the architecture allows local configurations, ie. images and data have not to be stored on a central, external, server but they can remain on servers at local institutions, therefore ensuring confidentiality and local administration. It is worth noting that although we do not rely on latest standard definitions in digital pathology ([1]), we plan to extend our software to support these standards once they will be implemented in the field. In the future, our architecture will also allow us to add new image formats without affecting the source code of the core application.

Regarding our preliminary evaluation with the CBIR algorithm, our results are promising for automatic term suggestion but further validation has to be conducted. Indeed, recognition rates for less frequent object types are lower, stressing the need for more manual annotations with respect to object types (e.g. mucosecreting cells were six times less frequently annotated than ciliated bronchial cells), and acquisition protocols (such as color stainings).

**Conclusions:** The proposed web software is generally applicable and its methodological choices open the door for large-scale distributed and collaborative image annotation and exploitation projects. Future work includes the integration and validation of general-purpose machine learning techniques to further facilitate annotation and quantification of specific visual phenotypes and to support their meta-analysis. We also plan to extend our framework to other types of multidimensional imaging data related to other diseases or biological processes, and we also intend to adapt its use for education purposes.

**Competing interests:** The author(s) declare that they have no competing interests.

**Authors' contributions:** RM, BS, and LR contributed to the general design of the software. BS and LR carried out most of the software programming work. NR, DC, XML and IS have been involved in functional specifications as well as acquisition and annotation of imaging data. RM coordinates the study and drafts the manuscript. All authors read, edited, and approved the final manuscript.

**Acknowledgements:** This work is funded by the research grant n° 1017072 of the Walloon Region (DGO6). RM is also supported by the

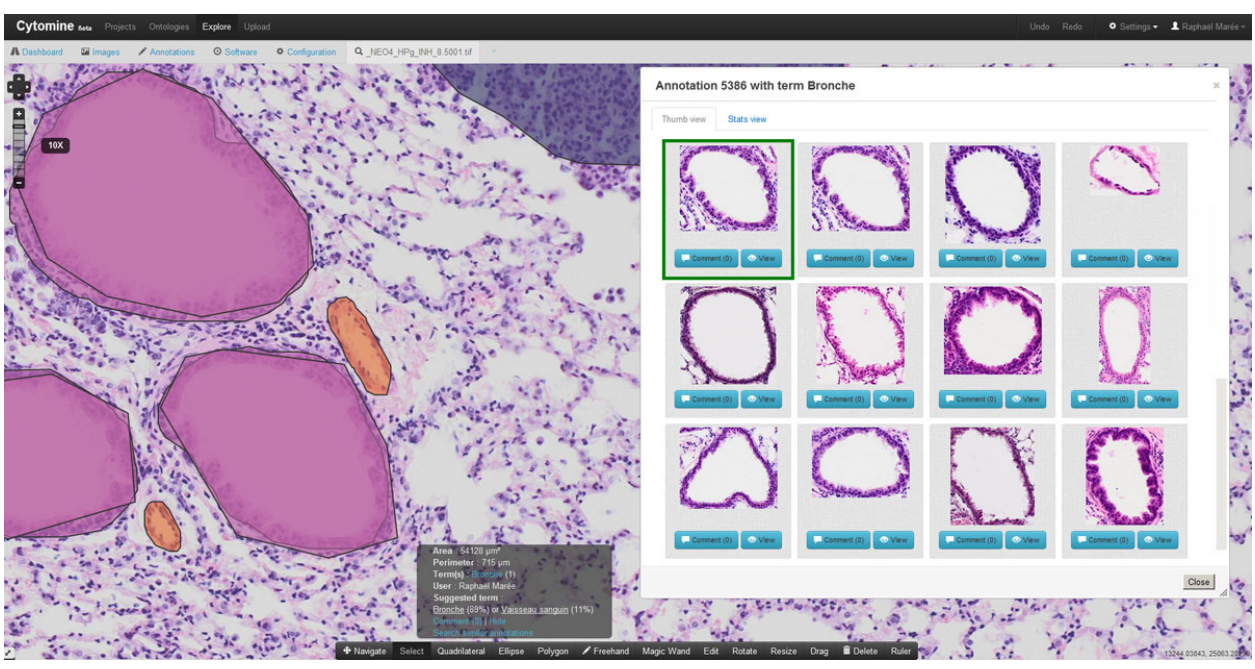



**Figure 2(abstract S26) Automatic ontology term suggestion** Once a user draw an annotation geometry, our CBIR algorithm is executed to retrieve the most similar annotations based on visual image content and suggest automatically an ontology term using average voting. The process is illustrated here with a bronchus of a lung tissue where the ten first retrieved annotations are bronchus somehow similarly oriented.

GIGA with the help of the Walloon Region and the European Regional Development Fund. The CMMI is supported by the European Regional Development Fund and the Walloon Region. XML is supported by the "Télévie" program of the "Fond National de la Recherche Scientifique" (FNRS). We thank Fabienne Perin, Christine Fink, Myriam Remmelink, and Sandrine Rorive for continuous software testing.

**References**

1. Daniel C, Macary F, Rojo MG, Klossa J, Laurinavicius A, Beckwith BA, Della Mea V: **Recent advances in standards for collaborative Digital Anatomic Pathology.** *Proc.10th European Congress on Telepathology and 4th International Congress on Virtual Microscopy, Diagnostic Pathology* 2011, **6(Suppl 1)**:526.
2. Della Mea V: **25 years of telepathology research: a bibliometric analysis.** *Proc.10th European Congress on Telepathology and 4th International Congress on Virtual Microscopy, Diagnostic Pathology* 2011, **6(Suppl 1)**:526.
3. Triola MM, Holloway WJ: **Enhanced virtual microscopy for collaborative education.** *BMC Medical Education* 2011, **11**:4.
4. Jeong W-K, Schneider J, Turney SG, Faulkner-Jones BE, Meyer D, Westermann R, Clay Reid R, Lichtman J, Pfister H: **Interactive Histology of Large-Scale Biomedical Image Stacks.** *IEEE Transactions on Visualization and Computer Graphics* 2010, **16(6)**:1386-1395.
5. Marée R, Denis P, Wehenkel L, Geurts P: **Incremental Indexing and Distributed Image Search using Shared Randomized Vocabularies.** *Proc. 11Th ACM International Conference on Multimedia Information Retrieval* 2010, 91-100.

## S27

### A simple mathematical model utilizing topological invariants for automatic detection of tumor areas in digital tissue images

Kazuaki Nakane[1*], Yasunari Tsuchihashi[2], Nariaki Matsuura[1]
[1]Osaka University, the graduate school of medicine, division of health science, Yamadaoka, Suita, Osaka, 565-0871, Japan; [2]Louis Pasteur center for medical research, Kyoto, department of clinical pathology research, Taniguchi Kakinouchi-Cho, Ukyo-Ku, Kyoto, 616-8012 , Japan
E-mail: k-nakane@sahs.med.osaka-u.ac.jp
*Diagnostic Pathology* 2013, **8(Suppl 1):**S27

**Background:** Pathological diagnosis starts usually from detection of morphological deviations of tissues and cells in question from their normal counterparts. Measurement and analysis of how they are deviated are the processes of actual pathological diagnosis and many of them can be achieved by computer using certain mathematical models. Computer assisted pathological diagnosis is now an issue of importance in the current situations of shortage of diagnostic pathologists in Japan. In the present study we propose a simple mathematical model to differentiate tumor tissues from their normal counterparts utilizing changes in *the Betti numbers* in tumorigenesis. We tested our method in several normal and tumor tissue images and preliminary results are reported.

**Materials and methods: Mathematical tools:** *The Betti numbers* are the numbers coined after Enrico Betti, an Italian mathematician of topology. *The Betti numbers* are one of the invariants in a homology. The invariant implies the quantity that is unchangeable by continuous transformation ([1]).

We treat two dimensional pathological images, *the Betti numbers* are consisting of two numbers. One is *b0* (*the 0-dimensional Betti number*), which is the number of such isolated solid component as each cell or cell nucleus. The other is *b1* (*the 1-dimensional Betti number*), which is the number of windows in the fenestrated area. That area is created by incomplete fusion of neighboring isolated solid component. The schematic illustration is shown in Figure 1.

**Morphological changes during the course of tumorigenesis:** Normal cells are usually controlled so as to grow and be reproduced only when a tissue composed of the cells requires new cells. That is, a new cell is generated when the old cell dies or is damaged. However, if the foregoing process of cell growth and cell division becomes out of order, the cells grow and divide, excessively. The excessive cells then produce a mass of tissues called "a tumor" or "a neoplasm". The disordered cell growth causes a variety of morphological characteristics of tumor tissues.

**Mathematical representation of morphological change in tumorigenesis:** Uncontrolled cell growth usually makes cell nuclei become larger. This causes a change in contact condition between components in a tissue. This change appears along with neoplastic changes and tumorigenesis. The following description deals with how to numerically express the changes.

As shown in Figure 1, when the number of contact points between individual solid components (cells or cell nuclei) increases, i.e., 7 solid components (*b0*) fuse each other through contact, 2 solid components with fenestration are created. They have 4 windows (*b1*). Namely, *b1* increase and the ratio *b1/b0* significantly change.

**Numerical procedure of our method:** Hematoxylin and eosin stained mucosal biopsy sections of colonic tumors are used as test samples. First they are taken as whole slide images and then binarized setting a threshold value depending on each HE stained condition. We regard the binarized image as a sum of *a simplical complex*. By using a free calculation software (cf. [2]), we can calculate *the Betti numbers*. We compare the values, *b1* and *b1/b0*, of colonic tumor tissues with those of normal tissues.

In Figure 2, we show a digital image of tissue of colon and its binarized image. In the tumor area, we can see many windows.

**Results and discussion:** *The Betti numbers* are obtained per unit binarized test image areas of normal and tumor tissues of colon. The values, *b1* and *b1/b0*, of representative each five unit areas of normal and tumor colonic tissues are listed in Table 1.

In normal tissue, *b1* is less than 1400 (565.6 average), *b1/b0* is less than 1.4 (0.251 average). If tissue is malignant, *b1* is large (3266.4 average), *b1/b0* is increased (3.79 average).

Based on the results obtained above, the indices for a normal tissue and a tumor tissue can be assumed.

Figure 3 is a result of this method. The image is divided into 7×7 unit areas and the marks are put on each unit area as Table 2.

**Conclusion:** *The Betti numbers* can express the degree of the connection of the figure. When the number of contact points between individual components increases, irrespective of their shape, *b1* and a ratio *b1/b0* significantly change. So we hypothesize that these numbers can be used as indexes to represent the cellular "accumulation" which is one of the characteristics of tumor tissue.

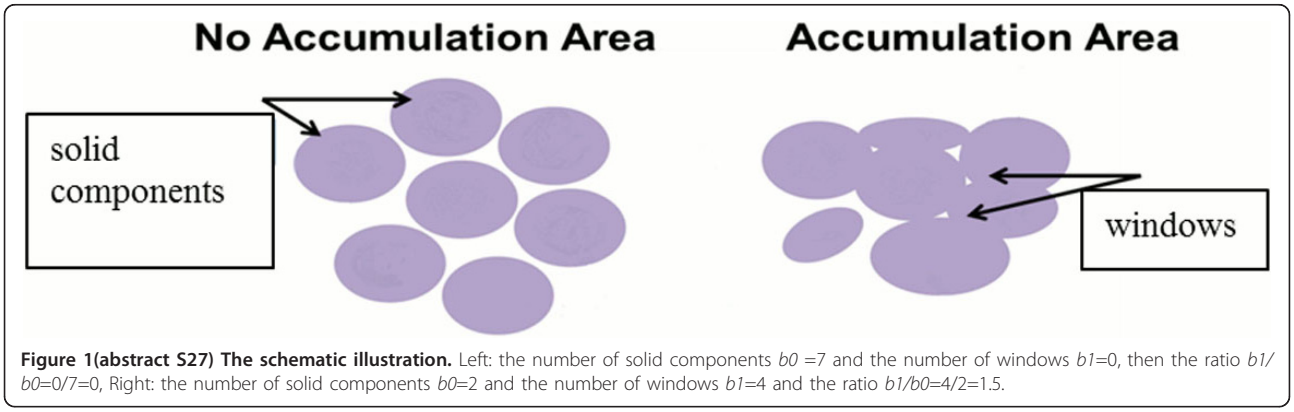



**Figure 1(abstract S27) The schematic illustration.** Left: the number of solid components *b0* =7 and the number of windows *b1*=0, then the ratio *b1/b0*=0/7=0, Right: the number of solid components *b0*=2 and the number of windows *b1*=4 and the ratio *b1/b0*=4/2=1.5.

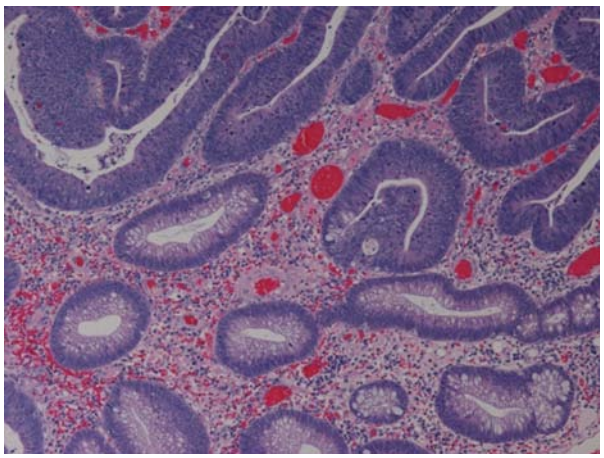


**Figure 2(abstract S27)** A digital image of tissue of colon.

**Table 1(abstract S27) *b1* and *b1/b0* of the normal and tumor tissue**

| Normal tissue | | | Tumor tissue | | |
|---|---|---|---|---|---|
| | *b1* | *b1/b0* | | *b1* | *b1/b0* |
| a015 | 1300 | 0.496 | a012 | 4045 | 2.16 |
| a017 | 233 | 0.106 | a034 | 3469 | 2.65 |
| a045 | 903 | 0.204 | a047 | 2377 | 5.23 |
| a049 | 283 | 0.204 | a067 | 3044 | 5.92 |
| a050 | 109 | 0.0689 | a070 | 3397 | 3.02 |
| Ave. | 565.6 | 0.251 | Ave. | 3266.4 | 3.79 |

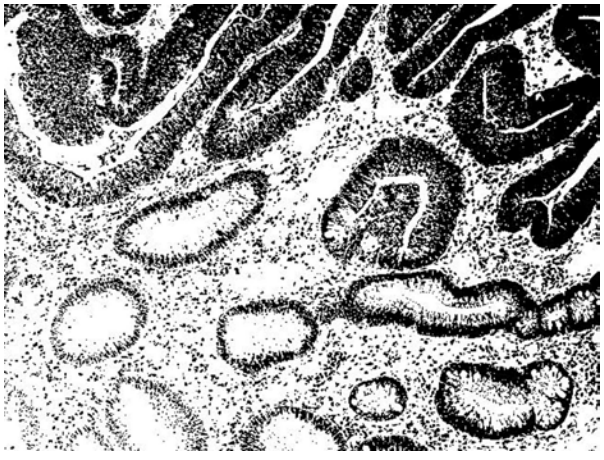


**Figure 3(abstract S27)** The binarized image of Figure 2.

The numerical results indicate that the difference in our indices can differentiate tumor tissues from their normal counterparts.

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** KN carried out most of experiments, participated in the design of the study and drafted the manuscript. TY and MN participated in the design of the study and helped the manuscript. All authors have read and approved the final manuscript.

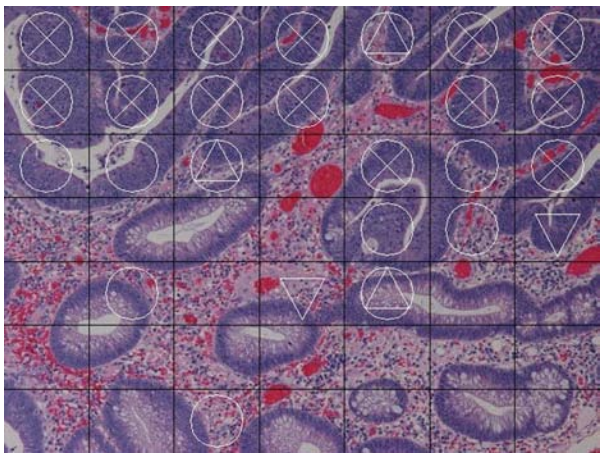


**Figure 4(abstract S27)** The image is divided into 7×7 and impose the marks as the result.

**Table 2(abstract S27) Marks are put on each unit area as the value of index**

| *b1* | None (~ 1400) | ▽ (1400 ~ 2000) | × (2000 ~ ) |
|---|---|---|---|
| *b1/b0* | None(~ 1.4) | △ (1.4 ~ 1.7) | ○ (1.7 ~ ) |

**Acknowledgement:** We would like to take this opportunity to thank Dr. Nagumo (research professor of graduate school of medicine, Osaka University) for her valuable advice from a pathological standpoint. Further, we would like to thank Prof. Suematsu J. Nobuhiko (graduate school of advanced mathematical science Meiji University) for his valuable advice on how to display calculated results on an image.

**References**

1. Arkowitz M: **"Introduction to Homotopy Theory (Universitext)".** Springer Press 2011.
2. **CHomP (Computational Homology Project).** [http://chomp.rutgers.edu/].

## S28

### Telepathology in cervical and breast cancer screening programmes

Romano Colombari[1*], Roberto Mencarelli[2], Alessio Gasparetto[3], Carla Cogo[4], Antonio Rizzo[5]
[1]Anatomic Pathology, Fracastoro Hospital, ULSS 20, Verona, Italy; [2]Clinical Pathology Department, Local Healthcare Authority, Rovigo, Italy; [3]Information Technology Department Local Healthcare Authority, Rovigo, Italy; [4]Cancer Registry Veneto Region, Italy; [5]Anatomic Pathology, S. James Hospital, ULSS 8, Castelfranco Veneto (Tv), Italy
E-mail: rcolombari@ulss20.verona.it
*Diagnostic Pathology* 2013, **8(Suppl 1):**S28

**Background:** This study is about the application of telepathology in external quality assurance (EQA) for cervical and breast cancer screening programmes. In Veneto Region, Italy, an organized effective mass-screening programme for cervical and breast cancer has been active since 1998. The previously experienced EQA with slide set circulation on a voluntary base by part of the Anatomic Pathology laboratories of this Region had failed because of several complications: slide's circulation around more than 20 Anatomic Pathology services takes a great deal of time; there is a high risk of slide loss and/or break down during the multiple slides' deliveries; patients' privacy and operators' anonymity can be regrettably compromised.
The EQA has recently become a mandatory requirement for accreditation of laboratories providing screening services within Veneto.
The whole image acquisition technology of cytological and histological slides, thank to innovative software, has become available in the last few years. These applications permit pathologists to storage digital images and to examine, on digital screen, virtual slides scanned in a far away

laboratory, using a viewer downloaded from internet in an easy and, in most cases, free way.
In 2008 a slide scanner (Aperio ScanScope-XT) was installed at Anatomic Pathology Laboratory in Rovigo, Italy. This technical tool, with the proper ICT (Information and Communication Technology) configurations and architecture, became available to the Anatomic Pathology Units in Veneto.
Such technology is still in early application stage, so that we evaluate its potential use in EQA for cervical and breast cancer screening programmes.
**Material and methods:** In 2008 a committee (one for cervical and one for breast cancer screening) has been elected among the pathologists involved in Veneto mass-screening programme for cervical and breast cancer to plan the EQA.
From 2009 to 2011 a year topic has been chosen among those hot in the field of cytopathology and histopathology of cervical and breast cancer screening.
Every year, in Veneto each laboratory participating in the project selected from routine files 1 to 3 cervical and breast cytologic and histologic samples. The selected samples were mailed to Clinical Pathology Department in Rovigo, where they have been anonymized and digitalized by using a ScanScope-XT. The images have been stored on a virtual slide repository available online for a web consultation. The personalized free access has been made available on website [1]. The experience with virtual slide technical tools was very different among the participant pathologists. No training, focused on virtual pathology, was given to the participant pathologists before the virtual slide-based EQA project started. Remote ICT support, when requested, was available to each participant. The project manager of each screening group collected the anonymized diagnoses. Each committee provided a gold standard diagnosis. The discordant cases have been proposed for plenary discussion in a year meeting, in the attempt to reach agreement among the participants.
**Results:** Twenty-three public laboratories out of twenty-four active in Veneto joined the project. From 2009 to 2011, virtual slides have been created from 98 cervical smears and from the 102 corresponding biopsies. Similarly, virtual slides have been created from 52 breast cancer samples, 48 breast needle core biopsies and 95 breast FNAC. On virtual cervical slides, a total of 1717 cytological diagnoses and 1719 histopathological diagnoses have been obtained; on virtual breast slides, a total of 1717 cytological diagnoses and 1822 histopathological diagnoses have been obtained.
In 2009 only 59% of the participant laboratories for cervical EQA and 80% for breast EQA could successfully evaluate the virtual slides. Most of the participants couldn't download and install the viewer or complained the low performances during slides' evaluation.
In 2009 some non-sense incoherent diagnoses have been registered.
In 2011 the virtual slides have been successfully evaluated by all participants.
**Discussion:** EQA has become an integral part of mass-screening programme development in Italy. Ideally, the scheme should include slide set circulation, but several obstacles prevent the adequate diffusion of such practice. The EQA based on circulating slides is still practiced but in only a few "niche" applications [2]. For this reason, the use of digital slides would represent a helpful alternative for the EQA [3,4].
The major issues limiting the use of virtual slide-based EQA in the first years of our EQA project did not involve image acquisition or quality but rather the pathologist's experience with virtual slide technical image management and several issues such as the pathologist's interface and the hospital's network. The need for standardization of technical elements of image has already been pointed out [5,6]. Moreover, very recently these technical standards in the contest of mass-screening programme have been established by a preeminent European committee [7].
Several challenges can be pointed for the next few years to allow the wide and easy application of virtual pathology in EQA: on one hand there is the need to focus on pathologist's training with virtual slide technical tools; on the other hand, there is the need to improve the hospital's network, to homogenize the security policies, to adequate the technical tools of each Anatomic Pathology Unit to the recently proposed standard [7].
These results are encouraging to purse this workgroup, able to involve the participants in screening programmes, with a very good cost/benefit ratio.
**List of abbreviations:** EQA: external quality assurance; ICT: information and communication technology.
**Competing interests:** The Authors declare that they have no competing interests.
**Authors' contributions:** RC and AR: scientific coordinators and project managers
RM: virtual slide quality supervisor
AG: ICT support
CC: continuing medical education regional coordinator
Veneto cervical cancer screening group: cervical screening committee
Veneto breast cancer screening group: breast screening committee
Veneto cervical cancer screening group: Romano Colombari and Enzo Bianchini, Laura Borghi, Loredana Bozzola, Concetta Chiarelli, Duilio Della Libera, Licia Laurino, Lorella Marchioro, Barbara Pertoldi, Antonella Pinarello, Paola Rossi.
Veneto breast cancer screening group: Antonio Rizzo and Enzo Bianchini, Laura Borghi, Roberta Boschetto, Giuseppe Briani, Andrea Caneva, Elisa Canova, Giovanni Capitanio, Duilio Della Libera, Stefania Dante, Orfeo Del Maschio, Roberto Di Pietro, PierPaola Gasparini, Licia Laurino, Genesio Leo, Gigliola Ludovichetti, Paolo Machin, Lorella Marchioro, Marcello Lo Mele, Enrico Orvieto, Salvatore Paolino, Ida Pavon, Antonella Pinarello, Quirino Piubello, Domenico Reale, Daniela Reghellin, Vittorio Rucco, Debora Tormen.
**Acknowledgements:** The study was supported by a grant of Veneto Region.

**References**

1. Online telepathology service ULSS 18 Rovigo. [https://servizi.azisanrovigo.it].
2. Confortini M, Bondi A, Cariaggi MP, Carozzi F, Dalla Palma P, Ghiringhello B, Minucci D, Montanari G, Parisio F, Prandi S, Schiboni ML, Ronco G: **Interlaboratory reproducibility of liquid-based equivocal cervical cytology within a randomized controlled trial frame work.** *Diagn Cytopathol* 2007, **35**:541-544.
3. Leong FJ, Graham AK, Schwarzmann P, McGee JO: **Clinical trial of telepathology as an alternative modality in breast histopathology quality assurance.** *Telemed J* 2000, **6**:373-377.
4. Lee ES, Kim IS, Choi JS, Yeom BW, Kim K, Han JH, Lee MS, Leong AS-Y: **Accuracy and reproducibility of telecytology diagnosis of cervical smears: a tool for quality assurance programs.** *Am J Clin Pathol* 2003, **119**:356-360.
5. Yagi Y, Gilbertson JR: **Digital imaging in pathology: the case for standardization.** *J Telemed Telecare* 2005, **11**:109-16.
6. Klaus K, Gortler J, Goldmann T, Vollmer E, Hufnagl P, Kayser G: **Image standards in tissue-based diagnosis (diagnostic surgical pathology).** *Diagn Pathol* 2008, **3**:17.
7. Ellis I: **Quality Assurance Guidelines for Breast Pathology Services (second edition).** *Sheffield NHS Breast Cancer Screening Programme* NHSBSP publicationISBN 978-1-84463-072-1 2011, n° 02.

## S29

### Automated classification of breast cancer morphology in histopathological images

Ville Ojansivu[1*], Nina Linder[1], Esa Rahtu[2], Matti Pietikäinen[2], Mikael Lundin[1], Heikki Joensuu[3], Johan Lundin[1]
[1]Institute for Molecular Medicine Finland-FIMM, Helsinki, Finland; [2]University of Oulu, Center for Machine Vision Research, Oulu, Finland; [3]Department of Oncology, Helsinki University Central Hospital, Helsinki, Finland
E-mail: ville.ojansivu@ee.oulu.fi
*Diagnostic Pathology* 2013, **8(Suppl 1):**S29

**Background:** The morphology of a breast cancer tumour, as examined through an optical microscope, is currently assessed visually by the pathologist in parallel with making the cancer diagnosis. The grade of differentiation, which describes how closely the morphology of the tumour resembles the corresponding healthy tissue of an organ, is undisputedly related to the outcome of breast cancer [1]. However, tumour grade is largely regarded as an unreliable prognostic factor due to its poor reproducibility [2]. The visually determined morphology is afflicted with a poor inter- and intra observer agreement, which prevents grade from being fully utilized as an important outcome predictor. The same pathologist may assign different grade to the same tumour when assessment is repeated, and different pathologists disagree to a substantial level when assessing the same tumour [3].
Computational diagnostic tools for estimating the morphological properties of cancer tissue would enable objective and reproducible

alternative for diagnosis. This could be achieved by fully utilizing the recent advances in digital microscopy and computer vision [4,5]. Some attempts have already been made for automated grading of histopathological breast cancer images, but these studies have covered only limited amount of data or produce just a partial grading [6,7]. We propose a texture based algorithm for automated classification of breast cancer morphology. The method uses the recently introduced LPQ [8] as well as LBP [9] descriptors and an SVM classifier. The LPQ and LBP descriptors each form a histogram representing the statistical texture properties and have been used earlier in many texture analysis applications which include surface inspection [9], tissue analysis [5], and face recognition [8], whereas SVM represents the state of the art among supervised learning based classification algorithms.

**Material and methods:** The image data set (n=1092) was extracted from a series of digitized, whole-slide tissue microarray (TMA) samples from a nationwide cohort of breast cancer patients, FinProg [10]. A single continuous area that contains only tumor tissue was defined in each representative tissue spot in the hematoxylin-eosin (HE) stained TMA samples. The original tissue spots fit into an approximately 1600 x 1600 pixel image while the size of the defined square areas was varying with dimensions in the range 400–1400 pixels. The images were scored by a human observer into three classes according to morphology: 1 (morphology resembling normal breast epithelium, extensive tubular formation, n=182), 2 (intermediate tubular formation, n=494), and 3 (morphology least resembling normal breast epithelium, no tubular formation; n=416). Examples of the three classes are illustrated in Figure 1.

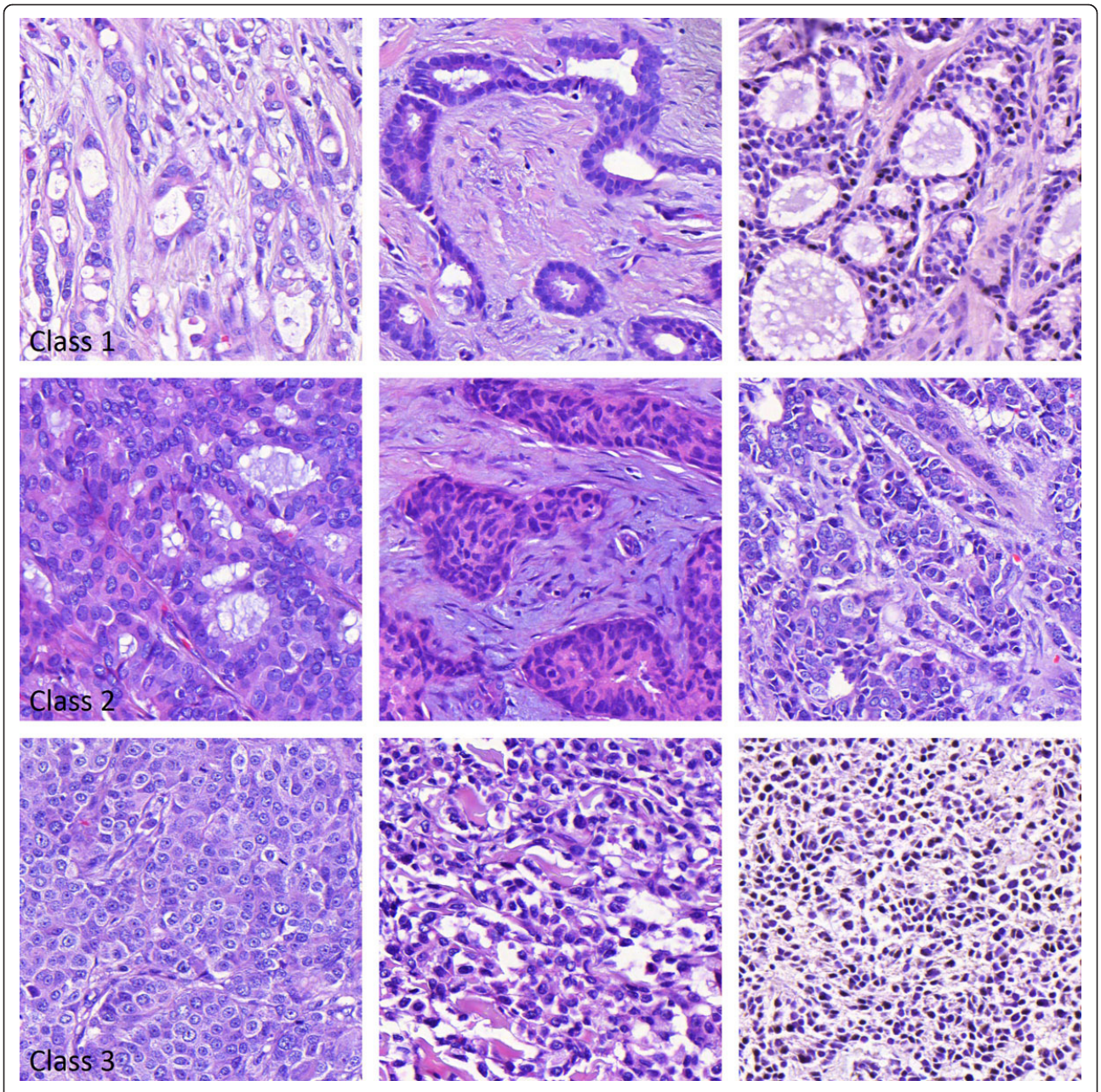



**Figure 1(abstract S29) Examples of tissue images from three morphological classes** Class 1 (top row): morphology resembling normal breast epithelium, extensive tubular formation, n=182; Class 2 (center row): intermediate tubular formation, n=494; and Class 3 (bottom row): morphology least resembling normal breast epithelium, no tubular formation, n=416. Images are classified into the three classes by a human observer.

The images were transformed to gray scale and represented by LBP [7] and LPQ [8] texture descriptors. The classification of the images into the three classes was done using three one-versus-rest SVM classifiers with a radial basis function kernel (RBF) combined with chi-square distance metric. The final class was chosen by selecting the largest of the scores produced by the individual SVM classifiers. Given the training samples and their classes, an SVM classifier learns a model for the data which aims to separate the classes in space with a margin. In testing phase, the SVM classifier assigns new data samples into the classes based on the learned model. In our experiments, the data was split into two halves for training and testing of the SVM classifiers. We did additional experiments with only the extreme class 1 and 3 samples. In this case, we used the same descriptors and a binary SVM classifier with an RBF kernel.

**Results and discussion:** The experiments were performed using different combinations of LBP and LPQ descriptor variants as well as by various scales of the images. The best classification results were achieved by combining the basic versions of LPQ and LBP descriptors with radius r=1 and number of samples p=8 into a 512-dimensional feature vector and using the original image scale 1:1. The receiver operating characteristic (ROC) curves illustrated in Figure 2, show the ratio of the "true positive" and "false positive" samples in classification when the threshold for each binary one-vs-rest SVM-classifier score is changed. The area under the ROC curve (AUC) is related to the fidelity of the classification result. The AUCs for the ROC curves were: class 1 (extensive tubule formation) vs. classes {2, 3}, 0.84; class 2 (moderate tubule formation) vs. classes {1, 3}, 0.65; and for class 3 (no tubule formation) vs. classes {1, 2}, 0.83. If each image is classified into the class with the highest SVM score, the total classification accuracy is 62.0%. The total classification accuracy was improved by 2% by using the LPQ descriptor in addition to the traditional LBP descriptor. It seems that the separation of intermediate class 2 from the classes 1 and 3 is the most challenging task. This is understandable since image content in class 2 samples is a mixture of the two neighbouring classes 1 and 3. If it would be enough to separate only the extreme morphological classes 1 and 3 neglecting the class 2, a single binary SVM classifier could be used. For this class 1 vs. class 3 classifier AUC is 0.95 which is remarkably better than the results for the one-vs-rest classifiers. The accuracy of class 1 vs. class 3 classifier is 90 % (when threshold=0 for SVM score is used). One option for better separation of class 2 could be to do the analysis for smaller image areas which would be classified as class 1 or 3. Then class 2 could be found as an appropriately selected mixture of these areas.

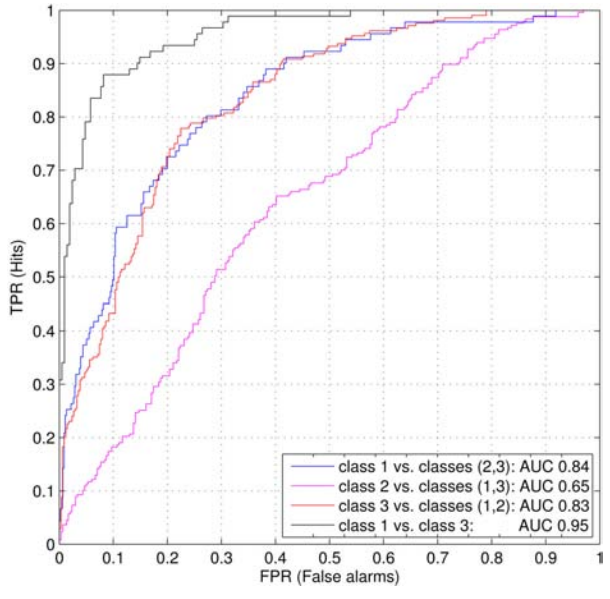



**Figure 2(abstract S29) ROC curves showing the classification performance** Receiver operating characteristic (ROC) curves for each of the three one-vs-rest SVM classifiers. Fourth ROC curve is for an SVM classifier separating classes 1 and 3. Also the relative areas under the ROC curves (AUC) are denoted which express the fidelity of the classification result.

**Conclusions:** Histological grade of breast cancer is regarded as an important prognostic factor, but not included in staging guidelines due to the subjective nature of the assessment process. In the current study, we propose a computer vision method based on texture features and a classifier utilizing supervised machine learning to discriminate between cancer morphology as determined by a human observer. The results obtained show that automated grading is feasible and that discrimination between different levels of tubule formation can be performed with moderate to high accuracy. By combining LBP and LPQ features it is possible to improve the discrimination accuracy compared to using only LBP alone. While the extreme morphological structures according to tubule formation in the breast cancer tissue are discriminated with high accuracy, the recognition of the intermediate class should still be improved.

**List of abbreviations used:** AUC: Area under the ROC curve; HE: Hematoxylin-eosin; LBP: Local binary pattern; LPQ: Local phase quantization; RBF: Radial basis function; ROC: Receiver operating characteristic; SVM: Support vector machine; TMA: Tissue microarray

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** VO designed, implemented, and tested the classification system and wrote related parts of the article. NL annotated and labeled the images used in the experiments and wrote the medical sections of the article. ER consulted significantly for the design of the algorithms. MP consulted for the design of the system. ML was responsible for handling and preprocessing of image data in the virtual microscope environment. HJ was the principal investigator of the FinProg study. JL was responsible for designing the multidisciplinary research setting as a whole. ER and JL contributed also by improving the article. All authors read and approved the final manuscript.

**Acknowledgements:** The study was kindly supported by the national Biomedinfra and Biocenter Finland projects.

**References**

1. Elston CW, Ellis IO: **Pathologic prognostic factors in breast cancer. I. The value of histological grades in breast cancer. Experience from a large study with long-term follow-up.** *Histopathology* 1991, **19**:403-410.
2. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI, Borgen PI, Clark G, Edge SB, Hayes DF, Hughes LL, Hutter RVP, Morrow M, Page DL, Recht A, Theriault RL, Thor A, Weaver DL, Wieand HS, Greene FL: **Revision of the American Joint Committee on Cancer staging system for breast cancer.** *J Clin Oncol* 2002, **20(17)**:3628-36.
3. Boiesen P, Bendahl PO, Anagnostaki L, Domanski H, Holm E, Idvall I, Johansson S, Ljungberg O, Ringberg A, Ostberg G, Fernö M: **Histologic grading in breast cancer: reproducibility between seven pathologic departments.** *Acta Oncol* South Sweden Breast Cancer Group 2000, **39(1)**:41-5.
4. Lundin M, Lundin J, Isola J: **Virtual microscopy.** *J Clin Pathol* 2004, **57**:1250-1.
5. Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Haglund C, Ahonen T, Pietikäinen M, Lundin J: **Identification of tumor epithelium and stroma in tissue microarrays using texture analysis.** *Diagn Pathol* 2012, **7**:22.
6. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC: **Automatic breast cancer grading of histopathological images.** *Conf Proc IEEE Eng Med Biol Soc* 2008, 3052-5.
7. Doyle S, Agner S, Madabhushi A, feldman M, Tomaszewski J: **Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features.** *IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 2008, **29**:496-499.
8. Rahtu E, Heikkilä J, Ojansivu V, Ahonen T: **Local Phase Quantization for Blur-Insensitive Image Analysis.** *Image and Vision Computing* 2012, doi: 10.1016/j.imavis.2012.04.001, Matlab code database [http://www.cse.oulu.fi/CMV/Downloads/LPQMatlab].
9. Ojala T, Pietikäinen M, Mäenpää T: **Multiresolution gray-scale and rotation invariant texture classification with local binary patterns.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002, **24(7)**:971-987, Matlab code database [http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab].
10. **The FinProg breast cancer database.** [http://www.finprog.org].

## S30

### An open-source, MATLAB based annotation tool for virtual slides

Riku Turkki[1,2*], Margarita Walliander[1], Ville Ojansivu[1], Nina Linder[1], Mikael Lundin[1], Johan Lundin[1,3]

[1]Institute for Molecular Medicine Finland (FIMM), P.O. Box 20, FI-00014 University of Helsinki, Helsinki, Finland; [2]Center for Machine Vision Research, Department of Computer Science and Engineering, P.O Box 4500, FI-90014 University of Oulu, Finland; [3]Division of Global Health, Department of Public Sciences, Karolinska Institutet, Stockholm, Sweden

E-mail: riku.turkki@helsinki.fi

*Diagnostic Pathology* 2013, **8(Suppl 1):**S30

**Background:** Computer aided analysis of virtual slide images has become an increasingly popular research topic and area of development. Novel applications are aimed for instance at automation of certain stages in sample assessment and assay readout (e.g. segmentation, detection, etc.) [2] or for reproducible measurement of a sample's visual appearance (e.g. grading, morphological classes etc.) [3].

Statistical learning methods are one of the key algorithms used in building these image analysis applications. Especially supervised algorithms (e.g. support vector machines (SVM), AdaBoost, k nearest neighbour (kNN), etc.) offer a way to train a classifier that can perform complex quantification tasks. The process requires annotation of labelled examples in order to describe the task and learn a model. Recently, image acquisition techniques for large-scale digitization of tissue samples have become common and require new methods to perform the annotation and associated labelling [4].

In this study we present an annotation tool that combines 1) direct interaction with a remote slide collection (a web based virtual microscopy application), 2) an interface to annotate points of interest in the virtual slides and 3) fast transition from annotation to development of image analysis methods.

**Methods:** The annotation tool is written in MATLAB (matrix laboratory) that is a cross-platform numerical computing environment (MathWorks, Natick, MA). MATLAB is a tool with a wide range of application areas and adopted by the image analysis community because of its versatile features. The programming language and its computing environment offers tools for testing and implementation of methods from a number manipulation to creating interactive graphical user interface (GUI) applications.

The annotation tool is based on exploiting two important properties of the current platform: 1) a *random code-stream access* featured image format (e.g. ECW or JPEG 2000) and 2) a *compression streamlining* protocol. The random code-stream compression makes it possible to extract sub-images from a large virtual slide file. Thus, the user can download only a field of interest, instead of loading the whole image file, which might be of gigapixel size. The other required property, the compression streamlining protocol, basically implements this sub-image extraction over standard hypertext transfer protocol (HTTP) in an efficient way.

With the advantages of the above-mentioned properties that a slide collection should have, it is reasonably straightforward to implement a MATLAB based tool to operate a remote database. The overall structure of the tool is illustrated in Figure 1. In a first phase the tool reads the metadata of a virtual slide file and deduces its height and width. Using a server's streamlining protocol, uniform resource locators (URL) are defined to extract the tiles from the pre-calculated coordinates within the virtual slide. In the annotation phase, the tiles are extracted and downloaded one at the time from the server and displayed to an annotator and requested for annotations. The locations and labels of the annotations are saved locally in MAT files.

**Results and discussion:** We implemented a MATLAB based annotation program capable of accessing a remote slide collection. The implementation is demonstrated within a previously described virtual microscopy environment [1,5] running image web server software (Erdas Inc., Atlanta Georgia), but is modifiable to pair with other platforms as well. The current virtual microscopy platform has proprietary server software (Image Web Server) that accepts both JPEG 2000 and ECW compressed files. The slides are made accessible via an *ImageX* protocol, which is implemented by the server. Other platforms, which have the ability to use JPEG 2000 files only, can utilize the JPIP (JPEG 2000 Interactive Protocol) to extract the metadata and tiles from a virtual slide over HTTP [6]. The code for the proposed tool is freely available [7].

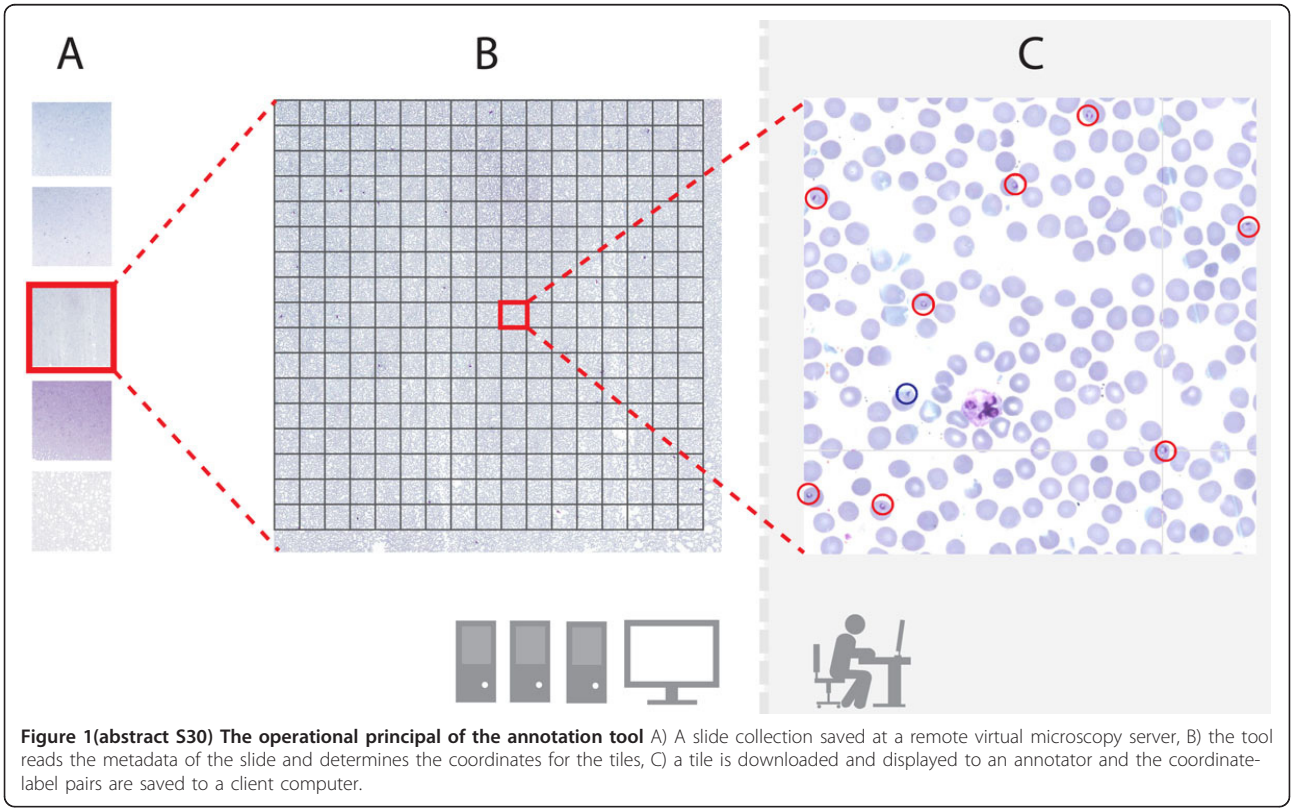



**Figure 1(abstract S30) The operational principal of the annotation tool** A) A slide collection saved at a remote virtual microscopy server, B) the tool reads the metadata of the slide and determines the coordinates for the tiles, C) a tile is downloaded and displayed to an annotator and the coordinate-label pairs are saved to a client computer.

The annotation tool keeps a record of all markings and enables the slides to be annotated in parts: it is possible to continue the annotation process from the beginning of a slide, or alternatively from a tile where the annotation was previously interrupted. The tool also keeps record of tiles that are already displayed to an annotator, which allows the areas that have not been annotated to be excluded from later processing if wanted. All the data are saved locally on the user's computer in standard MAT files.

The tiles are loaded one at the time from the image server and displayed to the annotator. The annotation window of an annotation process in a digitized thin blood smear film is illustrated in Figure 2. All the annotations are visualized with circles, which are coloured according to their labels. The tool can be set to handle at the most ten different labels.

The actual annotation process is kept as simple as possible to provide a fast and easy-to-use interface. The user is able to add and remove an annotation by simple mouse clicks: the left button for adding a new annotation and the right button to delete the nearest detected annotation. The keyboard's number keys are used to set the active category label and by keys N (next) and P (previous) the user can move between the tiles.

The work was motivated by our own need to annotate large virtual slide collections. In addition to the proposed applications, the tool can be applied to any kind of analysis of virtual slides. For instance, the tool offers a way to compare the agreement and accuracy of human annotators to recognize and locate objects, or it can be used to deliver one's annotations to a shared resource.

**Conclusions:** We have described a simple and effective way to combine a computational environment with a virtual microscopy environment [1] to allow fast transition and iteration between method development and data annotation. The code for the annotation tool is freely available [7] and can be modified to suit different server settings. The use of the tool is demonstrated on a virtual microscopy platform by providing example slides of thin blood films.

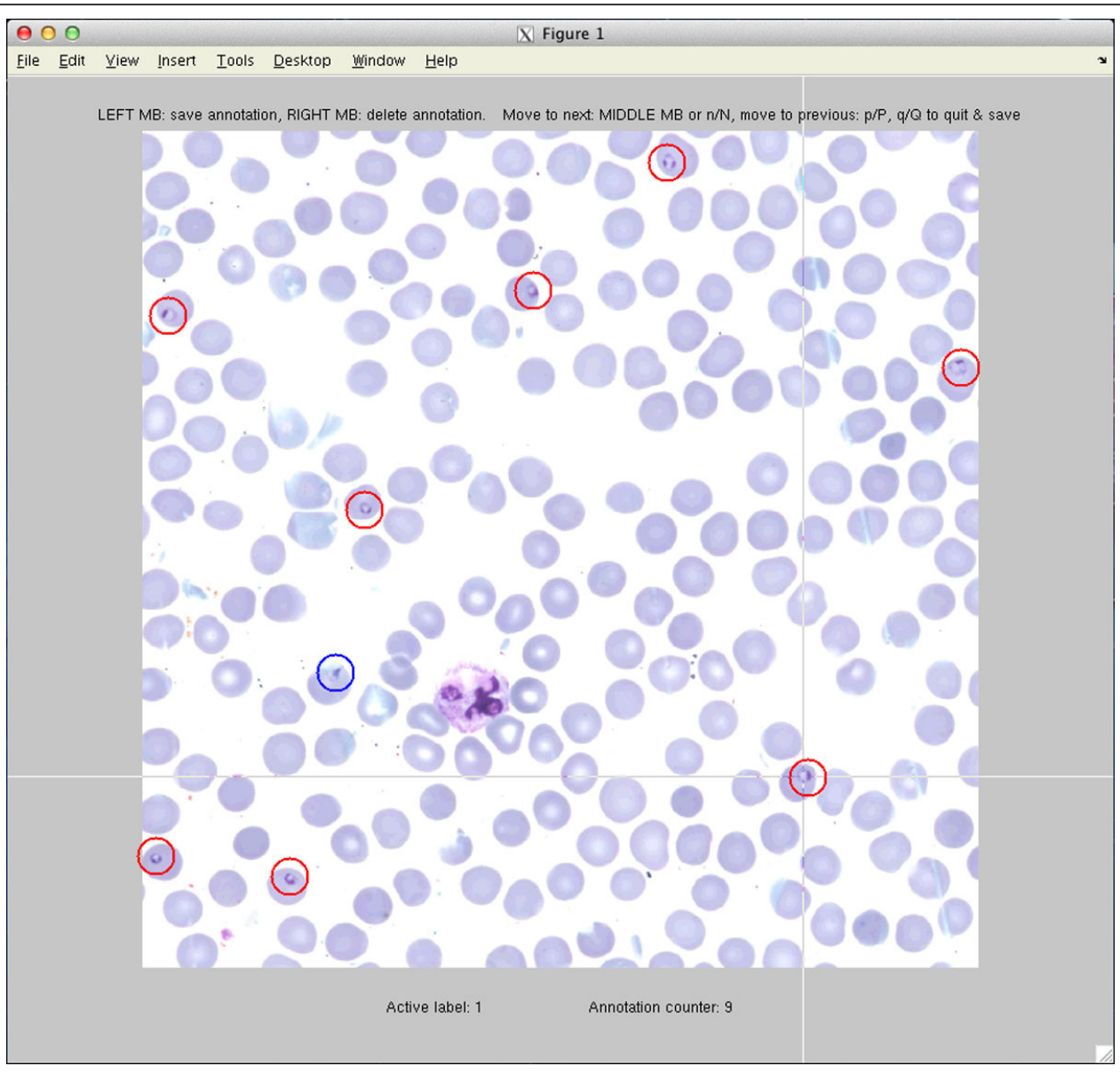



**Figure 2(abstract S30) The annotation window** Marked annotations are visualized with coloured circles to distinguish their labels. User can interact with the tool by using the mouse and the keyboard. The allowed user commands are displayed above the image.

**List of abbreviations:** ECW: Enhanced Compressed Wavelet; GUI: Graphical User Interface; HTTP: Hypertext Transfer Protocol; kNN: k-Nearest Neighbour; SVM: Support Vector Machine; URL: Uniform Resource Locator

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** RT planed and wrote the code for the annotation tool. MW and VO reviewed the implementations and suggested modifications. ML consulted on the use of the virtual microscopy platform. RT drafted the article and NL and JL contributed by supervision of the work and review of the manuscript. All authors read and approved the final manuscript.

**Acknowledgements:** The study was kindly supported by the national Biomedinfra and Biocenter Finland projects.

**References**

1. The WebMicroscope virtual microscopy environment. [http://www.webmicroscope.net/].
2. Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Ahonen T, Pietikainen M, Lundin J: Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. *Diagn Pathol* 2012, **7**:22.
3. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC: Automatic breast cancer grading of histopathological images. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* 2008, 3052-3055.
4. Kayser K: Introduction of virtual microscopy in routine surgical pathology-a hypothesis and personal view from Europe. *Diagn Pathol* 2012, **7**:48.
5. Szymas J, Lundin M: Five years of experience teaching pathology to dental students using the WebMicroscope. *Diagn Pathol* 2011, **6(Suppl 1)**:S13.
6. The description of JPIP-protocol. [http://www.jpeg.org/jpeg2000/j2kpart9.html].
7. The code for the proposed annotation tool. [http://fimm.webmicroscope.net/Research/Momic/tp2012].

## S31

### An oriented service multilayer architecture for virtual microscopy in mobile devices

Luis Martínez[1], Germán Corredor[1], Marcela Iregui[2], Eduardo Romero[3*]
[1]Department of System and Industrial Engineering, Faculty of Engineering, Universidad Nacional de Colombia, Bogotá, Colombia; [2]Multimedia Engineering Program, Faculty of Engineering, Universidad Militar Nueva Granada, Bogotá, Colombia; [3]Departament of Diagnostic Images, Faculty of Medicine, Universidad Nacional de Colombia, Bogotá, Colombia
E-mail: edromero@unal.edu.co

*Diagnostic Pathology* 2013, **8(Suppl 1):**S31

**Background:** The Virtual Slide (VS) is the constructed tool for interaction with a large amount of visual information, using for doing so devices designed to display and interact with the VS, i.e., search of Regions of Interest (RoIs), labeling specific RoIs in the VS, automatic VS or retrieval of certain RoIs [1]. Two main VS advantages, over a glass slide, are the information access and the data maintenance. Disadvantages are related with the computational cost [2]. Overall, slide storing and interaction is carried out from the same device used for display. Mobile devices are of course an extreme case of poor resources [3] and therefore clever navigation strategies are necessary to optimally interact with the VS.

Related to interaction with VS from mobile devices, there are two main issues to be addressed. The former is related to the storage and access to a large quantity of data, the latter concerns the reconstruction and display of the visual information. Several works have used the JPEG2000 compression standard to address the storage needs [4][5,6]. JPEG2000 is an image compression standard designed by the Joint Photographic Expert Group, based on the Discrete Wavelet Transform and the EBCOT encoder [7]. This standard provides several advantages, among others, compression efficiency, lossy and lossless compression and multidimensional data access, i.e., random and multiple resolution data representation and data organization in several quality layers [8]. Likewise, the granularity provided by the standard allows the retrieval of individual packets, guaranteeing transmission of just the number of bytes required to reconstruct particular regions of an image, instead of the whole slide [9].

As it was mentioned, a reconstruction task can be achieved by taking advantage of the granularity in the JPEG2000 standard. However, the data syntax described in the standard rules out the interactive construction of a valid data stream from arbitrarily ordered packets [10]. Moreover, reconstruction and display of the VS is still an open problem because of the high consumption of computational resource when decompressing the bitstream. A well designed architecture must therefore address the reconstruction task under the perspective of an optimal adaptation of the process policies to the problem.

**Material and methods: Experimental setup:** The proposed architecture was evaluated with a virtual slide of 36000x9200 pixels, each pixel corresponding to 0.67 $\mu m^2$. The original Virtual Slide had a size of 995 MB, and after JPEG2000 compression, of 226 MB. To run the storage layer, a distributed file system was deployed using 5 Linux nodes. The machines that form the network have limited processing capacity and low speed hard disks (1 GB of memory, processor of 2.2 GHz and disk of 7200 rpm). To run the data access and the proxy layers, two servers were selected; each with operating system OpenSUSE 11.4, 2.8 GHz 4-core processor and 5 GB RAM. Likewise, to run the Samsung Galaxy Tab 10.1 client, under an operating Android system 3.2, it was selected a device with a 1280×800 display size, 1 GB RAM memory and 1 GHz dual-core Nvidia Tegra 2 processor.

**Architecture overview:** The proposed architecture exploits the JPEG2000 granularity by dividing the main tasks of the data processing into four layers. The architecture and their main components are shown in Figure 1.

The data storage layer is charged of managing the compressed images. A JPEG2000 compressed image typically contains numerous embedded subsets, each standing for any of a large number of different spatial resolutions, image quality layers and spatial regions. This multidimensional access to data is defined in the JPEG2000 standard as spatially adjacent code-blocks, known as *precincts*. Each precinct is represented as a collection of *packets*, with one packet per quality layer, resolution level and component [10]. The logical structure of this compressed file is stored in an index file, along with the compressed image, and is herein used to navigate through the compressed file. The second, the data access layer, provides the required services to interact with the information stored in the previous layer. A loosely coupled architecture is maintained by providing the required services to interact with the minimal unit of information, i.e. services to retrieve specific image packets and services to retrieve the compressed image header. The services provided by this layer handle each request independently, thereby guaranteeing a simultaneous information access.

The proxy layer is the backbone of the interaction with the data stored in the first layer. This layer is responsible for two important tasks. The first task is to facilitate interaction with data retrieved by the data access layer, and the second one consists in providing efficient access to the previously requested packets. Provided that the present architecture is service oriented, this third layer receives and sends messages, from which the raw data must be extracted. For doing so, this layer has a decompression module, containing the functions to manage and to map the incoming messages from the data access layer. In addition, this layer implements a simple cache module, charged of checking and/or requesting the required packets to fulfill a requested region. Also, a communication API was designed for mobile devices because of the communication problems presented when using conventional web services since they are difficult to process in such limited devices.

Finally, the client layer is a standalone prototype, whose main function is to map the requested regions to list of packets, and reconstruct Virtual Slides, using information provided by the previous layers. The client uses the communication API and retrieves the required information for reconstruction. The client layer uses also a transformation module that allows the final display of inverse transformed wavelet coefficients.

**Results and discussion:** The proposed architecture was twice tested. Firstly, it was requested a variable region size, with constant resolution and quality values. Secondly, the requested regions were refined by requesting higher quality layers. Results show the advantage of retrieving and decompressing (decoding) a particular image region instead of the whole slide. In the first test, the time for resolving a requested region is proportional to the number of required packets to reconstruct it. These results are presented in the Table 1.

In the second test, the time between the requested layers is relatively small, probably because most of the relevant information is mainly compressed in the first layers, leaving small refinement details for the last ones. These results are presented in the Table 2.

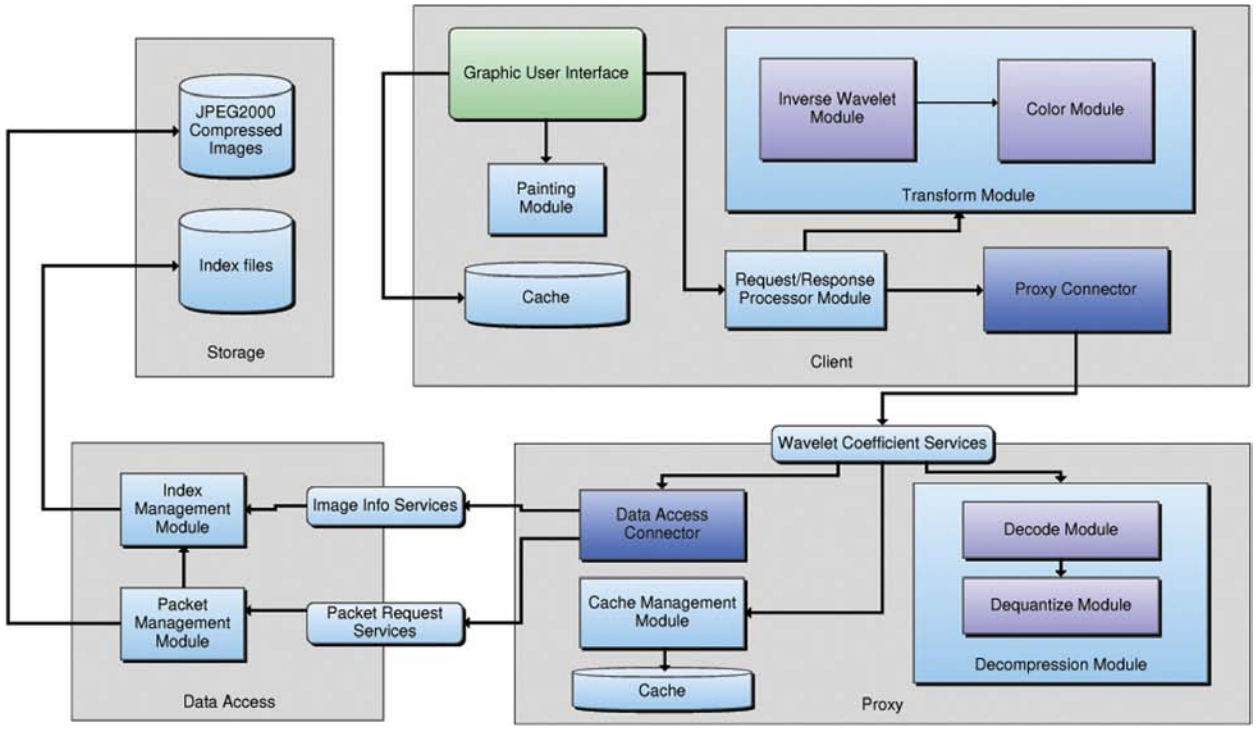



**Figure 1(abstract S31) Multilayer architecture distribution** The figure shows the distribution of the main components within different layers in the proposed architecture.

**Table 1(abstract S31) Requests with different region size**

| Window size | Transmission time (ms) | Reconstruction time (ms) |
|---|---|---|
| 1024x1024 | 34.9±5.65 | 162.6±4.7 |
| 2048x2048 | 79.6±6.64 | 244.1±1.1 |
| 4096x4096 | 471±68.16 | 613.7±36.79 |
| 8192x8192 | 1421±180.38 | 1881.8±46.83 |

The table shows the evolution of the transmission and reconstruction times while the client requests different region sizes of a virtual slide.

**Conclusions:** In this article, it was presented a distributed multi-layer architecture that supports interaction between its layers through a service-oriented scheme. It was shown that retrieval and reconstruction times are relatively slow using a refinement process by quality layers.
**List of abbreviations:** API: Application Programming Interface; ROI: Region of Interest; VM: Virtual Microscope; VS: Virtual Slide

**Table 2(abstract S31) Requests with refinement process**

| Quality layer | Transmission time (ms) | Reconstruction time (ms) |
|---|---|---|
| 1 | 83.1±25.39 | 245.9±5.49 |
| 2 | 94.4±25.55 | 249.7±1.16 |
| 3 | 97.3±25.66 | 250.3±1.57 |
| 4 | 99.5±24.13 | 250.5±0.53 |
| 5 | 103.4±25.07 | 254.9±6.66 |
| 6 | 99.8±7.22 | 258.2±8.57 |
| 7 | 108.9±24.85 | 268.9±27.09 |
| 8 | 124.5±51.03 | 265.5±28.48 |

The table shows the evolution of the transmission and reconstruction times while the client requests different layers for the same region of a virtual slide.

**Competing interests:** The authors declare that they have no competing interests.
**Authors' contributions:** LM implemented, deployed, and evaluated the web services for each layer and the results obtained in the reconstruction process. GC implemented and tested the client prototype, the communication API and the codestream transmission for Android Operating System. MI proposed, implemented and tested the index design, construction and parsing, and developed the fundamental ideas underlying this architecture. ER conceived the study, developed the fundamental ideas underlying this architecture, participated in the experimental design and was the director of the whole project. All authors read and approved the final manuscript.

**References**

1. Treanor D: **Virtual slides: an introduction.** *Diagnostic Histopathology* 2009, **15(2)**:99-103.
2. Glatz-Krieger K, Glatz D, Mihatsch M: **Virtual slides: high-quality demand, physical limitations, and affordability.** *Human Pathology* 2003, **34(10)**:968-974.
3. Agu E, Banerjee K, Nilekar S, Rekutin O, Kramer D: **A middleware architecture for mobile 3D graphics.** *25th IEEE International Conference. Distributed Computing Systems Workshops* 2005, 617-623.
4. Taubman D, Rosenbaum R: **Rate-distortion optimized interactive browsing of JPEG2000 images.** *Proceedings. International Conference on Image Processing* 2003, **3**:III- 765-768.
5. Descampe A, De Vleeschouwer C, Iregui M, Macq B, Marqués F: **Prefetching and caching strategies for remote and interactive browsing of JPEG2000 images.** *IEEE Trans Image Process* 2007, **16(5)**:1339-1354.
6. Iregui M, Gómez F, Romero E: **Strategies for efficient virtual microscopy in pathological samples using JPEG2000.** *Micron* 2007, **38(7)**:700-713.
7. ISO/IEC 15444-1, 2000: **Information technology JPEG2000 image coding system – Part 1: Core coding system.** *International Organization for Standardization* 2000.
8. Rabbani M, Joshi R: **An overview of the JPEG 2000 still image compression standard, signal processing.** *Image Communication* 2002, **17(1)**:3-48.

9. Iregui M, Meessen J, Chevalier P, Macq B: **Flexible access to JPEG2000 codestreams.** *23rd Symposium on Information Theory in the Benelux* 2002.
10. Taubman D: **Remote browsing of JPEG2000 images.** *Proceedings. International Conference on Image Processing* 2002, **1**:I-229-I-232.

## S32

### Label free technologies: Raman micro-spectroscopy and multi-spectral imaging for lymphocyte classification

Teddy Happillon[1†], Valérie Untereiner[1†], Abdelilah Beljebbar[1†], Cyril Gobinet[1†], Michel Manfait[1†], Sylvie Daliphard[2†], Pascale Cornillet-Lefebvre[2†], Xavier Troussard[3†], Jesus Angulo[4†], Santiago Velasco-Forero[4†], Véronique Saada[5†], Georges Flandrin[6†], Jacques Klossa[6*]
[1]MEDyC FRE/CNRS 3481, 51096 Reims, France; [2]CHU de Reims, 51100 Reims, France; [3]CHU de Caen, 14033 Caen, France; [4]CMM-ARMINES, 77300 Fontainebleau, France; [5]IGR-Institut Gustave-Roussy, 94805 Villejuif, France; [6]TRIBVN, 39, rue Louveau, 92320 Châtillon, France
E-mail: jklossa@tribvn.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S32

**Background:** Current diagnostic and prognostic approaches in oncology use morphological and molecular techniques which lead to patient personalized therapies. However, they are still complex and hard to standardize. This is also true for Chronic Lymphocytic Leukemia (CLL) that has been chosen for the IHMO project [1,2]. Simplifying diagnostic processes and making easier the standardization would be highly suitable. In order to develop such a simpler automated method, the IHMO project, funded by the French National Research Agency, proposed to develop a multimodal microscopy scanning platform that includes in a single machine a Raman micro-spectrometer (RMS) combined with a multispectral imager (MSI) [3,4].
RMS is a quick non-invasive and non-destructive technique for tissues and cells analysis [5]. It is very sensible to molecular changes and it could be used as a powerful diagnostic and prognostic tool when used in association with multivariate statistical methods. It is particularly useful for characterizing pathological tumors especially at the cellular level [6,7]. Multispectral imaging in the visible spectrum could confirm RMS classifications and provide new morphometric findings [8].
CLL disease is characterized by the proliferation of lymphocytes (lymphocytosis). It is the most common leukemia, preferentially affecting people aged over 50 years old. It is incurable and in most cases shows no clinical signs. Thus, it is often discovered by chance during a blood test. If necessary, morphological and immunological studies are led by analyzing blood smears colored with May-Grünwald Giemsa, by making a complete blood count and by computing a Matutes score. Such studies are necessary because it is impossible to distinguish a healthy cell from a cancerous one only using a conventional microscope.
**Material and methods:** Blood smears were prepared on classical glass slides commonly used in laboratories for microscopy; cells were localized by optical microscopy. A new multimodal machine which has been developed combining i) a 10 bands multi-spectral imager using the full spectrum of transmitted visible light ii) a Raman micro-spectrometer, equipped with a 532nm diode laser excitation source delivering 13.5mW of power on the sample; iii) a microscope stand with 40x and 150x lenses suited with an xyz motorized stage for scanning the blood smear, and localizing x-y coordinates of a representative series (~100 for each patient) of lymphocyte cells before registering Raman spectra on their nuclei and their individual multi-spectral images. An additional piezo actuator allowed for precise z stack recording. Figure 1 shows each step from the screening of the smears to the final results.
**Raman micro-spectroscopy:** More precisely, 997 polymorphonuclears, 95 monocytes and 1127 lymphocytes from 12 different blood samples have been considered in the first part of this study, and a Raman spectrum has been acquired on each of them. In the second part, 4257 spectra have been registered on lymphocytes of 49 leukemic patients suffering from hyper leukocytosis Chronic Lymphocytic Leukemia, and 2596 spectra have been recorded on lymphocytes of 27 healthy individuals. These spectra have been then studied, using around 90 cells per blood sample.
Raman data were first pretreated to erase contaminant information into the spectra and then were analyzed using a multivariate statistical method which put forward the relevant information needed to distinguish in a first time lymphocytes from polymorphonuclears. The spectra are then reduced to their relevant information and classified using a Support Vector Machine algorithm [9]. Then this algorithm was used to develop a classification model splitting leukemic and healthy lymphocytes; 3 sets of data were created, the first one, the training set, composed of the spectra from 6 leukemic patients (513 spectra) and 4 healthy individuals (315 spectra) which was used to create different prediction models, the second set, the validation set, composed of spectra from 33 leukemic patients (2820 spectra) and 13 healthy individuals (1106 spectra) was used to select the best model among all previously computed. Finally, the third set, the test set, composed of spectra from 20 different blood samples (2099 spectra) was used to validate the selected classification model in a blind way.
**Multi-spectral imaging:** After acquisition of Raman spectra, the slides are stained using a RAL_DIAGNOSTICS™ standardized staining protocol. Then, for each target cell, a multi-z/multi-spectral image is acquired: a z-stack of 100 nm spaced of 24 monochrome images for each of the ten wavelength bands. An algorithm combining mathematical morphology techniques and sparse regression was developed to produce a color RGB image with an extended-depth-of-focus from the multi-z/multispectral image. This color image is used to visually confirm the classification of cells characterized by RMS. The set of scalar images multi-Z/multi-spectral of each cell is also processed by mathematical morphology techniques to produce morpho-spectral texture descriptors reducing the 240 images from each color band to a set of 4 representative images. Such data reduction process allowed representing each cell by a 40x40 square symmetrical correlation matrix. Then, using tools from information geometry and multivariate statistics, cells are embedded into a dimensionality reduced space used to produce an unsupervised classification into two classes: leukemia patients vs. healthy individuals.
**Results and discussion:** First stage molecular classification with Raman spectrometry aimed at identifying the nucleated components from a blood smear, mainly polymorphonuclears and lymphocytes. Indeed such differentiation is straightforward with morphological methods like MSI even on unstained samples. Molecular classification gave a sensitivity of 99.3% and a specificity of 99.6%.
The second stage aimed at classifying spectra from leukemic and healthy lymphocytes and provided a sensitivity of 80% and a specificity of 100% on an extended set of blind samples.
It has to be mentioned that in a first step of this work, we tried to match the classification results to the percentages obtained with flow cytometry in term of T, B and NK lymphocytes of a blood sample. However, these results could not be compared with each other since in this study the information which is taken into account is obtained from the nucleus of each white blood cell. As a result, this is a molecular and nuclear information while, in classical cytology, this is the membrane of the cells which is considered.
Morpho-spectral texture classification from multispectral imaging in the visible is used to complement and consolidate the RMS data classifications and in on-going work will be used to identify specific morpho-spectral characteristics associated to leukemic cells (see figure 2).
**Conclusion:** IHMO project has demonstrated the power of Raman micro-spectroscopy coupled with supervised classification algorithm such as Support Vector Machines for cell classification and more precisely for the diagnosis of CLL. Morphological descriptors obtained from multi-Z and multispectral images provide another independent classification that still needs to be assessed.
The multimodal microscopy platform can be used more generally in the field of cytohematology, however application to cytological and histological pathology would need further developments and could take profit from new methods in data classification.
**List of abbreviations:** RMS: Raman Micro-Spectroscopy; MSI: Multi-Spectral Imaging; CLL: Chronic Lymphoid Leukemia; SVM: Support Vector Machines
**Competing interests:** The authors declare that they have no competing interests.
**Authors' contributions:** • TH carried out the acquisitions, pretreatments, the highlighting of relevant information into the spectra, the classification of the data and drafted the extended abstract.
• AB drove the implementation of pretreatments adapted to spectral data and found a way to highlight the relevant information into the spectra.
• VU realized the acquisitions of the data and drafted the extended abstract.

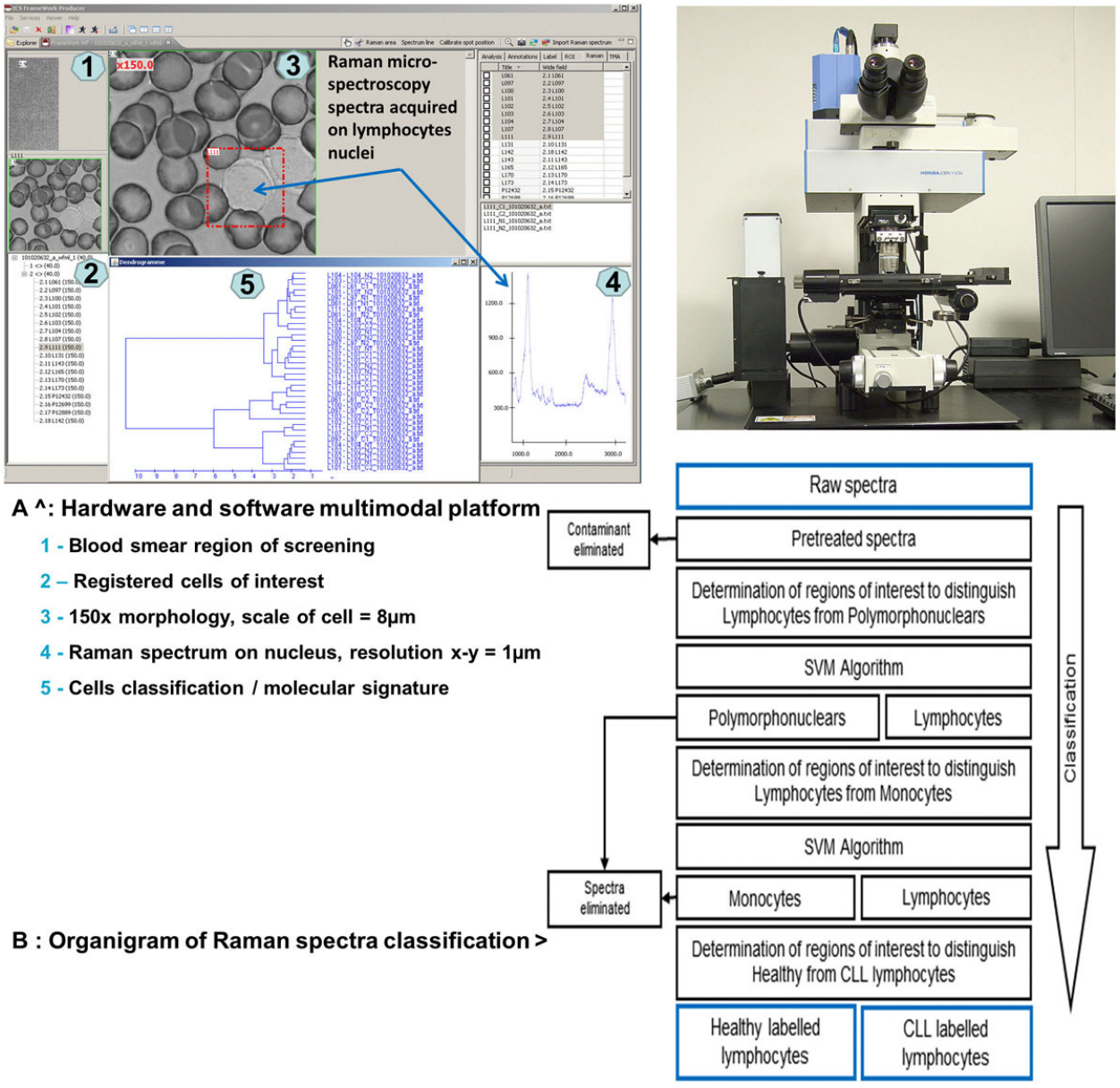



**Figure 1(abstract S32)** IHMO platform includes the Raman micro-spectrometer and the multi-spectral sources (hardware) and a software able to control the acquisitions and the classification process.

• CG contributes to the choice of the classification algorithm among several studied (supervised or not), and participated in the choice of each possible settings of these algorithms to improve the obtained results.
• JA and SVF developed the multi-spectral classification method and produced the multi-spectral classification.
• XT, JK and MM both designed and managed the full project with the contribution of GF.
• XT, SD, VS and PCL selected the best CLL patients, provided the corresponding samples and patient's data, and realized the flow cytometry for each of them.

**Acknowledgements:** • The French National Research Agency (ANR) for funding
• Chih-Chung Chang and Chih-Jen Lin, LIBSVM: a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm
• Clinical staff of the university hospital of Reims
• Rodolphe Dagiral and Valérie Coomans from Réactifs RAL for developing and providing standardizes MGG stains used for multi-spectral imaging classification
• Philippe Rideau, David Pointu and Claire Vernois from NIKON France for defining and installing appropriate microscopy hardware
• Emmanuel Froigneux from HORIBA JOBIN YVON for developing and providing the micro-spectroscopy Raman module

**References**

1. Cramer P, Hallek M: **Prognostic factors in chronic lymphocytic leukemia what do we need to know?** *Nat Rev Clin Oncol* 2011, **8(1)**:38-47.
2. Zenz T, Mertens D, Küppers R, Döhner H, Stilgenbauer S: **From pathogenesis to treatment of chronic lymphocytic leukaemia.** *Nat Rev Cancer* 2010, **10(1)**:37-50.
3. Velasco-Forero S, Angulo J: **Parameters selection of morphological scale space decomposition for hyperspectral images using tensor modeling.** *Proc. of SPIE symposium on Defense, Security, and Sensing: Algorithms and*

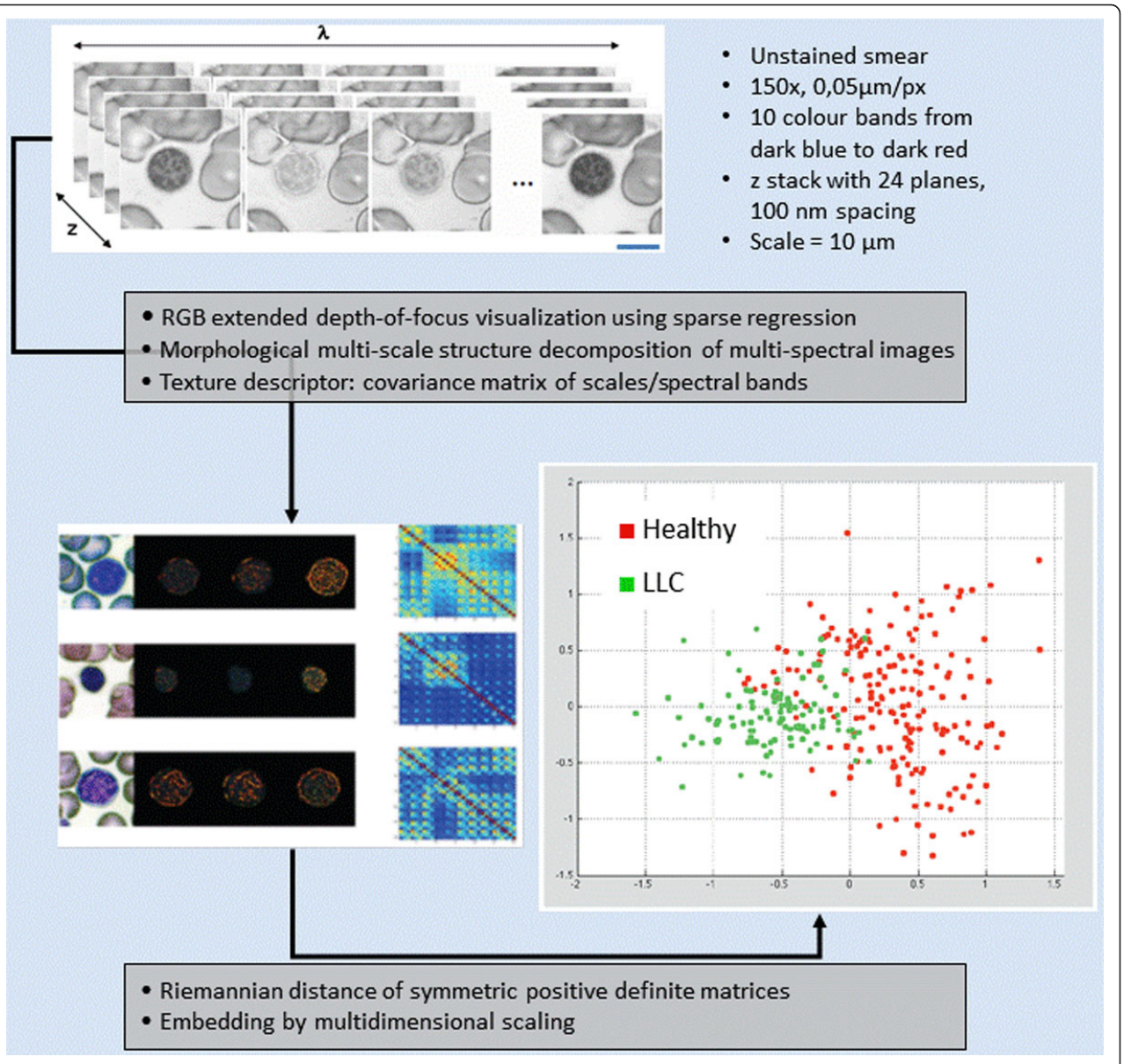



**Figure 2(abstract S32)** multi-spectral imaging feature extraction and classification method.

*Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVI, SPIE* 2010, **7695**:12.

4. Malik Z, Rothmann C, Cycowitz T, Cycowitz ZJ, Cohen AM: **Spectral morphometric characterization of B-CLL cells versus normal small lymphocytes.** *J Histochem Cytochem* 1998, **46(10)**:1113-8.
5. Anthony T: **Raman spectroscopy in biology.** NewYork: JohnWiley & Sons 1982.
6. Draux F, Jeannesson P, Beljebbar A, Tfayli A, Fourre N, Manfait M, Sulé-Suso J, Sockalingum GD: **Raman spectral imaging of single living cancer cells: a preliminary study.** *Analyst* 2009, **134**:542-8.
7. Krishna CM, Sockalingum GD, Vadhiraja BM, Maheedhar K, Rao AC, Rao L, Venteo L, Pluot M, Fernandes DJ, Vidyasagar MS, Kartha VB, Manfait M: **Vibrational spectroscopy studies of formalin-fixed cervix tissues.** *Biopolymers* 2007, **85**:214-21.
8. Angulo J, Klossa J, Flandrin G: **Ontology-based lymphocyte population description using mathematical morphology on colour blood images.** *Cell Mol Biol* 2006, **52(6)**:2-15.
9. Cortes C, Vapnik V: **Support-Vector Networks.** *Machine Learning* 1995, **20**:273-297.

## S33

### Highlighting peritumoral areas in human skin cancer biopsies by infrared micro-spectroscopy

David Sebiskveradze[1], Cyril Gobinet[1*], Valeriu Vrabie[2], Pierre Jeannesson[1], Olivier Piot[1], Michel Manfait[1]
[1]MéDIAN, CNRS FRE 3481 MEDyC, SFR Cap-Santé, Université de Reims Champagne-Ardenne, 51 rue Cognacq-Jay, 51096 Reims, France; [2]CReSTIC, Université de Reims Champagne-Ardenne, Chaussée du Port, 51000 Châlons-en-Champagne, France
E-mail: cyril.gobinet@univ-reims.fr
*Diagnostic Pathology* 2013, **8(Suppl 1):**S33

**Background:** Fourier transform mid-infrared (FT–IR) microspectroscopy is a label-free optical method based on the interaction between an incident light beam and matter. This vibrational spectroscopy permits to probe the biochemical composition of the analyzed sample and thus gives information about the structure of this sample. Associated with an

imaging system, FT–IR microspectroscopy of human tissues can be used as a very sensitive, non-destructive and non-subjective tool for the detection and localization of tumoral areas independently of visual morphology. Thus, FT–IR microimaging has demonstrated potential to provide clinically relevant diagnostic information in oncology [1-3].

The biochemical changes related to carcinogenesis between cancerous and surrounding tissue areas are subtle. As a consequence, IR hyperspectral images need to be processed by powerful digital signal processing and pattern recognition methods in order to highlight these changes [4,5]. To this end, an innovative fuzzy C-means (FCM) clustering-based algorithm was proposed in [6]. The real advantage of FCM is that it introduces the notion of nuance into the clustering of IR image pixels. Consequently, FCM allows considering the progressive transition between non-cancerous tissues and cancer lesions and reveals every nuance of intratumoral heterogeneity [6]. Moreover, the FCM-based algorithm proposed in [6] is fully automatic, i.e. the optimal clustering parameters such as the number of clusters are automatically determined. The main drawbacks with this algorithm are that it is very time consuming and that the transition areas can be difficultly seen. In this work, we thus propose solutions to these problems.

**Material and methods:** IR spectral images were acquired on 8 biopsies of formalin-fixed paraffin-embedded human skin carcinomas, squamous cell carcinomas (SCC, n=3) and basal cell carcinomas (BCC, n=5). The samples were selected by the pathologists from the tumor bank of the Pathology Department of the University Hospital of Reims (France).

From samples, 10-micron thick slices were cut and mounted on a calcium fluoride ($CaF_2$) (Crystran, Dorset, UK) window for FT–IR imaging without any particular preparation, especially no chemical dewaxing. First adjacent slices (5-µm thick) to those used for FT–IR analysis were stained with hematoxylin and eosin (H&E) for conventional histology. From these slices, the cancer outlines defined by the pathologists were drawn on the photomicrographs.

FT–IR hyperspectral images were recorded with a Spectrum Spotlight 300 FT–IR imaging system coupled to a Spectrum one FT–IR spectrometer (Perkin Elmer Life Sciences, France), with a spatial resolution of 6.25 µm and a spectral resolution of 4 $cm^{-1}$. Each spectral image, covering a substantial part of the biopsy, consisted of about 30 000 spectra.

The samples being analyzed without previous chemical dewaxing, the recorded FT–IR hyperspectral image were digitally corrected from paraffin spectral contribution thanks to an automated pre-processing method based on extended multiplicative signal correction (EMSC) [7]. Only the spectral variability of the molecular composition of the tissue is thus retained in the data sets.

In order to highlight the different biological structures of the samples from the weak inter-spectra differences, the EMSC-based pre-processed IR spectra were analysed by an upgraded version of the FCM clustering-based algorithm proposed in [6]. The innovations mainly consist in the breaking of the FCM algorithm as soon as the estimated clusters present some uninteresting characteristics and in the limitation of the number of computed FCM. These computational aspects will be described later in details in another article. Furthermore, in order to highlight the transition areas between different tissue structures, an entropy based interconnectivity measure between clusters has been defined and applied on the FCM results. For cluster assignment, each color-coded map was then provided to the pathologists for a comparison with the corresponding H&E-stained sections.

**Results and discussion:** After the application of the upgraded version of the FCM-based algorithm, the reconstructed color-coded clustering images allow recovering different histological structures automatically, particularly to precisely localize tumoral areas and their normal counterparts. Due to the limited length of the article, here we present only one representative case of SCC.

The images generated by the FCM-based algorithm are shown in Figure 1. After comparison with the histological image, each generated cluster can be assigned to a precise tissue structure: tumoral area (cluster 1), invasive front (cluster 2), dermis (clusters 3, 4 and 5) and epidermis (cluster 6). These results are identical to those presented in [6], except that the computational time is divided by 8 thanks to the innovations included into the original algorithm.

More than reproducing classical histology, our algorithm can give access to additional information about the assignment of the IR image pixels to the tissular structures. For each pixel, fuzzy clustering provides membership values, permitting to nuance their assignment. Such data are very valuable for the pixels located at the interface between tumoral tissue and its microenvironment. However, to ease the interpretation of transitional areas between tumor and marginal normal tissue, we developed an entropy based interconnectivity measure which is maximal when a pixel equally belongs to two clusters. Applied on the SCC IR image as shown on Figure 2, this interconnectivity measure shows that the invasive front (cluster 2) is tightly connected to the tumoral area (cluster 1) and that a surprising clear-cut difference between the invasive front (cluster 2) and the surrounding dermis (clusters 3, 4 and 5) exists. These areas cannot be identified on hematoxylin-eosin staining or by conventional clustering of IR data.

**Conclusions:** IR spectral microimaging associated with clustering techniques shows a great potential for the direct analysis of paraffin-embedded tissue sections of human skin cancers. These preliminary results show significant potential for probing tumor progression and for early determination of tumor aggressiveness in cutaneous cancers. Experiments are underway to define the molecular assignments of the spectral variations observed in these peritumoral areas. Furthermore, this approach could be applied to other human skin cancers such as melanoma.

**List of abbreviations:** FT–IR: Fourier transform mid-infrared; FCM: fuzzy C-means; SCC: squamous cell carcinoma; BCC: basal cell carcinoma; H&E: hematoxylin and eosin; EMSC: extended multiplicative signal correction

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** • DS realized the acquisitions of the data, developed the original FCM based algorithm, applied it on the IR images and drafted the extended abstract.

• CG developed the upgraded FCM based algorithm and the entropy based interconnectivity measure, and drafted the extended abstract.

• VV developed the original FCM based algorithm.

• PJ, OP and MM designed and managed the project

**Acknowledgements:** This study was supported by a grant of Institut National du Cancer (INCa), Canceropôle Grand Est. Ligue contre le Cancer, Comité de l'Aisne, INSERM PNR Imagerie and CNRS Projets Exploratoires Pluridisciplinaires are thanked for their financial support.

**References**

1. Wolthuis R, Travo A, Nicolet C, Neuville A, Gaub M-P, Guenot D, Ly E, Manfait M, Jeannesson P, Piot O: **IR spectral imaging for histopathological characterization of xenografted human colon carcinomas.** *Anal Chem* 2008, **80**:8461-8469.
2. Ly E, Piot O, Durlach A, Bernard P, Manfait M: **Differential diagnosis of cutaneous carcinomas by infrared spectral micro-imaging combined with pattern recognition.** *Analyst* 2009, **134**:1208-1214.
3. Travo A, Piot O, Wolthuis R, Gobinet C, Manfait M, Bara J, Forgue-Lafitte M-E, Jeannesson P: **IR spectral imaging of secreted mucus: a promising new tool for the histopathological recognition of human colonic adenocarcinomas.** *Histopathology* 2010, **56**:921-931.
4. Beljebbar A, Dukic S, Amharref N, Bellefqih S, Manfait M: **Monitoring of biochemical changes through the C6 gliomas progression and invasion by Fourier transform infrared (FTIR) imaging.** *Anal Chem* 2009, **81**:9247-9256.
5. Untereiner V, Piot O, Diebold MD, Bouché O, Scaglia E, Manfait M: **Optical diagnosis of peritoneal metastases by infrared microscopic imaging.** *Anal Bioanal Chem* 2009, **393**:1619-1627.
6. Sebiskveradze D, Vrabie V, Gobinet C, Durlach A, Bernard P, Ly E, Manfait M, Jeannesson P, Piot O: **Automation of an algorithm based on fuzzy clustering for analyzing tumoral heterogeneity in human skin carcinoma tissue sections.** *Lab Invest* 2011, **91**:799-811.
7. Ly E, Piot O, Wolthuis R, Durlach A, Bernard P, Manfait M: **Combination of FTIR spectral imaging and chemometrics for tumour detection from paraffin-embedded biopsies.** *Analyst* 2008, **133**:197-205.

## S34

### Label free technologies 3: infrared imaging applied to paraffinized tissue microarrays for colon cancer diagnosis

Ganesh D Sockalingum[1*], Jayakrupakar Nallala[1], Marie-Danièle Diebold[2], Cyril Gobinet[1], Olivier Piot[1], Valérie Untereiner[1], Michel Manfait[1]
[1]MEDyC, FRE CNRS/URCA 3481, France; [2]CHU Reims, France
E-mail: ganesh.sockalingum@univ-reims.fr
*Diagnostic Pathology* 2013, **8(Suppl 1):**S34

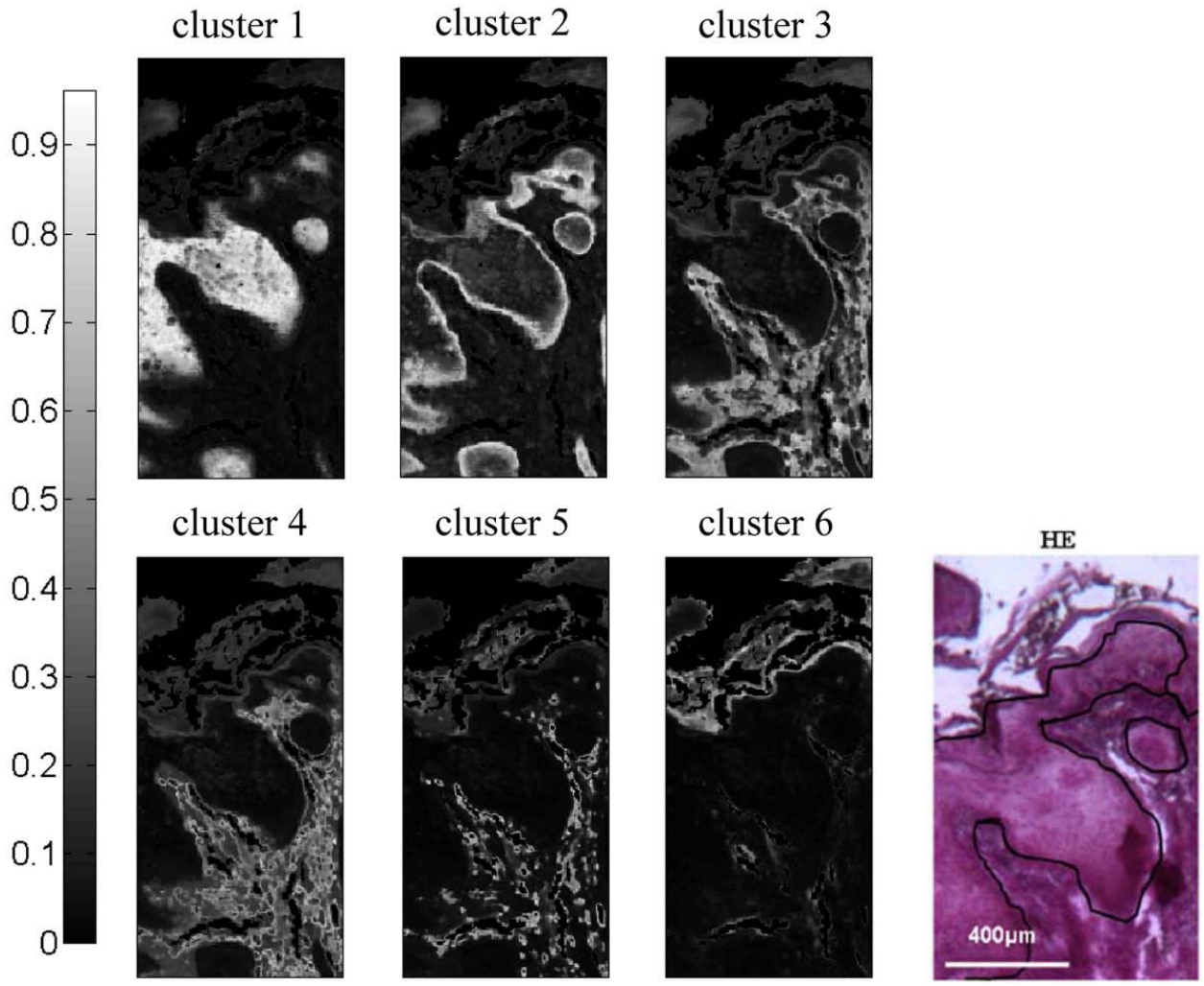



**Figure 1(abstract S33) Upgraded Fuzzy C-means (FCM) images on the Fourier transform mid-infrared (FT–IR) data set of the human skin squamous cell carcinoma (SCC) sample and the corresponding hematoxylin and eosin (H&E)-stained section.** Assignment of the clusters: 1 (tumor); 2 (invasive front); 3, 4 and 5 (dermis); 6 (epidermis). The color bar represents the membership value for each pixel. In the corresponding H&E-stained section, SCC is outlined.

**Introduction:** Colorectal cancers are the third most common type of cancers globally, affecting both sexes [1]. As of now, histopathology is the gold standard method for colon cancer diagnosis. Newer technologies are the important need of the hour to complement the existing approaches, for better understanding the onset and progression of the disease. Infrared (IR) imaging could be a potential candidate method because of its capability to probe non-destructively and in a label-free manner the intrinsic chemical bonds present in the tissue, thus giving a "spectral fingerprint" of its composition and structures. To this end, we have developed IR spectral histopathology with the aims to: (a) examine, the molecular changes between normal and tumoral colon tissues, (b) exploit its potentials to identify new diagnostic markers to complement conventional histopathology, and (c) develop an algorithm as an automatic diagnostic tool for tumor prediction, directly on paraffinized samples, without chemical de-waxing, staining or any further preparation.
**Methods:** Nine normal and 25 tumoral colon tissue sections (3mm diameter x 10 µm thick) embedded in the form of paraffinized tissue microarray, stabilized in an agarose matrix were analyzed directly by IR imaging. To avoid chemical deparaffinization, a modified Extended Multiplicative Signal Correction (EMSC) method [2] was used to digitally neutralize the spectral interferences of paraffin and agarose, and to preserve only the biological information from the tissue. Corrected spectra were classified using k-means to construct color-coded images using the HPS stained sections as reference. Linear Discriminant Analysis (LDA) was then used to construct a prediction model for identification of blind samples.
**Results:** EMSC permitted mathematical correction of the spectral interferences originating from paraffin and agarose. K-means classification allowed to identify and to distinguish important histological features of the colonic tissues such as crypts, lamina propria, tumor, etc. When compared to HPS stained images, after whole slide image analyzed with crop and score Calopix module from TRIBVN or through pathologist control, color-coded spectral images not only reveal features representative of the biochemical make up of the tissues, but also highlight additional features like intra-tumoral heterogeneity and tumor-associated stroma, which are difficult to discern by conventional histopathology. The LDA prediction model was promising since an average sensitivity of 91% was achieved in the identification and prediction of tumoral tissues.
**Discussion:** IR imaging allowed differentiating and detecting normal and tumoral colon tissue features based on their intrinsic biochemical information. This chemical-free approach on paraffinized tissue biopsies combined with multivariate statistical image analysis opens a new avenue

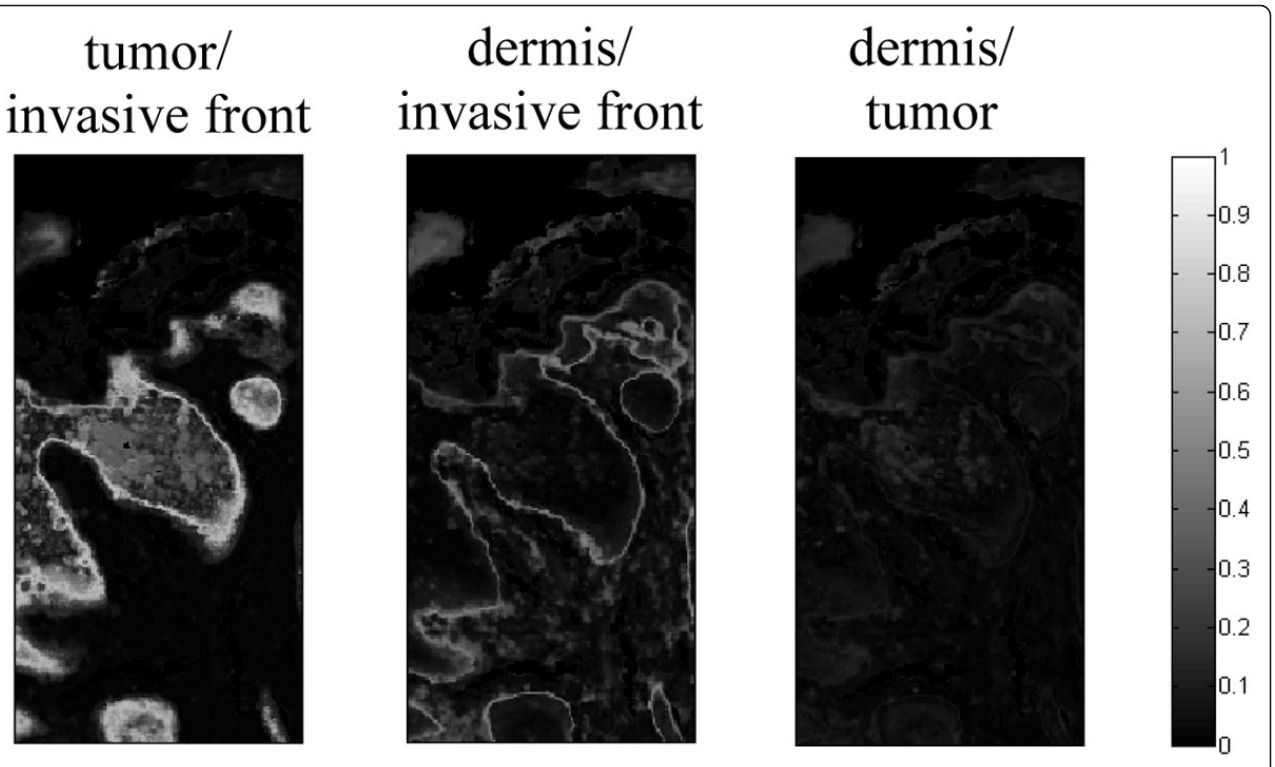



**Figure 2(abstract S33)** Images representing the entropy based interconnectivity measure computed on the upgraded Fuzzy C-means (FCM) images estimated on the SCC sample from Figure 1.

for numerical spectral histopathology and appears as a promising tool for colon cancer diagnosis. Further work to improve the model and to predict tumors in blind samples is ongoing.

**References**

1. **GLOBOCAN.** International Agency for Research on Cancer (IARC) 2008.
2. Ly Elodie, Piot Olivier, Wolthuis Rolf, Durlach Anne, Bernard Philipe, Manfait Michel: **Combination of FTIR spectral imaging and chemometrics for tumour detection from paraffin-embedded biopsies.** *Analyst* 2008, **133**:197-205.

## S35

### CaseConferencing: telecom resource used for an original approach to on-going teaching through case expertise

Jacques Klossa[1*], Frédérique Capron[2], Dominique Algalarrondo[3], Gilles Le Nahour[2], Jean-François Pomerol[1]
[1]TRIBVN, France; [2]CHU La Pitié, France; [3]Orange, France
E-mail: jklossa@tribvn.com

*Diagnostic Pathology* 2013, **8(Suppl 1):**S35

**Introduction:** In the pathology field, digital slide conferencing capabilities are obviously a highly necessary functionality when multiple participants need to simultaneously view the same digital slide from multiple, remote locations e.g. for frozen section teleconsultation, second opinion on complex cases or e-learning. In this presentation we present an expertise approach that uses teleconsultation to transfer the knowledge from the expert to requesters. In that context, we will also analyze advantages and perspectives in integrating a general purpose solution to a business specialty (pathology) workflow.

**Methods:** The idea from the department of Pathology in La Pitié University Hospital center is to take advantage from whole slide image (WSI) teleconsultation to provide an on-going teaching service that will enhance the interpretation level of unspecialized pathologists. Slides are sent to La Pitié for expertise; they are immediately scanned and sent back the same day to the emitter service for storage.

For such application the Orange Multimedia Conference module which is a general purpose piece of software has been integrated to the TRIBVN Calopix™ platform. Calopix™ platform is a dedicated pathology platform which is integrated in the hospital workflow between LIS (Laboratory Information System) and PACS. Multimedia Conference is a web conference service widely used and many thousands conferences are opened daily. To meet the healthcare needs the video module has been enhanced for high image quality, large screen compatibility and exchanges traceability. Such collaboration brings secure data transmission and image compression knowhow to pathology business.

**Results:** The resulting CaseConferencing module allows recently scanned WSI, i.e. for frozen sections, to be shared instantly without any previous server upload between the conference organizer (expert) and up to 25 participants (case emitters and other interested pathologists). The organizer uses the pathology image workstation that allows classical pan and zoom functions as well as annotation or image analysis tools. Invited participants are informed by email and may participate through the use any web browser on PC or Mac. Among main functionalities, session leadership can be transferred to any participant and real time annotations are automatically stored on the organizer PC.

CaseConferencing guarantees higher quality images transmission and speed for a known transmission channel in addition to ensure security and annotation traceability. Such high quality transmitted image allows clear visualization of very thin features like one pixel thick overlays showing image analysis results. Visualization latency is highly dependent on communication network configuration. It has been generally considered as good enough for WSI sharing and in most cases latency was clearly beyond one second. Such service can be efficiently used as a complement of cooperative asynchronous applications commonly used in the field of telepathology when WSI can be shared through a remote server.

The service is currently used since 2011 by the pathology lab in La Pitié University Hospital in Paris to deliver an expertise service toward private and general hospitals. The purpose is to define guideline describing such expertise services so that it could be extended to other university hospitals.

**Discussion:** Visualisation and diagnostic clearly allows on screen diagnostic follow up. However higher "fluidity" of pan and scan would be required to achieve the same feeling as when viewing a slide through a microscope. Another demand that will be answered is the publishing of the session report which has been stored on the organizer PC.
The main purpose of this communication was to evaluate how a platform dedicated to pathology could take profit from existing multipurpose information and telecom tools. Such integration proved to be good enough for case conferencing application. The next step would be to apply such concepts to meet another telepathology issue which needs to share expert knowledge. More specifically, in the Cloud Computing context, the idea would be to use the PaaS (Platform as a service) layer to take profit from the stored patient data information in conjunction with consolidated formalized specialist knowledge to drive WSI exploration and to produce automated pre-annotation that will make easier and quicker their on-line consultation.
**Competing interests:** The authors declare that they have no competing interests.
**Authors' contributions:** • FC and GLN provided the original ideas and carried out the evaluation
• JFP in conjunction with DA drove the implementation of the CaseConferencing application to telepathology
• JK and DA contributed to the technical implementation and extended the "CaseConferencing" to "FlexMIm" cloud implementation

## S36

### A visual model approach to extract regions of interest in microscopical images of basal cell carcinoma

Ricardo Gutiérrez[†], Eduardo Romero[*†]
CIM&LAB – Telemedicine Centre, Universidad Nacional de Colombia, Carrera 30 No. 45-03, Medicine Faculty, Building 471, Bogotá, Colombia
E-mail: edromero@unal.edu.co
*Diagnostic Pathology* 2013, **8(Suppl 1):**S36

**Background:** The virtual microscopy is a discipline that emulates the interaction between an expert with a microscopical sample upon a high resolution digital slide [1]. This type of technology is used for medical training and medical education, but so far it has been exclusively used in research environments because of the large computational requirements [2]. For instance, after digitizing *1cm2* of a physical slide at a level of ×20 magnification, the resulting virtual slide amounts to a 4GB [3], that must be processed, transmitted, explored and analyzed in real time. This picture can be even worsen since a typical laboratory takes hundreds of slides each week [4].
Overall, a typical pathologist does not explore the entire slide, but instead she/he focus her/his analysis on a few number of visual fields or regions of interest (RoI). In consequence, recognition of RoI in microscopical images may be a potential source of knowledge in many diagnostic tasks. Such RoIs would introduce new learning paradigms that would be used in medical education, medical training and diagnosis assistance. In addition, a precise determination of such regions can highly reduce the computational and transmission charge of informative regions from a sample. However, automatic recognition of such regions is really a challenging task because of the inherent randomness of tissue's cutting, color tissue properties and tissue orientation.
In spite of these difficulties, the pathologist efficiently recognizes regions of interest in several domains by fusing image and task dependent information into a unique framework. This paper proposes a novel automatic approach to recognize RoIs by emulating the processing of the human visual system (HVS), not only modeling the preattentional process but also integrating it with high level processes. Hence, this paper extends our previous work [5] by including structural information about the relationships between several objects and texture recognition as higher cortex functions. These processes are necessary to minimally perceive the core of a scene, just as it is carried out within the pathologist memory [6], and therefore, to identify relevant regions for diagnosis.
**Material and methods: Experimental setup:** The model was tested with a total of 115 histological microscopical fields of view of different types of basal cell carcinoma, sampled from 25 randomly chosen patients. Each biopsy was formalin-fixed and stained with Hematoxylin-Eosin dyes. Microscopical fields were digitized with a Nikon eclipse E600 system, through a coupled Nikon DXM1200 camera, and stored in JPEG format at a 1280 × 1024 resolution using microscope magnifications of ×4, ×10, ×20 and ×40. An expert pathologist, with at least five years of experience, selected the digitized fields of view and manually segmented relevant regions. We use 20 images to extract textons of size 32 × 32 pixels from RoIs and background for the object recognition task, and 95 images to test the entire algorithm.
**Method overview:** During a standard exploration of a histopathological sample, the pathologist integrates two types of information sources, namely, 1) the visual field content itself ("bottom-up" source), and 2) the knowledge involved in the specific task of diagnosis ("top-down" source). The HV S fuses together these sources using at least four different brain areas, namely, 1) the V1 cortex which assigns a local relevance to the visual input, 2) the V2 cortex which is responsible of gathering together these relevancies as simple shapes, 3) the V4 cortex which actively regulates the V2 excitation, and finally, 4) the inferotemporal gyrus which integrates the function of the previous areas by recognizing complex shapes and their purposes in the scene and probably by retrieving a slight image representation composed of a few objects actually recognized, the relations between them and the background, and basic information about the background texture [6]. The approach proposed herein models the function of the HV S in four steps, as shown in Figure 1. Firstly, it assigns a local relevance by integrating information from basic features as orientation, color and intensity at multiple scales, as previously described by Itti *et al*. [7] (V1 cortex function). Secondly, the model segments the image by taking into account the proximity and similarity between pixels and mixing up the conspicuity maps into the resulting regions (V2 cortex function), but also adding a map of the intrinsic structural disorder which models the specific task knowledge that regulates the attention over each structure (V4 cortex function) [5]. Thirdly, this intermediate map is used to feed a module that efficiently looks for basal cell carcinoma by comparing the pattern composition with a data base of carcinoma regions, starting with the highest attentional regions (inferotemporal gyrus function). Once the algorithm sets up the first carcinoma region within the image, the other carcninoma regions are defined as the most similar regions using an Euclidean metrics of the different basic features: color, entropy and image intensity.
Then, this intermediate map is used to feed a module that efficiently looks for basal cell carcinoma by comparing the pattern composition with a data base of carcinoma regions, starting with the highest attentional region (inferotemporal gyrus function). Once the algorithm sets up the first carcinoma region within the image, the other carcninoma regions are defined as the most similar regions using an Euclidean metrics of the different basic features: color, entropy and image intensity.
**Structure search and recognition:** During a searching task of an undetermined number of targets, the pathologist determine whether she/he is familiar with the scene, using simple features as color and texture information, possibly recognizing some relevant objects and their relations [6,8]. If this is the case, she/he efficiently looks for the targets, following the saliency of the scene and then using the target contextual relationships [8,9]. Likewise, our proposal follows the most saliency levels of the map obtained in the local recognition step [5], that is say the third stage, when the system has determined a target. We extract then some textons from RoIs previously marked (object recognition) and some from the background (basic background features extraction) to discard irrelevant structures. A classical kNN algorithm was used, using the low level features captured before as input [10]. In case of finding a target, its relations with the other structures in the image are determined, comparing how different its internal features are. Formally, the distance between features (df) within the target structure (st) and the i-structure (si) is defined as the Euclidean mean distance of the intensity features (I), orientation (O), color (C) and entropy(H) inside each structure, formally, 0 identifying similar structures as part of the searched targets using a simple threshold. Note that this algorithm addresses the intensive computations to a few structures, until it reaches the first target.
**Results and discussion:** The method herein presented was compared with a previous one [5] that outperformed the state of the art. The model proposed uses as model a generic "top-down" information source [6], taking advantage of its relation with the context. The inferotemporal gyrus function improves the RoI selection, as illustrated in Figure 2. Note that the model proposed is more selective to the basal cell carcinoma structures since it is able to carefully identifies specific low level mixtures

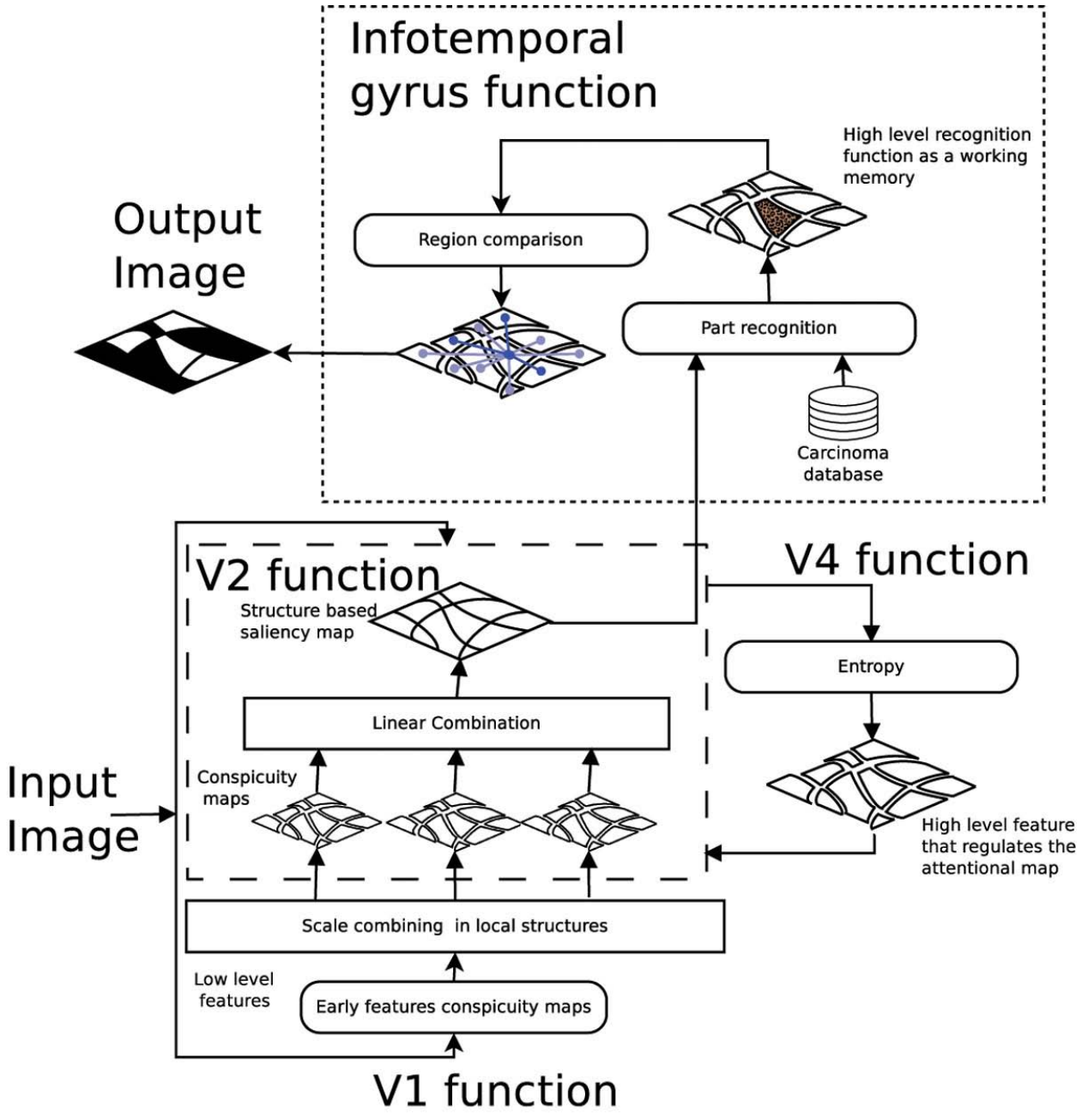



**Figure 1(abstract S36) The proposed model** The proposed model fuses the bottom-up and top-down information. Firstly, assigning a level of conspicuity based in a set of early characteristics, secondly, it gather together this information into simple shapes, thirdly, by regulating the saliency of each shape by the texture information as it is reported a relevant characteristic to diagnosis of cancer. Finally, our method uses the resulting saliency map to look for carcinoma structures using a database and a similarity measure over its neighborhood.

present in the first target reached and choose a clever similarity rule to define the discriminative threshold.

Also, the proposed model is more robust to distractors at several scales of magnification of the virtual slide. Our actual approach was tested with images at different scales of magnification, without any restriction at the training stage. The current method improves the results in specificity and reduces the variability of the results. These results are presented in the Table 1.

**Conclusions:** This paper presented a novel methodology to find RoI based on the human visual system. This differs from our previous approach by the inclusion of a stage of region recognition and evaluation of inter-region similarity. These characteristics let us improve the RoI extraction since the selection criteria are modified by a knowledge database.

**List of abbreviations:** HVS: Human Visual System; kNN: k-Nearest Neighbour; RoI: Region of Interest

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** RG developed the algorithms and evaluated the results of the visual model. ER conceived the study, developed the

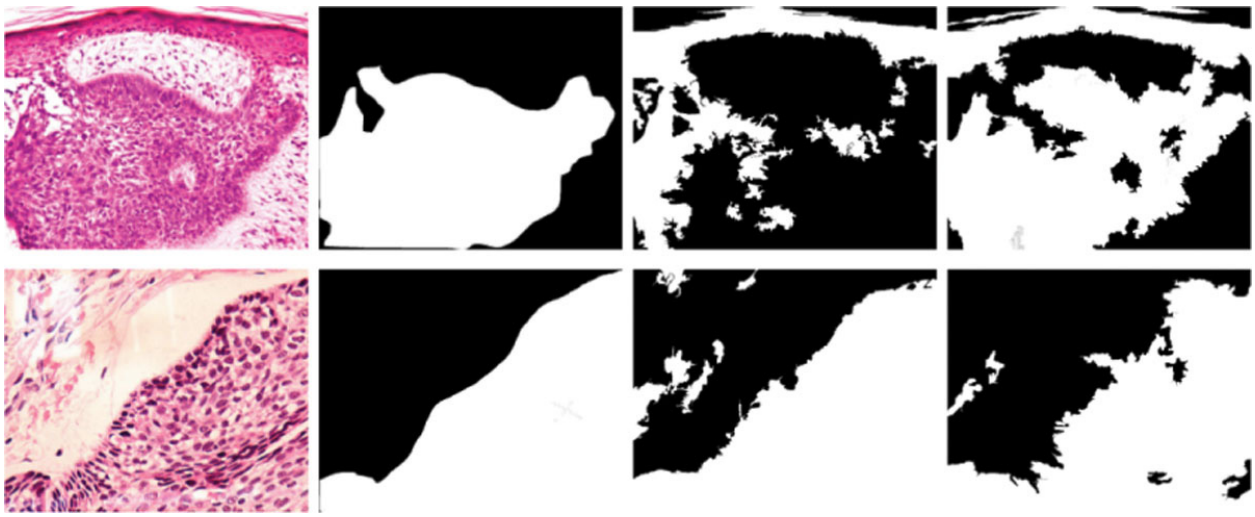


**Figure 2(abstract S36) Comparison between the previous work and the current one** From left to right, first column present the original images, the second one the ground truth marked by the pathologist, the third column presents the results obtained with Gutiérrez *et al.* (2011) strategy, and finally, the fourth column shows the results obtained with our method.

**Table 1(abstract S36) Statistical results**

| | Mean (variance) [%] | |
|---|---|---|
| | Gutiérrez *et al.* (2011) | Ours |
| Sensitivity | 86.6 (27.5) | 80.8 (17.8) |
| Specificity | 37.6 (23.7) | 63.6 (19.0) |

Sensitivity and Specificity results computed over several magnifications.

fundamental ideas underlying this model, participated in the experimental design and was the director of the whole project. All authors read and approved the final manuscript.

**Acknowledgements:** This work was partially funded by de project "Sistema para la recuperación de imágenes médicas utilizando indexación multimodal" Grant: 110152128767 from "CONVOCATORIA COLCIENCIAS 510 de 2010".

**References**

1. Catalyürek U, Beynon MD, Chang C, Kurc T, Sussman A, Saltz J: **The Virtual Microscope.** *IEEE Trans. Inf. Technol. Biomed* 2003, **7(4)**:230-248.
2. Tsuchihashi Y, Takamatsu T, Hashimoto Y, Takashima T, Nakano K, Fujita S: **Use of virtual slide system for quick frozen intra-operative telepathology diagnosis in Kyoto, Japan.** *Diagnostic Pathology* 2008, **3(Suppl 1)**:S6.
3. Iregui M, Gómez F, Romero E: **Strategies for efficient virtual microscopy in pathological samples using JPEG2000.** *Micron* 2007, **38(7)**:700-713.
4. Doyle S, Monaco J, Madabhushi A, Lindholm S, Tomaszewski J: **Evaluation of effects of JPEG2000 compression on a computer-aided detection system for prostate cancer on digitized histopathology.** *IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 2010, 1313-1316.
5. Gutiérrez R, Gómez F, Roa-Peña L, Romero E: **A supervised visual model for finding regions of interest in basal cell carcinoma images.** *Diagnostic Pathology* 2011, **6(26)** [http://www.diagnosticpathology.org/content/6/1/26].
6. Wolfe JM: **What do you know about what you saw?** *Current Biology* 1998, **8(9)**:303-304.
7. Itti L, Koch C, Niebur E: **A model of saliency-based visual attention for rapid scene analysis.** *IEEE Transactions on patern analysis and machine intelligence* 1998, **20**:1254-1259.
8. Wolfe JM, Horowitz TS, Michod KO: **Is visual attention required for robust picture memory?** *Vision Research* 2007, **47**:955-964.
9. Kunar MA, Flusberg S, Wolfe JM: **The role of memory and restricted context in repeated visual search.** *Perception & Psychophysics* 2008, **70(2)**:314-328.
10. Cover T, Hart P: **Nearest neighbor pattern classification.** *Information theory, IEEE Transactions* 1967, **13**:21-27.

## S37

### Automated segmentation of blood cells in Giemsa stained digitized thin blood films

Margarita Walliander[1*], Riku Turkki[1,2], Nina Linder[1], Mikael Lundin[1], Juho Konsti[1], Ville Ojansivu[1,2], Taru Meri[3], Ville Holmberg[3], Johan Lundin[1]
[1]Institute for Molecular Medicine Finland (FiMM), Finland; [2]Center for Machine Vision Research, Department of Computer Science and Engineering, University of Oulu, Finland; [3]Haartman Institute, Finland
E-mail: margarita.walliander@helsinki.fi
*Diagnostic Pathology* 2013, **8(Suppl 1):**S37

**Background:** Assessment of erythrocytes and leucocytes in thin blood films can be used as an inexpensive diagnostic aid in a series of disease states, e.g. infections, anemia and hematological malignancies. Manual counting of cells is still considered the gold standard for example to establish the level of parasitemia in malaria. However, manual cell counting is time consuming and subject to variability [1]. We here propose an image analysis method that is a combination of adaptive histogram thresholds and morphologic characteristics for the segmentation of red blood cells (RBCs) and white blood cells (WBCs) in digitized thin blood films. The method is implemented on a virtual microscopy platform, the Webmicroscope [2].

**Methods:** Ten Giemsa stained thin blood films were digitized with a microscopy slide scanner (Axio Imager Z2, Carl Zeiss MicroImaging, Jena controlled by Metafer software, MetaSystems, Altlussheim) using a 63x objective with a numerical aperture (NA) of 1.4 (Plan-Apochmat, Carl Zeiss, Jena) and oil immersion. Image acquisition was performed with a monochrome CCD camera with a 1360x1024 pixel sensor and a pixel size of 6.45 μm (CoolCube 1, MetaSystems, Altlussheim), a 1.0 camera adapter and illumination with an RGB illuminator (MetaLED Z, red 619 nanometer, green 515 nanometer, blue 465 nanometer, MetaSystems, Altlussheim). The pixel size in the digital images was approximately 0.10μm and the original TIFF images were converted into a wavelet format (Enhanced Compression Wavelet, ERDAS/ Intergraph, Norcross, GA) and transferred to a virtual microscopy image server (http://fimm.webmicroscope.net/Research/Momic/tp2012) [3]. Approximately five-hundred (473 – 505) fields of view from each blood film sample were captured and stored in the database. Five of the samples were infected with *Plasmodium falciparum* and five were non-infected control samples.

The described method (Fig. 1) involves 1) separation of background and foreground, 2) recognition of objects that compose the foreground and 3) cell counting (i.e. RBCs and WBCs).

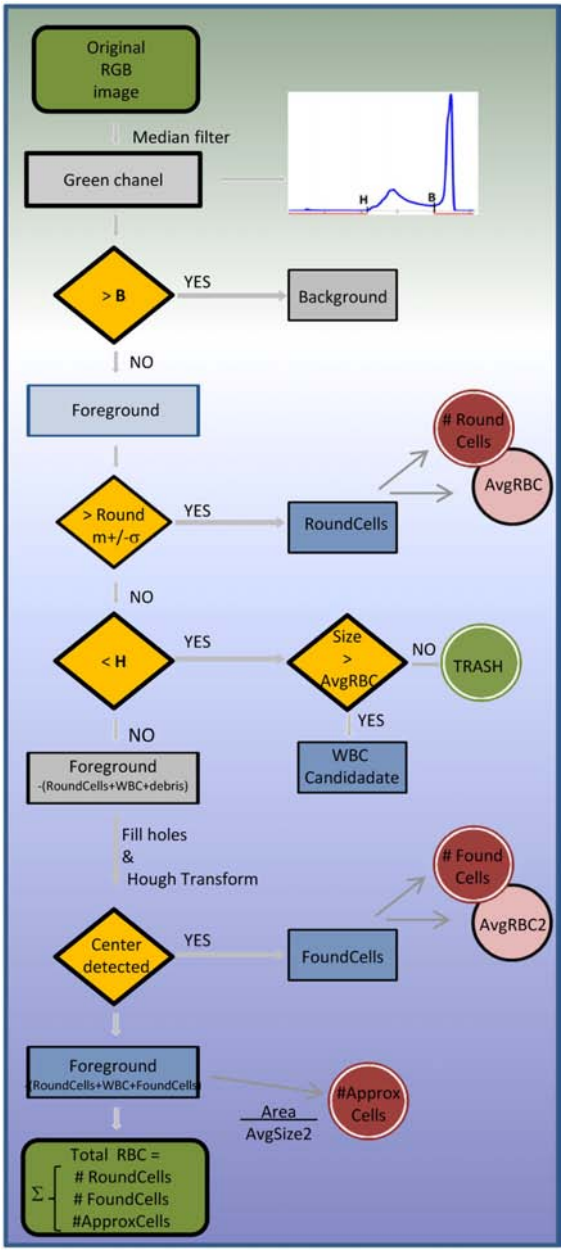



**Figure 1(abstract S37) Flowchart showing the cell segmentation process** Overview of the cell segmentation and counting process for red and white blood cells. From the original image the green channel is selected and smoothed. Dynamic thresholding allows the separation of the image in foreground and background. The foreground is split in to *RoundCells*, *WBC Candidates*, *FoundCells* and the remaining part being deformed red blood cells and clumps of red blood cells. These subimages are used to count the objects of interest i.e. red blood cells and white blood cells.

**Image preprocessing:** As a preprocessing step for each thin blood film sample, the green channel was selected from the original RGB image [4] and smoothed by applying a median filter 3X3 to reduce the' salt and pepper' noise [5]. The green channel is extracted using a color deconvolution between the original image and a vector [0,1,0].

**Adaptive histogram thresholds:** Let each pixel of the preprocessed image have intensity levels in *[0, 1, 2, …, L-1]* with *L= 256*. The number of pixels with intensity level *i* is denoted by $n_i$, $\forall$ *i=0, 1, 2,…, L-1*, where

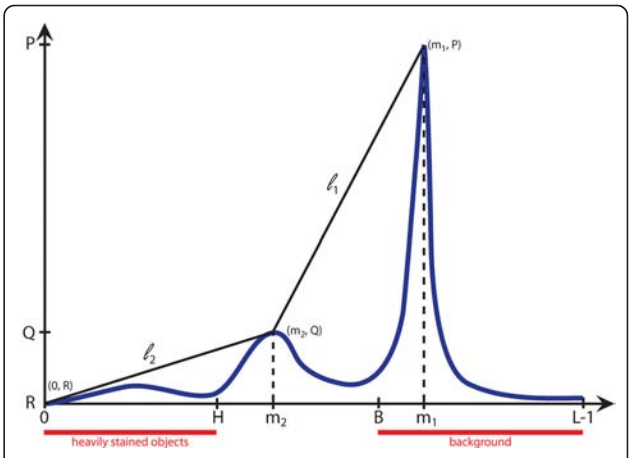



**Figure 2(abstract S37) Bimodal histogram depicting adaptive thresholds location, background intensity range and heavily stained objects range** A gray level histogram that corresponds to the green channel of a typical field of view after applying a median filter (3X3). The mode of the histogram are located at $m_1$ and $m_2$. B and H are the adaptive thresholds. $l_1$ is a line defined between the modes of the histogram and $l_2$ is a line defined between the origin of the histogram and the mode at $m_1$.

$N = \sum_{i=0}^{L-1} n_i$ is the total amount of pixels. We defined the histogram distribution as $p(i) = n_i/N$, $p(i) \geq 0$, where $\sum_{i=0}^{L-1} p(i) = 1$.

For any monolayer stained blood film, the histogram is bimodal. A typical histogram shape for a monolayer thin blood film is shown in Fig. 2.

There are two local maxima located at $m_1$ and $m_2$, where $m_2 < m_1$ and $P = p(m_2) < p(m_1) = Q$.

A threshold to differentiate the background from the foreground *B*, is defined by finding the maximum distance between the histogram and the line $l_1$ described between $(m_1, P)$ and $(m_2, Q)$ as shown in Fig. 2.

$$d(p(B), l_1) = \max\{d(p(x), l_1) \,|\, x \in [m_2, \ldots, m_1]\}$$

with

$$l_1 = \left\{(x, y) \,|\, y = ax + b,\ \forall x \in [m_2, \ldots, m_1], a = \frac{P - Q}{m_1 - m_2}, b = Q - m_2 * \frac{P - Q}{m_1 - m_2}\right\}$$

In a similar manner, a threshold to find heavily stained objects *H*, was defined by finding the maximum distance between the histogram distribution and the line $l_2$ described between *(0, R)* and $(m_1, P)$, with $R = p(0)$;

$$d(p(H), l_2) = \max\{d(p(x), l_2) \,|\, x \in [0, \ldots, m_2]\}$$

and

$$l_2 = \left\{(x, y) \,|\, y = ax + b,\ \forall x \in [0, \ldots, m_2], a = \frac{Q}{m_2}, b = 0\right\}$$

The image background is given by Eq. (1), while the heavily stained objects are described by Eq. (2).

$$Background = \{f(x, y) > B \,|\, f(x, y) \in D\}, \quad (1)$$

$$Heavily\ Stained\ Objects = \{f(x, y) < H \,|\, f(x, y) \in D\}, \quad (2)$$

where = gray level Image.

**Table 1(abstract S37) Results comparing manual and automated cell counting**

| | Red Blood Cells | | | White Blood Cells | | |
|---|---|---|---|---|---|---|
| Sample | Annotated | Automated | Error % | Annotated | Automated | Error % |
| I1 | 3145 | 3160 | 0.476948 | 20 | 20 | 0 |
| I2 | 4048 | 4058 | 0.247036 | 34 | 34 | 0 |
| I3 | 2796 | 2782 | 0.500715 | 22 | 22 | 0 |
| I4 | 2972 | 2958 | 0.471063 | 30 | 32 | 6.66 |
| I5 | 3047 | 3042 | 0.164096 | 77 | 75 | 2.59 |
| C1 | 3396 | 3389 | 0.206125 | 75 | 75 | 0 |
| C2 | 3491 | 3482 | 0.257806 | 55 | 56 | 1.81 |
| C3 | 3093 | 3087 | 0.193986 | 42 | 42 | 0 |
| C4 | 3197 | 3206 | 0.281514 | 49 | 50 | 2.04 |
| C5 | 3513 | 3514 | 0.028466 | 71 | 72 | 1.38 |
| TOTAL | 32698 | 32678 | 0.061166 | 476 | 477 | 0.21 |
| % | 100 | 99.9388 | 100 | 99.79 | | |

Results comparing manual and automated counting for red blood cells and white blood cells (WBCs). Red blood cells were annotated in a region equivalent to 30 fields of view per film, while the annotations for WBCs were performed on 500 fields of view per film. The samples I1-I5 are *Plasmodium falciparum* infected cases, C1-C5 are non-infected controls.

**RoundCells separation:** Using the image histogram and based on its bimodal shape, two important thresholds were extracted (Fig. 2). The first threshold (B) defines a binary separation between the image background and foreground. From the image foreground, all objects with roundness bigger than 0.6 were selected and the area of each of the objects was measured. The mean diameter to be 7.52μm and standard deviation of 0.06μm for the whole set of objects was calculated. Finally, only the subset of objects with an area equal to $m$ +/-$\sigma$ (3848+/-688 pixels) was chosen and defined as *RoundCells*. From this set of round objects of similar size, the average diameter was calculated and used to define a representative red blood cell, designated as *AvgRBC* (diameter ~7μm) and to establish limit diameters for WBCs (~7-21 μm) and platelets (~2-3μm).

The second threshold (H), defines the heavily stained objects in the foreground (i.e. WBC, platelets, artifacts and debris). The heavily stained objects larger than *AvgRBC* are the *FoundWBCs*.

**Detection of circular shapes by Hough transform:** Hough transform is calculated on gray level images that contain only the regions of interest while the remaining is set to zero. The region of interest is composed by foreground without *RoundCells*, *FoundWBCs* and debris. The maximization of Hough transform for a radius interval $\left[\frac{r}{2}, \frac{3r}{4}\right]$ is performed, where $r$ = *radius* (*AvgRBC*). The result is a set of accumulations of hits (votes). The accumulations are concentrated around the centers of the circular shapes. Hough transform detected cells are filtered by selecting the pixel with the maximum vote and deleting all the pixels with less than 20% of the votes. Thus the selection of nearly circular shapes is ensured. Finally, a morphological opening is performed to discard accumulations with less than 50 pixels. The remaining objects are centers of *FoundCells*.

**FoundCells detection:** After subtracting the *RoundCells* and the heavily stained objects from the original image, to compensate for the holes left from the subtraction of the platelets, debris and parasites, a morphologic filling was performed. By using Hough transform, circular shapes were detected in the grayscale image and designated as *FoundCells*, resolving the center positions of the nearly circular objects. A second representative red blood cell *AvgRBC2* was defined from the area of *FoundCells*.

**ApproxCells detection:** After subtracting the *RoundCells*, the *FoundWBCs* and the *FoundCells*, the remaining image contains fragments of RBCs and deformed RBCs which Hough transform was not able to define as circular shapes. The total area covered by these objects, named *ApproxCells* was divided by the area of *AvgRBC2* which is estimation for the number of cells that still remain without being counted *#ApproxCells*.

Finally, the total number of RBCs in the image is calculated by summing up the partial results;

$$\#TotalRBC = \#(RoundCells) + \#(FoundCells) + \#(ApproxCells) \quad (3)$$

**Results and discussion**

RBCs were manually annotated in 30 fields of views per thin blood film and WBCs were annotated in the entire data set (Table 1). The results from the manual counting and automated counting are shown (Table 1.) Using the annotated fields of view, automated quantification of RBCs and WBCs was compared against the manual annotations and RBCs showed an overall error rate of 0.06%, WBCs counting showed an overall error rate of 0.21%. A test for the automated counting of RBCs and WBCs was performed on whole slides of thin blood films and approximately half a million red blood cells and 477 white blood cells were counted (Table 2.). Previously published studies have addressed the separation and counting blood cells, but fixed thresholds for colors, sizes and intensity values restrict the use to particular data sets [6-10]. Here, we make use of adaptive thresholds for size and intensity values, which converges to a solution.

**Conclusions:** The segmentation of RBCs and WBCs is an easy task for a human observer. Humans have the ability of distinguishing large number of colors, shades and hues, also estimating shapes and size similarities while referring to prior knowledge, making global and local comparisons simultaneously. However, performing large scale quantification is a time consuming and tedious task.

We present an unsupervised tool for separating the foreground from the background in Giemsa stained thin blood films and an automated cell counter for RBCs and WBCs. The segmentation of blood cells in thin blood films can be used as a pre-processing step to specify the regions of interest for a secondary algorithm, e.g. the detection of malaria parasites in RBCs, morphological analysis of RBCs and WBCs and follow-up during treatment of hematological malignancies or measurement of response to chemotherapy.

**List of abbreviations:** RBC: Red blood cell; WBC: White blood cell; #: Number of cells in

**Competing interests:** The authors declare that they have no competing interests.

**Acknowledgements:** The authors wish to thank Elisabet Tyyni for sample preparation and analysis. The study was kindly supported by the national Biomedinfra and Biocenter Finland projects.

**Table 2(abstract S37) Results of automated red blood cell counting on whole slides of thin blood film**

| | Cell counting | | | | |
|---|---|---|---|---|---|
| sample | #TotalRBC | #RoundCells | #FoundCells | #ApproxCells | AvgRBC diameter μm |
| I1 | 59333 | 17935 | 40512 | 886 | 7.6804 |
| I2 | 70236 | 16458 | 52600 | 1178 | 7.5457 |
| I3 | 43973 | 23068 | 20475 | 430 | 7.7613 |
| I4 | 46980 | 14237 | 32438 | 305 | 7.8404 |
| I5 | 46090 | 13670 | 31918 | 502 | 7.5258 |
| C1 | 57760 | 18379 | 38993 | 388 | 7.3918 |
| C2 | 53645 | 23704 | 29669 | 272 | 6.987 |
| C3 | 45462 | 22290 | 22669 | 503 | 7.5385 |
| C4 | 51605 | 18108 | 32674 | 823 | 7.2068 |
| C5 | 58029 | 16656 | 41098 | 275 | 7.7327 |
| TOTALS | 533113 | 184505 | 343046 | 5562 | |

Partial and total amount of red blood cells, automatic counting on 500 fields of view per sample. The first five samples are *Plasmodium falciparum* infected cases I1-I5, samples C1-C5 are non-infected controls.

**References**

1. O'Meara WP, Barcus M, Wongsrichanalai Ch, Muth S, Maguire JD, Jordan RG, Prescott WR, McKenzie FE: **Reader technique as a source of variability in determining malaria parasite density by microscopy.** *Malaria Journal* 2006, **5**:118.
2. Linder E, Lundin M, Thors C, Lebbad M, Winiecka-Krusnell J, Helin H, Leiva B, Isola J, Lundin J: **Web-based virtual microscopy for parasitology: a novel tool for education and quality assurance.** *PLoS Negl Trop Dis* 2008, **2(10)**:e315, Epub 2008 Oct 22.
3. **The WebMicroscope virtual microscopy environment.** [http://www.webmicroscope.net/].
4. Wermser D, Haussman G, E LC: **Segmentation of blood smears by hierarchical thresholding.** *Computer Vision, Graphics and Image Processing* 1984, **25**:151-168.
5. Gonzalez RC, Woods RE: **Digital Image Processing.** New Jersey: Prentice Hall 2002.
6. Kumar R, Joseph DK, Sreenivas TV: **Teager Energy based blood Cell Segmentation.** *14th International Conference on Digital Signal Processing 2002* **2**:619-622.
7. Tek FB, Dempster Ag, Kale I: **Computer vision for microscopy diagnosis of malaria.** *Malaria Journal* 2009, **8(53)**.
8. Dorini LB, Minetto R, Leite NJ: **White Blood cell segmentation using morphological operators and scale-space analysis.** *SIBGRAPI, 2007. XX Brazilian Symposium on Computer Graphics and Image Processing* 294-304.
9. Ramesh N, Dangott B, Salama ME, Tasdizen T: **Isolation and two-step classification of normal white blood cells in peripheral blood smears.** *Journal of Pathology Informatics* 2012, 149-170.
10. Purwar Y, Shah S, Clarke G, Almugairi A, Muehllenbachs A: **Automated and unsupervised detection of malaria parasites in microscopic images.** *Malaria Journal* 2011, **10**:364.

## S38

### Pap smear cell image classification using global MPEG-7 descriptors

Luz H Camargo[1†], Gloria Diaz[2†], Eduardo Romero[3*]
[1]Faculty of Engineering – Cra. 7 No.40 53, Universidad Distrital Francisco José de Caldas, Bogotá D. C. – Colombia; [2]Faculty of Engineering – Cra. 5 No.21 38, Universidad Central, Bogotá D. C. – Colombia; [3]Bioingenium, National University of Colombia, Cra 30 No 45 03-Ciudad Universitaria, Faculty of Medicine - Building 471, National University of Colombia, Bogotá DC - Colombia
E-mail: edromero@unal.edu.co
*Diagnostic Pathology* 2013, **8(Suppl 1):**S38

**Background:** Several strategies have been previously applied for classifying cervical cytology cells, all pursuing a nucleus segmentation. Sanchez sets regions [1] using a simple threshold [2], a procedure broadly adapted to different techniques: a local adaptive segmentation nuclei procedure [3], seed growing [4], mathematical morphology [5], a Hough transform [6], and active contours [7]. Jantzen and Dounias propose several cell features as morphometric descriptors, including the nucleus and cytoplasm areas, nucleus/cytoplasm proportion, nucleus and cytoplasm brightnesses, smaller and larger nucleus/cytoplasm diameters, nucleus and cytoplasm roundness, nucleus and cytoplasm perimeters, nucleus position, nucleus/cytoplasm maxima and minima. Nevertheless, these morphometric characteristics require a previous accurate segmentation, hardly achieved by human intervention using commercial software such as CHAMP (Cytology and Histology Modular Analysis Package, Aarhus, Denmark) or DIMAC (Digital Image Company) [8,9].

**Methods:** Rather than attempting to detect some of the previously reported morphometric features, the present investigation used two global MPEG-7 color descriptors, *Color Layout* and *Scalable Color*, and one texture descriptor, the *Edge Histogram descriptor*, as the representation space and two supervised classification algorithms (SVM and KNN) that divide the different classes.

**Classification based on global MPEG-7 descriptors:** The cell classification approach is carried out using color and texture MPEG-7 descriptors, thereby attempting to capture information related with the particular color spatial location and global color distribution of both the nucleus and cytoplasm. The texture descriptor stands for the particular borders of both nucleus and cytoplasm and their intrinsic relationships. These global characteristics are not evaluating the classical morphometric features, but they are using nucleus and cytoplasm visual primitives as discriminant factors.

**Color layout:** This descriptor, typically used in the YCrCb color space, captures the spatial color distribution in an image or an arbitrary region. Basically, the color layout descriptor uses representative colors on an grid, followed by a Discrete Cosine Transform (DCT) and an encoding of the resulting coefficients. The feature extraction process consists of two parts; grid based representative color selection and DCT transform with quantization. Specifically, an input image is divided into 64 blocks, their average colors are derived and transformed into a series of coefficients by performing a conventional DCT. A few low-frequency coefficients are selected using zigzag scanning and quantized to form a Color Layout Descriptor [10].

**Scalable color:** This descriptor is extracted from a color image histogram in the hue-saturation-value (HSV) color space. This histogram, constructed with fixed color space quantization, is projected into a set of Haar bases so that the obtained coefficients constitute a scalable color representation. The histogram values must be normalized and non linearly mapped into a 4-bit integer representation, giving higher weight to small values. The Haar transform is applied then to this histogram version with two basic operators: sum and difference bin neighbor, decomposing the histogram into low and high frequency subbands [10].

**Edge histogram:** This descriptor captures the spatial edge distribution, a very useful feature for image matching, even though the underlying texture may not be homogeneous. A given image is first sub-divided into sub-images and local edge histograms, for each of these sub-images, are computed. Edges are then coarsely grouped into five categories: vertical, horizontal, 45 diagonal, 135 diagonal, and isotropic (nonorientation specific). Thus, each local histogram has five bins corresponding to the above five categories. The image partitioned into 16 sub-images results in 80 bins. These bins are nonuniformly quantized using 3 bits/bin, resulting in a descriptor with size of 240 bits [10].

**Classification models:** The classification method used a classical K-Nearest Neighbor algorithm and a Support Vector Machine. The proposed approach was evaluated under a 10-fold experimental setup.

**The k-NN decision rule:** The k-nearest neighbor method is an intuitive method that classifies unlabeled samples based on their similarity with samples in the training set. Given the knowledge of N prototype features (vectors of dimension ∑) and their correct classification into M classes, the k-NN rule assigns an unclassified pattern to the class that is most heavily represented among its k neighbors in the pattern space, under some appropriate metric. In this work euclidean distance was used.

**The SVM algorithm:** A support vector machine (SVM) is a classification model that finds an optimal separating hyperplane that discriminates two classes. A SVM is a linear discriminator, however it can perform non-linear discriminations thanks to the fact that this is a kernel method. In this work, it is used a SVM version that uses sequential minimal optimization algorithm. The multi-class classification problem is solved using a one vs. all strategy: a binary classifier for each class by labeling the class samples as positive examples and other class samples as negative ones. The final decision is set to the class having the largest decision function among all classes.

**Results and discussion: Database:** Two databases composed of images with single cells, from the Herlev University Hospital, Denmark, were used (http://labs.fme.aegean.gr/decision/downloads). Skilled cyto-technicians and doctors manually classified each cell in 2 classes: abnormal and normal and then subclassified into seven classes. Each cell was examined by two cyto-technicians, and difficult samples also by a doctor. In case of disagreement, the sample was simply discarded (Byriel, Martin, Norup, Jantzen). Finally there are two database:

B_1 data contains 500 cells with the following distribution:

1. Normal: columnar epithelial, parabasal squamous epithelial, intermediate squamous epithelial, superficial squamous epithelial.
2. Abnormal: mild squamous non-keratinizing dysplasia, moderate squamous non-keratinizing dysplasia, severe squamous non-keratinizing dysplasia.

The B_2 data contains 917 cells with the following distribution:

1. Normal: superficial squamous epithelial, intermediate squamous epithelial, columnar epithelial.
2. Abnormal: mild squamous non-keratinizing dysplasia, moderate squamous non-keratinizing dysplasia, severe squamous non-keratinizing dysplasia, squamous cell carcinoma in situ intermediate.

**Experimental Setup:** Two classification algorithms were assessed (KNN and SVM). Each classification model was tuned independently for its own particular set of parameters as follows: k-NN was assessed by varying the k nearest neighbors between 1 and 15 with increment steps of two. SVM used two kernel types were evaluated; radial basis function (RBF) and polynomial functions. For the RBF kernel, the $\gamma$ parameter was varied from 0.00 to 0.90 with increment steps of 0.10, while the polynomial kernel degree was set at 1, 2 and 3. The regularization parameter Complexity 1.0. Evaluation was carried out with both 2-class (normal and anormal) problems. A conventional 10-fold cross validation was performed for every parameter combination.

**Results:** The effects of different levels of complexity were evaluated, but do not show important variations. Figure 1 shows the percentage error for the method proposed in the 2-class problem for B_1 datasets. Better performance was achieved by the KNN classifier with k 15. The B_2 database shows the performance of SVM with radial base kernel, for values of 0.01 (Figure 2). In both database these results were obtained with the edge histogram descriptor.

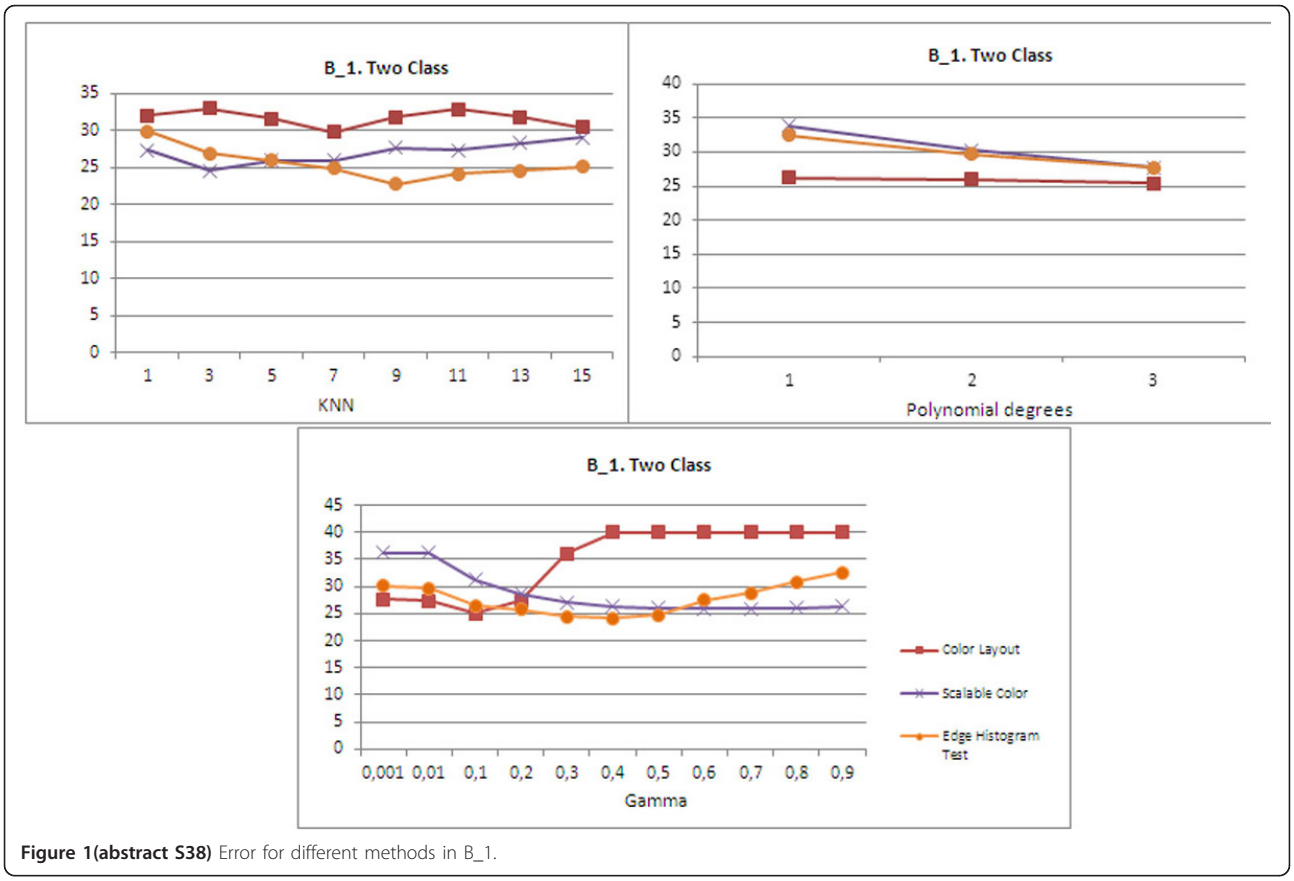



**Figure 1(abstract S38)** Error for different methods in B_1.

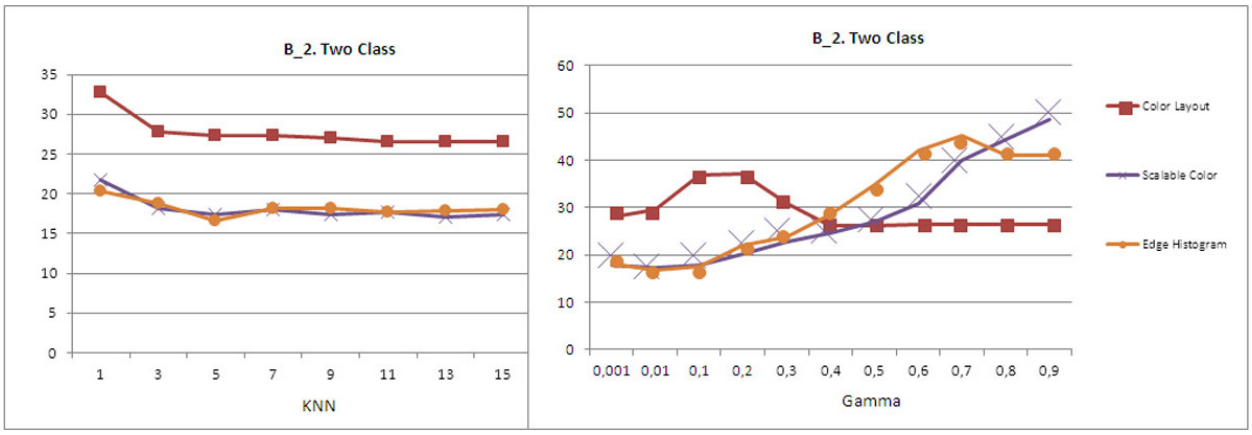



**Figure 2(abstract S38)** Error for different methods in B_2.

**Discussion:** A new strategy for assisting the diagnosis of pap smears which requires no segmentation was proposed, implemented and evaluated. Instead of attempting to segment cells and its structural components, we propose to characterize the internal cell structure using well known MPEG-7 global descriptors. Results are very promising, however evaluation of other specific features and classification algorithms can improve the classification performance.

**Conclusions:** This work presented an original method for discrimination of each class: normal and abnormal. Future work is to evaluate the classification of the seven classes.

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** All authors read and approved the final manuscript.

**Acknowledgements:** This study was supported partially by the Research Headquarters Address Bogota - DIB, National University of Colombia, QUIPU code 202010011343, HERMES code: 8521.

**References**

1. Sánchez A: **Aplicaciones de la visión artificial y la biometría informática.** Dykinson S.L., (Madrid), U. R. J. C. 2005.
2. Passariello : **Imagénes médicas. Adquisición, Analisis, procesamiento e Interpretación.** *Ediciones de la Universidad Simón Bolívar* 1995.
3. Li Z, Najarian K: **Biomedical image segmentation based on shape stability.** *IEEE International Conference on Image Processing* 2007, 281-284.
4. Romberg J, Akram W, Gamiz J: **Image segmentation using region growing.** 1997 [http://www.owlnet.rice.edu/~elec539/Projects97/WDEKnow/index.html].
5. Lassouaoui N, Hammami L: **Genetic algoritms and multifractal segmentation of cervical images.** *Proc. Of. IEEE-EURASIP 7th International Symposium on Signal Processing and its Applications* 2003.
6. Garrido NP: **Applying deformable templates for cell image segmentation Pattern Recognition** *Pattern Recognition*. 2000, **33**:821-832.
7. Bamford P, Lovell B: **Method for Accurate Unsupervised Cell Nucleus Segmentation.** *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE* 2001.
8. Byriel J: **Neuro-fuzzy classification of cell in cervical smears.** *Master's thesis* Technical University of Denmark 1999.
9. Martin E: **Pap-smear classification.** *Master's thesis* Technical University of Denmark 2003.
10. Manjunath B, Ohm J, Vasudevan V, Yamada A: **Color and Texture Descriptors.** *IEEE Transactions on circuits and systems for video technology* 2001, **11**:703-712.

## S39

### E-education for medical Students using WSI in Egypt

Essam Ayad
Departament of Pathology. Faculty of Medicine, Cairo University & Italian Hospital in Cairo, Egypt
E-mail: essamayad@yahoo.com

*Diagnostic Pathology* 2013, **8(Suppl 1):**S39

**Introduction:** Classic practical Teaching of pathology for both under & post graduates requires microscopes and glass slides. Medical schools in developing countries are confronted with many challenges, among them, the large number of the students especially if this is compared with the limited resources of the medical schools. Furthermore, the high cost of acquisition and maintenance of microscopes. As a consequence, more than one student may have to share a microscope in the teaching laboratory (Figure 1).

Furthermore, glass slide specimens may also be shared among students since preparation of new specimens and replacement of broken slides are often expensive as well. Moreover, the time for laboratory hours is limited, and micro-scopes are often locked up and unavailable for use after class hours.

Although the medical schools in developed countries do not suffer a lot from the same problems but the reported trial for applying digital pathology techniques has shown to provide advantages of this digital technique over the usual method of teaching histology and pathology [1,2]. Few developing countries applied the digital pathology in the form of telepathology in clinical practice [3,4]. The use of digital pathology telepathology in education [5], second-opinion consultations [6] and primary diagnosis [7] has been reported in several journals. The early trials for applying digital pathology in medical education using the older techniques like the static & dynamic forms were not greatly satisfactory for the teachers or the students. Using the WSI technology changed the teaching circumstances completely. The entire microscopic glass slides were scanned and changed to digital files available on the Web server that can be viewed on a computer monitor with a Web browser. The dynamic and interactive mode of viewing images (horizontal and vertical movement of images, and zooming in and out) simulates viewing a glass slide under the microscope (virtual microscope). Its interactive features and facility to view the image on a large monitor promotes group interaction and discussion. Images can be viewed anytime with an Internet or Intranet connected desktop computer, portable or tablet computer, or even through smartphones anywhere. Its main challenge may be the requirement for high bandwidth. Implementing digital pathology in low-resource areas remains a challenge even today. Access to the Internet on academic networks is often slow and expensive. Other barriers include: the high cost of equipment for digitizing glass slides and limited student access to computer workstations especially after class hours. Very few trials tried to explore the student perspective of digital pathology especially in developing countries [8]. In this paper, we attempt to determine the student's attitudes in one medical school towards digital pathology.

**Methods:** Teaching glass slides were provided by the Department of Pathology, Cairo University. The whole set of slides were scanned using Bioimagene iScan 2, then the JP2 files gained were uploaded on the computer network in the pathology department computer lab and the Grand Student Library in the Faculty of Medicine, Cairo University.

**Figure 1(abstract S39)** showed more than one student may have to share a microscope in the classic teaching laboratory.

The images were configured for display on Apple devices (iPhone, iPod and iPad) using a secure PIN given for the students included in this pilot trial.

Thirty five medical students (male = 29, female = 6; age range, 20–22 years, mode = 21 years) were selected as one group of practical lessons. After verbal consent was obtained, they were instructed to view both classic practical pathology teaching course using the glass slides and another digital teaching using the virtual slides available in the Pathology department computer network and that uploaded on the Faculty severer. Figure (2).

The participants used a first generation iPad (Apple) tablet computer and get connected to the medical school's shared wireless network with average tested speeds of 120 Kilobits/sec download and 41 Kilobits/sec upload. The Website was configured for devices that did not require Adobe Flash player; furthermore the image browsing program was also uploaded so there was no need to download any additional programs to see the images. By the end of the digital pathology teaching course, the students were asked to complete a paper-based evaluation. They reported their acceptance of the trial and rated their experience based on their perception of ease of navigation, access speed, and preferences comparing the digitized images seen in the department computer network and that seen from Web server through the wireless access in the faculty library or at home using the internet accessibility in a scale (1-Strongly Disagree, 2-Disagree, 3-Somewhat agree, 4-agree, 5-Strongly Agree). Comments were encouraged but not required. No personally identifiable information was collected. The students finally passed two practical exams; a routine exam using the microscope glass slides & another digital exam using the computer lab in the department uploaded with the virtual slides Figure [3].

The grades of both exams were collected and statistically analyzed & compared. Furthermore, we compared both grades [classic & digital] of this pilot group with the results of the rest of the student who received only classic glass-slides practical teaching.

**Results:** More than 90% of the participating medical students rated the local computer network as significantly faster and responsive to their school needs. The participants also evidently preferred using the wireless connection in the library rather than reaching the website at home using the internet. Although not required, all of the students provided written comments to almost all of the questions posed. Their observations generally revealed that the images from the local server loaded faster and of the almost seamless quality in viewing the slides displayed. They also validated that virtual slides simulate the experience of examining glass slides under an optical microscope. Some general comments common to both servers were: "dealing with the pictures on the computer screen was much easier than the glass slides and the microscope", "interactive", "pathology studies became much easier & interesting", "wonderful quality pictures". The slides were labeled with the diagnosis and the description of the main diagnostic features but all these data could be seen when needed. Concerning the specific comments on the servers partially reflect the bandwidth deficiency that slowed the display of digital images from the server. They found the local server to be "fast loading & zooming in local server".

These results reflected that the understanding of the practical pathology material learned [by microscopes & classic glass slides] was evidently better & easier for student received digital teaching also as the students reported in their comments about the clarity of the images and how was it easy to navigate through the virtual slides.

**Discussion:** Although Internet users in Egypt are more than 35% of the population, the bandwidth in most places is still less than 2 Mb/s. In our study, the network in the medical school, which is situated at the computer network in the pathology department or the Grand Student Library in the Faculty of Medicine, Cairo University is rated at 8 Mb/s. Although the medical students found their experience to be generally satisfactory in this study, digital pathology implementation in developing countries may still be limited by slow connections to the Internet [8]. However, we believe that the slow connectivity or unavailability of Internet access will improve soon in most of the developing courtiers so it need not hamper the use of digital pathology in these countries. The increasing capacity and lowering costs of these portable storage devices


**Figure 2(abstract S39)** showed the Faculty severer website.

make them suitable alternatives to accessing virtual slides through the Internet. The use of digital slides in medical education makes the study of pathology a meaningful yet interesting endeavor for students [10]. In this study, we received positive comments that showed how students perceived the use of digital pathology in medical school. The students also suggested that the option to switch between labeled and unlabeled versions was useful and easier to navigate and reaching the diagnosis. This is a feature that has been suggested in the past through a previous trial [8]. In developed countries, local area networks are often assumed to be robust and adequately managed. However, it does not usually prevail in low-to-medium income countries where reliability is a continuing challenge. On top of the expense of network hardware, there is also the cost of hiring and retaining competent technical personnel to install, configure, and keep the local network running at optimal levels.

Although local access to the digital slides received better reviews from the students in this study, the hidden cost of keeping a reliable network should be considered when planning the appropriate placement of servers and wireless networks. Since the study was done in a single medical school in Egypt, we hope that it can be universally applicable. However, since many of the conditions mentioned earlier prevail in developing countries, our study highly encourage the suggestions on the implementation of digital pathology in low-resource areas. Estimation of the minimum acceptable bandwidth for viewing virtual slides needs further studies. The cost of scanning slides is still prohibitive in most developing countries. Collaboration between academic centers in developed countries with digital scanning equipment and universities in developing countries with teaching slides that may be rare and unusual in developed countries, could enhance medical education for all [8]. Ultimately, it may be cost-effective for a university to send their teaching slides to partner institutions by mail or courier who can digitize them for sharing across participating schools.

**Conclusion:** Our results show that access to the server either through the wireless connection in the computer network in the pathology department or the Grand Student Library in the Faculty of Medicine, Cairo University or through the internet at home was satisfactory but the local server was deemed faster and preferred by a majority of participants in this study. Virtual slides, accessible through a local server or portable drive may be a solution to the high bandwidth requirement of digital pathology or in places where the Internet in unavailable. Collaborations with universities in developed countries would enhance

**Figure 3(abstract S39)** showed a participating student in the digital exam.

**Table 1(abstract S39) showed the comparison between the student's grades in both classic & digital practical exams. The results indicated the better understanding of the practically taught material through the digital form rather the classic one**

| | Classic | Digital |
|---|---|---|
| Number of students | 35 | 35 |
| Range [out of 10] | 3-10 | 6-10 |
| Mean | 8.1 | 8.9 |
| Percentage | 81 | 89 |

**Table 2(abstract S39) showed the comparison between the trial group grades [classic & digital] and the grades of the rest of the student who received only classic glass-slides practical teaching**

| | Trial Group | Rest of Students |
|---|---|---|
| Number of students | 35 | 140 |
| Range [out of 10] | 3-10 | 3-10 |
| Mean | 8.1 | 7.33 |
| Percentage | 81 | 73.3 |

image collections for teaching for institutions and these can be shared with others.

**Competing interests:** The author declares that they have no competing interests.

**References**

1. Dee F, Heidger P: **Virtual slides for teaching histology and pathology.** *Virtual Microscopy and Virtual Slides in Teaching, Diagnosis, and Research* Taylor & Francis Group, CRC Press, Boca Raton 2005, 151-159.
2. Bloodgood RA, Ogilvie RW: **Trends in histology laboratory teaching in United States medical schools.** *Anat Rec B New Anat* 2006, **289**:169-175.
3. Brauchli K, Jagilly R, Oberli H, Kunze KD, Phillips G, Hurwitz N, *et al*: **Telepathology on the Solomon Islands–two years' experience with a hybrid Web- and email-based telepathology system.** *J Telemed Telecare* 2004, **10(Suppl 1)**:14-17.
4. Pagni F, Bono F, Di Bella C, Faravelli A, Cappellini A: **Virtual surgical pathology in underdeveloped countries: The zambia project.** *Archives of Pathology & Laboratory Medicine* 2011, **135**:215-219.
5. Weinstein RS: **The education of professionals.** *Hum Pathol* 2003, **34**:415-416.
6. Molnar B, Berczi L, Diczhazy C, Tagscherer A, Varga SV, Szende B, *et al*: **Digital slide and virtual microscopy based routine and telepathology evaluation of routine gastrointestinal biopsy specimens.** *J Clin Pathol* 2003, **56**:433-438.
7. Evans AJ, Chetty R, Clarke BA, Croul S, Ghazar-ian DM, Kiehl T-R, *et al*: **Primary frozen section diagnosis by robotic microscopy and virtual slide telepathology: The Uni-versity Health Network experience.** *Hum Pathol* 2009, **40**:1070-1081.
8. Fontelo Paul, Faustorilla John, Gavino Alex, Marcelo Alvin: **Digital pathology – implementation challenges in low-resource countries.** *Analytical Cellular Pathology* 2012, **35**:31-36.
9. Fontelo P, DiNino E, Johansen K, Khan A, Ackerman M: **Virtual microscopy: Potential applications in medical educa-tion and telemedicine in**

countries with developing. *System Sciences, 2005. HICSS 05. Proceedings of the 38th Annual Hawaii International Conference on (2005)* 153c.

10. Kumar RK, Freeman B, Velan GM, De Permentier PJ: **Integrating histology and histopathology teaching in practical classes using virtual slides.** *Anat Rec B New Anat* 2006, **289**:128-133.

## S40

### Towards the integration of digital cytology in the tablet technologies

Daniele Giansanti[1*], Marco Pochini[2], Enrico Giarnieri[2], Maria Rosaria Giovagnoli[2]
[1]Istituto Superiore di Sanità, Roma, Italy; [2]Università Sapienza, Facoltà di Medicina e Psicologia, Roma, Italy
E-mail: daniele.giansanti@iss.it
*Diagnostic Pathology* 2013, **8(Suppl 1):**S40

**Background:** The today's technologies allow an easy way to share information relevant to image by means of different tools. Many solutions are today available that allow to the cytologist to exchange digitalized information about glasses. It is thus arising the need of objective methodologies, such as the ones based on HTA that we have proposed in a previous study focused in tele-pathology [1], to investigate the potentialities of the new technologies in D-CYT. It should be also considered that the D-CYT will have a great impact in the work organization as the interaction with the glasses is radically changing [2] giving also new chances to e-learning [3-5].
In a typical today's architecture there is
a) an *on-site-server* with the scanner or alternatively a tele-pathology *third-party-server* (in site or in remote as a Web service such as the Leeds' centre [6] at http://www.virtualpathology.leeds.ac.uk/index.php), for the creation of the *virtual glasses* named Digital Slides (DS)s,
and
b) low cost (or free) *Light-client applications* proprietary software tools (PC-client application not asking enough hard disk space, not consuming dynamically PC resources, such as RAM, SWAP ecc., simple and user-friendly) for the navigation on the DS which can be installed in remote clients.
This new methodology is rapidly largely spreading and it is becoming the core aspect of the formation for future qualified personnel and is more and more asking for user-friendly and effective ICT solutions.
The *Tablet-technology-ICT-solution* is recently widely increased as a user-friendly and effective tool to remotely share image information. Thanks to this technology it is possible to navigate into an image (pan, zoom-in, zoom-out) using only the hand-fingers without typing the keyboard or the software interfaces' keys. This way of image navigation is going to further improve the application of Telemedicine to DP and in particular in D-CYT with particular reference to the remote and/or cooperative decision and diagnosis in cytology.
**Material and methods:** The methodological flow clearly arises by inviting the reader to navigate in the Public WEB of the University of Leeds named "Virtual Pathology at the University of Leeds" where several digital-slides are available for public use at the URL: http://www.virtualpathology.leeds.ac.uk/public/common_slides.php
The reader can try to navigate the DS by means of a PC using the interface of Spectrum Web Viewer (by Aperio) or by means of a Tablet system using the fingers and *subjectively* consider thus the differences. It is thus clear that there is thus strongly the need of considering both the state of art of the Tablet technologies and the design of an *objective methodology* to assess the technology. The methodological flow faced thus the two basic issues:
• Analysis of the state of art of the tablet technologies.
• Investigation of HTA solutions [1] to assess the technology (Tablet and applications of D-CYT) both in terms of performances and acceptance for the relevant introduction in Telemedicine [6-8], [9,10].
**Results and discussion: Analysis of the state of art of the tablet technologies:** The analysis with the focus to D-CYT returned that the tablet technologies could be grouped into: *wearable* tablets, *portable* tablets and *not-portable* tablets.
- The *wearable* tablets comprehend the Smart-phones i.e the devices that can be embedded in a pocket (Figure 1 ).
- The *portable* tablets comprehend the *A4 A4/2* tablets such as the Apple Ipad i.e the devices that can be embedded in a 24-hours-suitcase (Figure 2).

**Figure 1(abstract S40)** Navigation using the WEB-Scope by Aperio using a wearable tablet.

- The *not-portable* tablets comprehend the very large touch tablets such as the XDesk or Microsoft Surfaces, i.e the devices that cannot be self carried.
The first two systems, that are widely used for many different purposes allow to reach everyone in the world, therefore represent a chance for the remote consulting in D-CYT (Figure 1,2).
The last systems, focusing for example on the Epson XDesk, represent a powerful ICT solution for cooperative analysis and discussion of cases of virtual cytology.
In details, the Epson XDesk is an interactive table; some call it a coffee table because you can put anything on the surface of the table, it works by projections, with the very latest technologies on that. This desk is also compatible with Bluetooth communication protocol and as soon as you put your phone or camera on the surface of the table, the XDesk will be able to see all your files and pictures on the desk. By natural interface pictures on the table can be managed freely and resized, zoomed in and out by finger movements as the iPod touch does, only on a fair larger scale. The Epson XDesk has a 52-inch screen and a 1024x768 touch screen display. It represents the appropriate high technology solution for cooperative discussions (Figure 3,4,5), clinical audit and ultimately the future direction of cooperative virtual microscopy environment. Furthermore it could represent a tool suitable to recover the inheritance of some ICT solutions (such as Pap-Net) for large screening in cytology

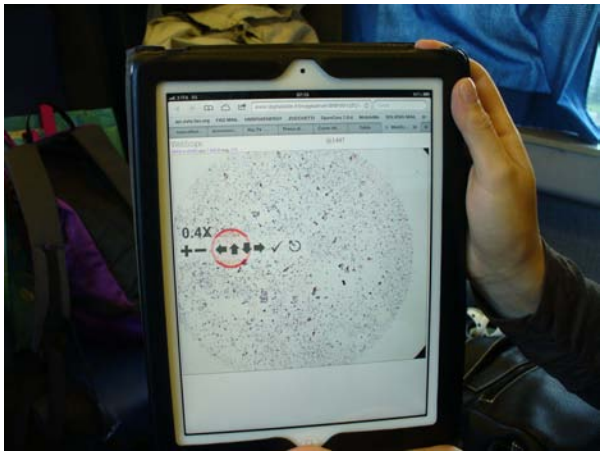



**Figure 2(abstract S40)** Navigation using the WEB-Scope by Aperio using a portable tablet.

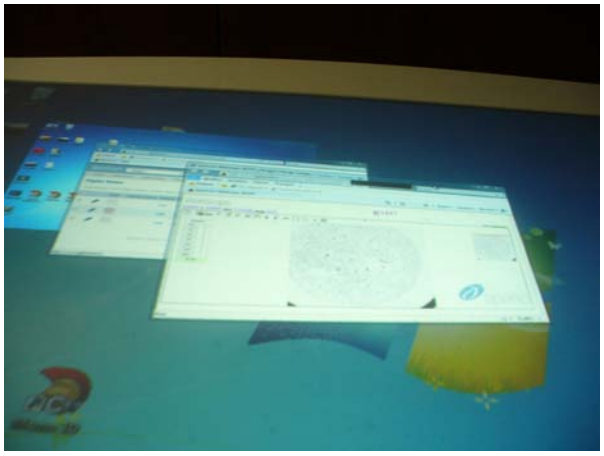


**Figure 3(abstract S40)** Cooperative scenario in D-CYT with the Xdesk : The Windows.

**Figure 4(abstract S40)** Cooperative scenario in D-CYT with the Xdesk : The virtual-slide.

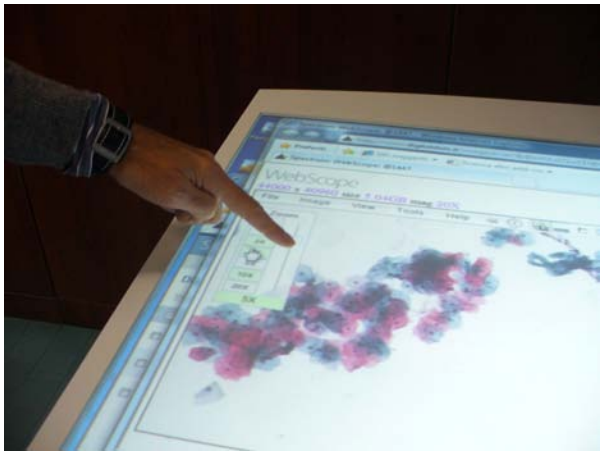


**Figure 5(abstract S40)** Cooperative scenario in D-CYT with the Xdesk : The navigation based on the fingers.

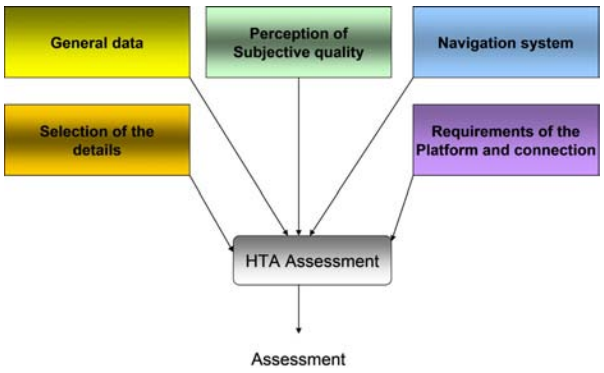



**Figure 6(abstract S40)** Key points investigated by means of the HTA tool.

abandoned because in the first applications the technology was not ready, for computer assisted cervical/vaginal cytology diagnosis [7].

**Investigation of HTA tools to assess the technology:** As a core aspect is the introduction of the new technologies for the D-CYT in the Hospital, specific studies of Health technology assessment should be performed focused in the new technologies for the digital-cytology. This is a basic aspect for the Health Care Systems.

As a HTA tool we recovered the interactive word based tool we have proposed in [1] for applications in tele-pathology. This tool had been proposed before the diffusion of the tablet technologies and was thus conceived for the assessment of *PC-based-technologies* for D-CYT. However the tool is portable, as it is, on *tablet based technologies*. Figure 6 elucidates the contents of the HTA tool, each element contains a number of questions with a scoring based on 4 levels. Starting from this tool the investigation will be widened considering also:

1. Further issues relevant to studies of HTA, as from specific experience of mondial networks of HTA such as the EUNETHTA (http://www.eunethta.net/).
2. Integration of specific studies on HTA over the NET, as the focus is the D-CYT is the communication over the WAN/LAN.

**Conclusions:** Tablet technologies have been reviewed with the focus to the perspectives of the D-CYT. A HTA specific tool has been proposed to assess the performances and acceptance of the applications in D-CYT based on tablet technologies. Possibilities and limitations of the three different tablet technologies will be deeply investigated by means of the proposed HTA methodology on experts and students approaching the new scenario of digital-cytology.

**List of abbreviations used:** HTA: Health Technology Assessment; DP: Digital Pathology; D-CYT: Digital Cytology; DS: Digital Slide

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** DG MRG wrote the manuscript as major contributors. MP has prepared the review on the wearable and portable and not portable Tablet technologies and is currently caring the experimentation study. EG has reviewed and improved the scientific content and rationale of the manuscript. All authors have read and approved the final manuscript.

**References**

1. Giansanti D, Castrichella L, Giovagnoli : **The design of a health technology assessment system in telepathology.** *Telemed J E Health* 2008, **14(6)**:570-5.
2. Giovagnoli MR, Giarnieri E, Carico E, Giansanti D: **How do young and senior cytopathologists interact with digital cytology?** *Ann Ist Super Sanita* 2010, **46(2)**:123-9.
3. Giansanti D, Castrichella L, Giovagnoli MR: **New models of e-learning for healthcare professionals: a training course for biomedical laboratori technicians.** *J Telemed Telecare* 2007, **13(7)**:374-6.
4. Giansanti D, Castrichella L, Giovagnoli MR: **Telepathology training in a master of cytology degree course.** *J Telemed Telecare* 2008, **14(7)**:338-41.
5. Giansanti D, Castrichella L, Giovagnoli MR: **Telepathology requires specific training for the technician in the biomedical laboratory.** *Telemed J E Health* 2008, **14(8)**:801-7.
6. Giansanti D, Grigioni M, Giovagnoli MR: **Virtual microscopy and digital cytology: fact or fantasy? Preface.** *Ann Ist Super Sanita* 2010, **46(2)**:113-4.

7. Giansanti D, Grigioni M, D'Avenio G, Morelli S, Maccioni G, Bondi A, Giovagnoli MR: **Virtual microscopy and digital cytology: state of the art.** *Ann Ist Super Sanita* 2010, **46(2)**:115-22.
8. Giansanti D, Cerroni F, Amodeo R, Filoni M, Giovagnoli MR: **A pilot study for the integration of cytometry reports in digital cytology telemedicine applications.** *Ann Ist Super Sanita* 2010, **46(2)**:138-43.
9. Giansanti D, Morelli S, Macellari V: **Telemedicine technology assessment part II: tools for a quality control system.** *Telemed J E Health* 2007, **13(2)**:130-40.
10. Giansanti D, Morelli S, Macellari V: **Telemedicine technology assessment part I: setup and validation of a quality control system.** *Telemed J E Health* 2007, **13(2)**:118.

## S41

### Enhanced imaging workflow for anatomic pathology: implementation of a knowledge actor

Jacques Klossa[1*], Thomas Schrader[2], Christel Daniel[3]
[1]TRIBVN, France; [2]University of Applied Sciences Brandenburg, Germany; [3]Assistance-Publique Hôpitaux de Paris, France
E-mail: jklossa@tribvn.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S41

**Introduction:** In Pathology, the development of slide scanners offers the possibility to acquire whole slide images (WSI) from specimen stored on glass slides and this is the beginning of a new paradigm in this medical domain. Two DICOM supplements were defined by the DICOM WG26 in order to store and display such big amount of data. These WSI images are mainly used for teaching, for research and for collaborative and cooperative applications. However, automated integration in the hospital healthcare enterprise (IHE) needs much more detailed information for i) integrating the digitalization process into the general diagnostic process in a Pathology Department, in other words, to define the WSI "acquisition modality work list", and ii) integrating efficiently image analysis into the diagnostic process, e.g. to create "evidences" from a WSI.
**Methods:** A pathologist who reads a slide uses an iterative process which includes exploration and analysis phases, each at varied magnification. In comparison, a slide scanner acquires the whole slide or a pre-defined region of interest. In practice, we know that medical questioning may require more sophisticated possibilities like combining low RGB resolution scan (e.g. 5x to 20x) with higher resolutions and multi-z imaging in some specific region which could be addressed by some slide scanners provided one can build the corresponding modality (ies) work list. The same issues happen for "evidence creation" through an image analysis workflow. Workflow can become even more complex when using a multimodal scanner which adds to the current modalities (RGB bright field and fluorescence) complementary ones like quantitative phase imaging for label free unstained samples and Raman micro-spectroscopy for molecular signature acquisition.
We have 2 choices for solving such issues: either the process is embedded in the scanner software inside a proprietary solution, or the process can be driven by a scenario following the medical needs and assigns to each modality unitary task to be executed.
**Results:** Let us try to scenarize the fairly simple use case of "Malaria". Following the current guidelines, diagnostics and personalized care aims at parasitemia counting and species identification on red blood cells (RBC): step 1 of the workflow asks for the acquisition of 200 fields at 100x and step 2 asks for individual classification of the retrieved infected RB cells. Such description is medically oriented and needs to be traduced in unitary technical tasks. Clearly, there is a need for some additional knowledge that explains how to select the fields to acquire, how to identify infected RBC and what information should be acquired on each RBC: color, MS images with or without z stack.
**Discussion:** We are in favor of implementing in the workflow the second solution that avoids proprietary solutions and which could be achieved by adding to the IHE workflow a knowledge actor (KA) that will produce and orchestrate the scenario. KA would ask to the order filler patient and case information's in addition to the diagnostic questions. Based on the corresponding guidelines, the KA could then orchestrate a sequence of unitary work lists which can be executed by the microscopy scanner in conjunction with the evidence creator actor. Such solution would clearly separate medical guidelines from modality embedded software making the workflow more scalable.
Trying to benefit from previous IHE integration, we would suggest implementing such an actor in a way similar to the Treatment Management System (TMS), an information system that manages oncology information and is responsible for the scheduling of radiotherapy activities.
**Competing interests:** The authors declare that they have no competing interests
**Authors' contributions:** • CD, TS and JK have developed the need for a new actor in the IHE workflow to answer the issue associated with modality driving and evidence creation.

## S42

### Quantitative phase imaging and Raman micro-spectroscopy applied to Malaria

Jacques Klossa[1*†], Benoit Wattelier[2†], Teddy Happillon[3†], Dominique Toubas[3,4†], Lucie de Laulanie[2†], Valérie Untereiner[3†], Pierre Bon[2†], Michel Manfait[1†]
[1]TRIBVN, 39, rue Louveau, 92320 Châtillon, France; [2]Phasics, Campus de l'Ecole Polytechnique, 91128 Palaiseau, France; [3]MEDyC FRE/CNRS 3481, 51096 Reims, France; [4]CHU de Reims, Laboratoire de parasitologie-mycologie, 51100 Reims, France
E-mail: jklossa@tribvn.com
*Diagnostic Pathology* 2013, **8(Suppl 1):**S42

**Background:** Malaria is due to parasitism of red blood cells (RBC) by protozoan parasites of the genus Plasmodium. Three main parameters have to be determined for patient treatment: parasite species, the rate of infected blood cells (parasitemia), and development stage. Even if a series of laboratory techniques are available, a suited treatment needs microscopy skills [1]. Microscopic observation needs a specialist and is time consuming (e.g. observation of hundreds fields of view at 100x immersion objective) and automating 100x slide scanning of white light imaging of thin film stained blood smears is not straightforward.
Seeking for an easy to automatize alternative, we thought that combining two microscopy techniques: quantitative phase imaging for quick detection and Raman micro-spectroscopy for molecular characterization could appear to be an efficient multipurpose solution.
Raman micro-spectroscopy is a good candidate to identify molecular species in microscopic. Laser light is focused in a tiny volume. Due to molecular vibrations, part of the light is non-linearly scattered to longer (Stokes) or shorter (Anti-Stokes) wavelength. The wavelength shift is directly linked to the vibration energy. The scattered light spectrum contains lines typical of the molecular binding in the focused volume. However imaging is very time consuming as each pixel needs to be acquired individually (between 1 and 10s per pixel).
Quantitative phase imaging is an imaging technique that measures the optical path difference of light travelling through different part of a semi-transparent medium. If a biological tissue has local different index of refraction, we obtain an image that reflects these index changes. This kind of maps usually has a contrast more than two orders of magnitude higher than bright field microscopic images. In addition, they give information about the relative index change in the tissue.
Combining a morphological technique suited for rare events detection with a molecular technique suited for local molecular signature acquisition could provide an easy way to address automation of parasite parameters acquisition.
In the following sections, we present our first level attempt for proof of concept.
**Material and methods: Sample preparation and hardware:** We used for our experiments the solution developed during the IHMO program (see "Label free technologies: Raman micro-spectroscopy and multi-spectral imaging for Lymphocyte classification" in the current publication). It uses the Calopix TRIBVN$^{TM}$ software platform integrated to the hospital workflow and off the shelves components NIKON FN1$^{TM}$ microscope fitted with a Raman micro-spectroscopy (RMS) HORIBA module whose excitation source is a 532nm diode laser. In addition a commercial quadriwave lateral shearing interferometer (QWLSI Phasics SID4Bio$^{TM}$) is plugged on one of the microscope exit port; it measures the quantitative phase shift in the visible spectral band. This is an easy-to-integrate and compact solution (dimension of a simple camera) and

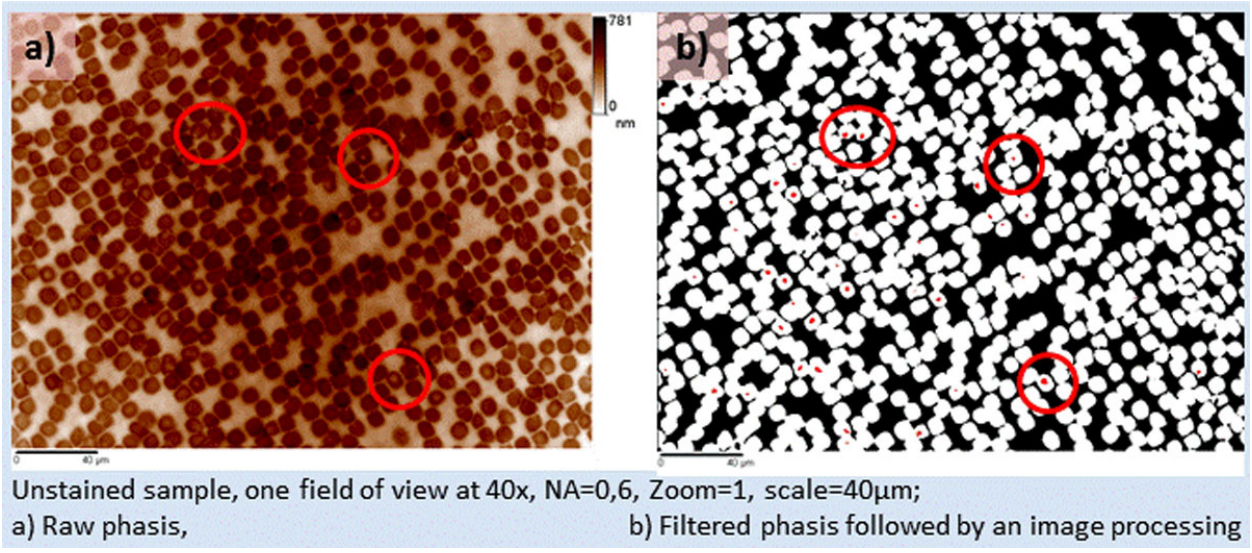



**Figure 1(abstract S42)** parasite characterization: a) microscope image, b)quantitative phase image, c) RMS spectra comparison between infected RBC and healthy RBC after hemoglobin mean spectrum subtraction.

gives a 300x400px$^2$ (px=pixel) phase and intensity image with a lateral pitch of p=29.6 µm in the image plane, which corresponds to 0.74µm/px for 40x magnification and 0.197µm/px for 150x magnification.

Blood smears are prepared on a standard glass slide with no slide cover and without any previous staining to avoid contamination of the recorded Raman spectra with fluorescence contribution. Spectra are acquired on infected RBC and finally, slides are stained with May-Grünwald Giemsa for morphological validation.

**Quantitative phase imaging:** The presence of the parasite inside the RBC implies a synthesis of hemozoin (complex molecular structure) from hemoglobin. These two molecules present a different refractive index; therefore the parasite should appear with a different quantitative phase inside the RBC. We have checked this assumption at high magnification (150x) and we found out that the parasite induces lower phase values (figure 1). Afterward, we studied large fields of view at lower magnification: 40x proved to be efficient: see a typical field of view figure 2. In order to automatically detect the infected cells, we first applied a high pass filter. This removes the slowly varying cell shape while keeping the plasmodium image. Then low phase values due to the plasmodium were detected by three-level segmentation. One level corresponds to the extracellular medium; the second level corresponds to the RBC cytoplasm containing hemoglobin and the third one to the hemozoin. The second level segmentation leads to the number of the RBC contained in the field. The third level segmentation gives the number of plasmodia and thus of infested RBCs. The ratio between these two numbers is the parasitemia. We have studied 10 fields of a single blood smear containing falciparum parasites.

**Raman micro-spectroscopy:** Raman spectra are recorded on parasite from infected RBC localized thank to QPI at 40x. Raw spectra need first a pre-treatment phase to make them eligible in the classification process. Indeed, the presence of hemoglobin into the cells, and more precisely the fact that hemoglobin contains cyclic structures, implies distortions of the baseline of each spectrum. This distortion effect is called a fluorescence background. To correct the spectra, a function based on a polynomial estimation and correction of the baseline is used for each spectrum independently of each other. Then, the normalization function, standard normal variate, is used on each spectrum. This function eliminates the variation of the absolute values into the spectra, and makes them comparable avoiding scale differences. Then, a first classification could be realized. For such purpose, we used the classical Hierarchical Clustering Analysis (HCA). HCA is based on the Euclidean distance between each observation (each spectrum) and from these distances a hierarchical binary tree is created, putting forward the different groups of spectra. Representative peaks of the hemozoin biocrystal, previously identified in literature, are taken into account to classify the spectra, by reducing them to the corresponding region of interest. In a second classification, and to improve the first results, a representative spectrum of pure hemoglobin, which has been obtained by averaging RBC spectra from other samples, has then been estimated and subtracted from each recorded spectrum, thanks to a mean squared based function.

**Results and discussion: Quantitative phase imaging results:** QPI analytical process detected an average of 640 blood cells per field at 40x (standard deviation of 23) and an average of 4.1% of infected cells (standard deviation is 1.1%) for sample labeled as a 3.5% parasitemia.

**Raman micro-spectroscopy results:** The first classification realized considered the pre-treated spectra reduced to the representative peaks of the hemozoin biocrystal (i.e. 1380-1385, 1594-1597 and 1644-1646 cm-1). After a HCA classification the full 35 spectra obtained from sane red blood cells were classified in the first cluster of the HCA dendrogram, but only 20 spectra on the 34 registered on parasites were present in the second cluster, which gave a sensitivity of 100% and a specificity of 58.8%.

To improve these first results, a second classification was realized, considering every spectrum without the contribution of the hemoglobin signal as mentioned previously. Then, the full 35 spectra from the parasites where present in the first class of the dendrogram and the full 34 spectra from sane red blood cells were classified into the second cluster, giving a sensitivity and a specificity of 100%.

**Discussion:** This study proved the ability of QPI to detect RBC and infected RBC at low magnification without any previous staining. Each detected infected RBC can then be easily confirmed with RMS spectra acquisition. Making such global process easy to standardize and easy to automate.

The proposed concept is very efficient on this practical use case thanks to QPI efficiency that allows quick retrieving of a representative cell population without any previous staining. However during further studies, detection potential of poorly differentiated parasite should be compared to 100x classical immersion microscopy on stained sample.

RMS highlights thin molecular information and has been 100% sensible and specific in the current study. However, the presence of hemozoin into the parasite could be variable in quantity or even nonexistent which in further studies could lead to spectra misclassifications. Finally, it will also be necessary to differentiate species: this should be assessed on following studies while using different high resolution (150x) techniques on infected cells: i) RMS spectra, ii) optical

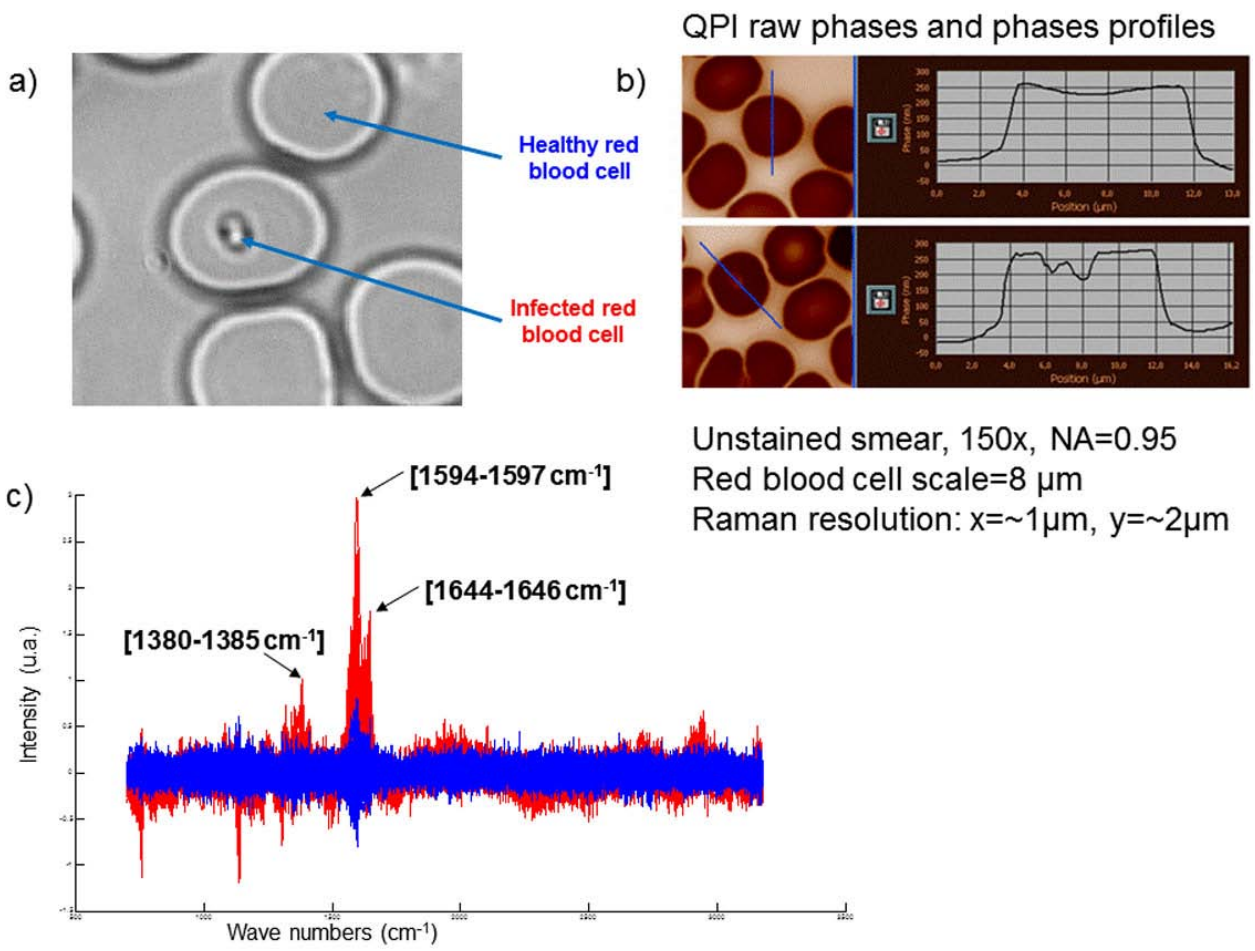



**Figure 2(abstract S42)** blood smear exploration and parasite detection with quantitative phase imaging.

volume measure with QPI and iii) multi-spectral classification (see IHMO project).

Following further investigations, combination of quick detection with Quantitative Phase Imaging and selective molecular imaging with Raman Micro-Spectroscopy has a good potential to automate parasite detection and counting. Such toolkit could be applied to any kind of parasite, bacteria or bacteria inside vesicles and it also could be extended to other cytology use cases and likely to the field of histopathology.

**List of abbreviations:** HCA: ; QPI: Quantitative phase imaging; Px: pixel; QPI: Quantitative phase imaging; QWLSI: ; RBC: Red blood cell; RMS: Raman micro-spectroscopy

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** • TH carried out the acquisitions, pretreatments, the highlighting of relevant information into the spectra, the classification of the data and drafted the extended abstract.

• VU realized the acquisitions of the Raman spectra and drafted the extended abstract.

• PB, LdL and BW developed the multi-spectral classification method and produced the multi-spectral classification.

• JK, BW and MM both designed and managed the study.

• DT selected the infected patients, provided the corresponding samples and patient's data, and did microscopy assessments.

**Reference**

1. Noppadon Tangpukdee, Chatnapa Duangdee, Polrat Wilairatana, Srivicha Krudsood: **"Diagnosis: A Brief Review".** *Korean J Parasitol* 2009, **47(2)**:93-102.

## S43

### Impact of tumor heterogeneity on disease-free survival in a series of 368 patients treated for a breast cancer

Myriam Oger[1*], Mohamed Allaoui[2], Nicolas Elie[3], Jacques Marnay[2], Paulette Herlin[1], Benoît Plancoulaine[1], Jacques Chasle[1], Véronique Becette[1], Catherine Bor-Angelier[1,2]

[1]Imagin' Team of EA 4656 in François Baclesse Cancer Centre, 3 avenue du Général Harris, 14076 Caen, France; [2]Pathology laboratory, François Baclesse Cancer Centre, 14076 Caen, France; [3]Plateau d'Histo-Imagerie Quantitative, SF ICORE, CMABIO, University of Caen Basse-Normandie, 14076 Caen, France
E-mail: myriam.oger@gmail.com

*Diagnostic Pathology* 2013, **8(Suppl 1):**S43

**Introduction:** Tumor heterogeneity [1-4] is an old concept but its impact on the cancerogenesis process is poorly understood. Breast cancer is a noteworthy model for its frequency, and for the diversity of its phenotypes and of its evolution. This study examines the influence of the heterogeneity of tumor proliferation on disease-free survival of patients with a breast carcinoma.

**Material and methods: Histological slides:** The study involved a series of 368 patients from the François Baclesse Cancer Centre (Caen) treated for a breast carcinoma between 1991 and 1995, whitout neoadjuvant therapy and with a follow-up of more than 15 years. The table 1 contains the description of the series.

Histological sections, representative of each tumor, have been stained with the anti-phosphohistone-H3 antibody (PHH3: Ser10, MILLIPORE®,

**Table 1(abstract S43) Univariate Analysis of Disease Free Survival – 368 Eligible Patients. In grey: Follow up (2011)**

| Variable | No. of patients | *P* value |
|---|---|---|
| Age | | |
| Mean = 58.7 yr | | 0.400 |
| Menopauses | | |
| Yes | 237 (65%) | |
| No | 124 (34%) | |
| Localization | | |
| Right breast | 176 (47.8%) | |
| Left breast | 187 (50.8%) | |
| Synchronous bilateral | 5 (01.4%) | |
| Tumor size | | |
| Mean = 25 mm | | 0.010 |
| Surgery | | |
| Tumorectomy | 212 (57.6%) | |
| Mastectomy | 141 (38.3%) | |
| Biopsy | 15 (04.1%) | |
| Excision quality | | |
| Satisfying | 202 (57.2%) | |
| Unsatisfying | 47 (13.3%) | |
| Unspecified | 104 (29.5%) | |
| Histological type | | |
| ICC | 297 (80.71%) | |
| ILC | 37 (10.05%) | |
| SBR grade | | |
| Grade 1 | 50 (14%) | |
| Grade 2 | 180 (49%) | 0.002 |
| Grade 3 | 137 (37%) | |
| Mitotic index (/1.7mm ) | | |
| Mean = 10 mitosis | | <0.0001 |
| Tumor vascular emboli | | |
| Yes | 293 (79.6%) | |
| No | 75 (20.4%) | |
| Lymph node metastasis | | |
| Yes | 153 (44%) | <0.0001 |
| No | 195 (56%) | |
| Hormone receptor status (at least 1) | | |
| Yes | 265 (73%) | 0.030 |
| No | 98 (27%) | |
| Visceral or lymph node (other than axillary) metastasis | | |
| Yes | 136 (37%) | |
| No | 232 (63%) | |
| Local recurrence | | |
| Yes | 66 (18%) | |
| No | 302 (82%) | |
| Oncological event (Metastasis and/or local recurrence) | | |
| Yes | 157 (42.5%) | |
| No | 211 (57.5%) | |
| Death | | |
| Yes | 90 (24.46%) | |
| No | 278 (75.54%) | |

dilution 1/600) [5,6]. With this specific immuno-stain, cells presenting mitosis figures are more easily identifiable (Figure 1).

**Acquisition:** Histological slides have been scanned with a high resolution slide scanner to obtain virtual slides with a final resolution of 0.5 µm (ScanScope® CS from Aperio Technologies (20x NA 0.7 objective). The true color images obtained (color RGB 24 bits) have been saved in the tiled pyramidal TIFF file format.

**Region of interest (ROI):** Before the automatic image analysis, the user can discard "normal" tissue surrounding the tumor by drawing a region of interest on the high resolution virtual slide with the Aperio ImageScope® software.

**Image processing:** The image processing was performed in two steps on a personal computer with a 1.6 GHz Pentium IV processor and a 1 GB of random access memory (RAM). The first step being a sub-sampling of virtual slide done with a specific algorithm 'Daubechies' second moment orthogonal wavelet decimation developed in C++ language which creates a low resolution image of the virtual slide (divided by 8: from 0.5µm to 4µm/pixels). In a second step, the low resolution image is automatically processed thanks to chaining operators of image analysis toolbox software (Aphelion, ADCIS).

In addition to estimating the frequency of mitotic figures, the program detects "hot spots" and measures 9 features representing the tumor heterogeneity, including the Haralick texture features and Fisher's index. The zones of influence of each stained nuclei have been determined using Voronoï's pavement principle. When nuclei are close, the size of pavements is small, highlighting the "hot spots".

**Feature selection:** A principal component analysis has been done in order to select the most relevant features.

**Statistic analysis:** These features have been statistically analyzed, combined with classic clinic-pathological prognostic factors (age, tumor size, grading, mitotic index, vascular emboli and metastatic lymph nodes).

**Results: Principal component analysis:** Thanks to the principal component analysis (PCA) 4 features representing tumor heterogeneity have been chosen then combined into three new features: CP1, CP2 and CP3, corresponding to the three principal directions of the PCA.

The four selected features are:

- 2 Haralick's texture indexes (correlation and energy);
- Fisher's index;
- variance of the size of Voronoï pavements.

The variance of the size of Voronoï pavement (named Voronoï) and the Fisher's index are regional features whereas the Haralick's texture indexes are local features. Indeed, Voronoï and Fisher features are "cutting" the tissue into pieces and analyzing each of them compared to the others, whereas Haralick is dealing with relations between neighbor pixels, each pixel representing a cell at this resolution.

**Prognostic study:** In the analysis of prognostic factors, disease free survival was used as the end point.

**Univariate statistical analysis (DFS):** Univariate analysis of disease free survival was performed with the features of age, tumor location, initial tumor size, pathologic lymph node status (N), histological type, SBR grade, mitotic index, vascular emboli, metastatic lymph nodes and hormone receptor status. The results are shown in Table 1 for usual features, in Table 2 for heterogeneity features.

The CP2 feature correlated highly with disease free survival, whereas the variance of the Voronoï pavements was borderline significant.

**Multivariate statistical analysis (Cox):** The above features that correlated with disease free survival in univariate analysis were combined with clinic-pathologic factors and included in the multivariate analysis. Cox's regression analysis highlighted 3 independent prognostic factors: tumor heterogeneity feature CP2 (RR = 1.46; p = 0.03), mitotic index (RR = 1.71; p = 0.004) and lymph node metastasis (RR = 2.20, p < 0.0001) correlated highly with disease free survival.

The construction of this model has individualized 3 groups of patients: 0 factor, 1 or 2 factors and 3 poor prognostic factors (mitotic index > 10, lymph node metastasis in the axillary dissection, upper tercile of CP2; p < 0.0001).

Disease free survival according to this model is shown in Figure 2.

**Discussion and conclusion:** To characterize tumor heterogeneity in the presented series of breast cancer, 9 features were computed. 4 nonredundant of them have been selected by principal component analysis (PCA).

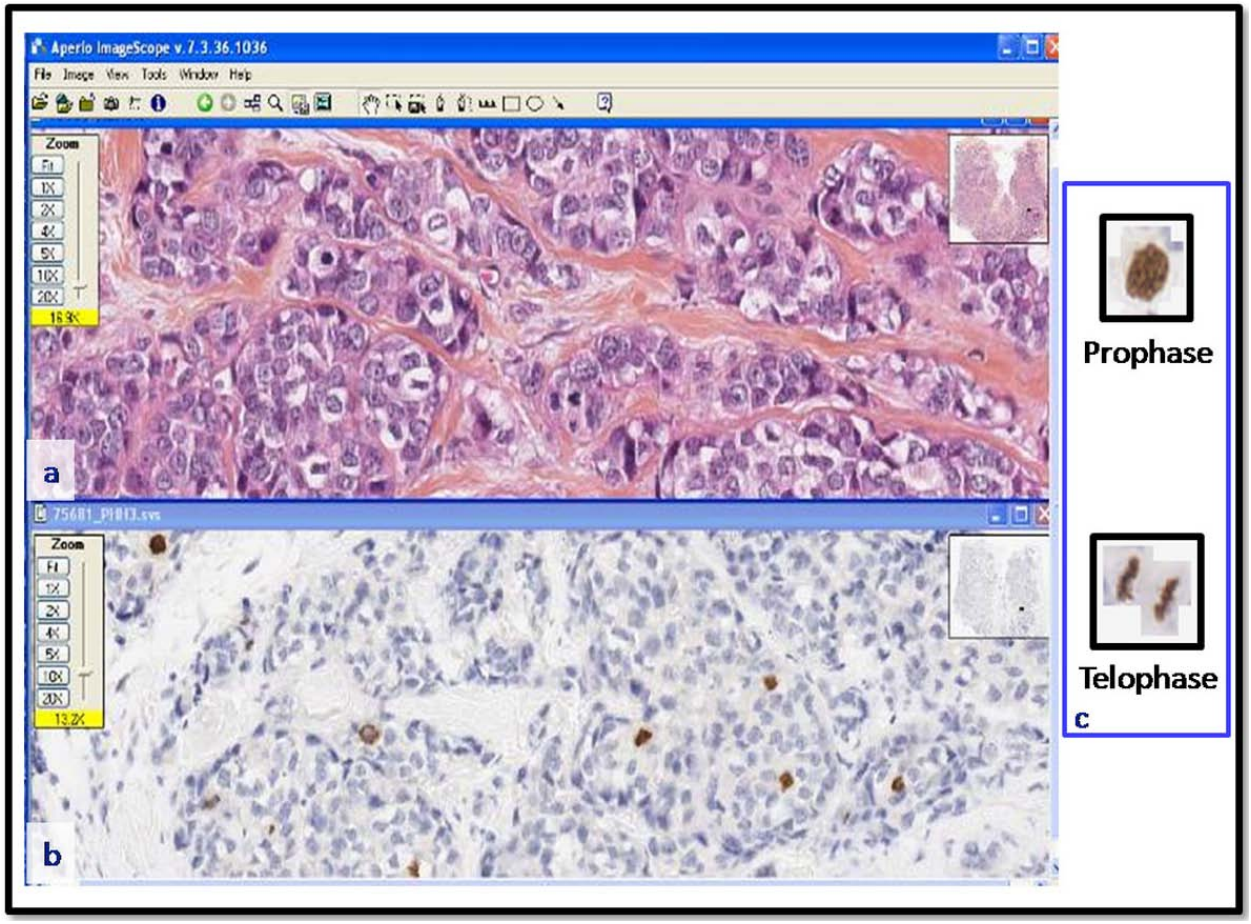



**Figure 1(abstract S43) dyes used to stain the histological sections. a**: histological slide stained with HES; **b**: histological slide of the same case immunohistochemically stained with PHH3 [6-8] (anti-phosphohistone-H3 antibody); **c**: thanks to this specific immunohistochemical, identification of cells with mitosis figures, from prophase to telophase, is improved.

PCA was also used to create 3 new composite features: CP1, CP2 and CP3, corresponding to the 3 principal directions of the PCA.

The univariate analysis made for each feature from image analysis has first highlighted that only the combination CP2 and Voronoï's feature had a prognostic value. It has to be noted that a high value of heterogeneity index is associated with a poor prognosis.

In multivariate analysis, CP2 was found to be an independent prognostic feature just like the mitotic index and the lymph node status. The lymph node status is a well-known clinical factor; the two other features are intrinsic factors of tumor growth, at cellular level for mitotic index and at the tissue level for heterogeneity.

Surprisingly, age, tumor size, Scarff and Bloom Grade and hormone receptor status are of secondary importance compared to these 3 features.

**Table 2(abstract S43) Results of the univariate analysis**

| Variables | P value |
|---|---|
| Voronoï | 0.040 |
| Normalized variance of density | 0.170 |
| Energy (Haralick) | 0.090 |
| CP1 | 0.500 |
| CP2 | 0.016 |
| CP3 | 0.670 |

This result encourages to confront the heterogeneity feature CP2 to clinic information, such as recent or late oncologic event or the nature locoregional or distant visceral of the recurrence, and to the absence of lymph node metastasis.

**Competing interests:** The authors declare no competing interests.

**Authors' contributions:** MO, MA, NE, JM, PH, BP and CB defined the research theme, analyzed the data and interpreted the results. NE, PH and BP designed and carried out experiments for the computation the heterogeneity features. JC, VB and CB collected the data for classical clinic-pathological prognosis factors such as histological type and grade of the tumor, size or mitotic index . VB drawn the Regions of Interest for all the images. MO carried out the experiments to select the most relevant heterogeneity features and create the new CP2 feature. JM carried out all the statistical analyses. MO and MA wrote the paper.

All authors read and approved the final manuscript.

**References**

1. Dodd LG, Kerns B-J, Dodge RK, Layfield LJ: **Intratumoral Heterogeneity in Primary Breast Carcinoma: Study of Concurrent Parameters.** *Journal of Surgical Oncology* 1997, **64**:280-288.
2. Sharifi-Salamatian V, de Roquancourt A, Rigaut JP: **Breast carcinoma, intratumour heterogeneity and histological grading, using geostatistics.** *Analytical Cellular Pathology* IOS Press0921-8912 2000, **20**:83-91.
3. Bertucci F, Birnbaum D: **Reasons for breast cancer heterogeneity.** *Journal of Biology* 2008, **7**:6.
4. Fisher B, Redmond CK: **Evolution of Knowledge Related to Breast Cancer Heterogeneity: A 25-Year Retrospective.** *Journal of Clinical Oncology* 2008, **26(13)**:2068-2071.

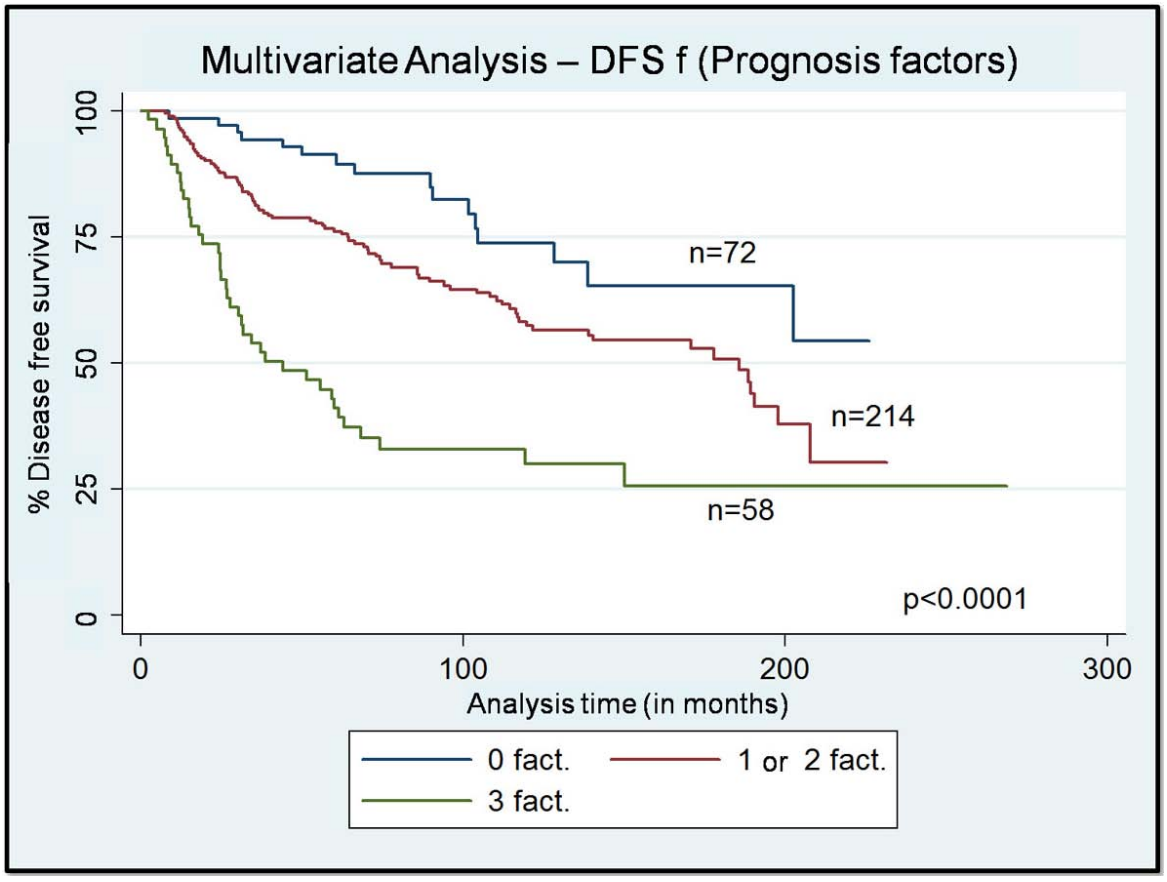



**Figure 2(abstract S43) Model build using the results of the multivariate analysis.** Multivariate analysis highlighted 3 independent prognostic factors: tumor heterogeneity (CP2), mitotic index, lymph node metastasis. The construction of this model has individualized 3 groups of patients: 0 factor, 1 or 2 factors and 3 poor prognostic factors (mitotic index > 10, lymph node metastasis in the axillary dissection, upper tercile of CP2; p < 0.0001).

5. Hans F, Dimitrov S: **Histone H3 phosphorylation and cell division.** *Oncogene* 2001, **20**:3021-3027.
6. Bossard C, *et al*: **Phosphohistone H3 labelling for histoprognostic grading of breast adenocarcinomas and computer-assisted determination of mitotic index.** *J Clin Pathol* 2006, **59(7)**:706-10.
7. Gosme M: **Modélisation du développement spatio-temporel des maladies d'origine tellurique, thèse de doctorat.** Rennes: Université de Rennes I 2007, 198.
8. Fisher RA: **The use of multiple measurements in taxonomic problems.** *Ann. Eugenics* 1936, **7**:179-188.
9. Marcelpoil R, Usson Y: **Methods for the study of cellular sociology: Voronoi diagrams and parametrization of the spatial relationships.** *J. Theor. Biol* 1992, **154**:359-369.
10. Haralick RM, Shanmugam K, Dinstein I: **Textural features for image classification.** *IEEE Transaction on Systems, Man and Cybernetics* 1973, **3(6)**:610-621.

## S44

### Webconference mixed with virtual slides as a pedagogical tool to improve pathology practice in the French Midi-Pyrenees area

ML Ranty[1*], C Guilbeau-Frugier[1], M Jacob[2], P Gil[2], E Uro-Coste[1], Y Nicaise[1], MB Delisle[1]
[1]Service d'Anatomie Pathologique et d'Histologie-Cytologie, Hôpital de Rangueil, 1 avenue Jean Poulhès TSA 50032 31403 Toulouse, France; [2]Toulouse Paul Sabatier University, DTSI, 118 Route de Narbonne 31062 Toulouse cedex 09, France
E-mail: quintyn.ml@chu-toulouse.fr
*Diagnostic Pathology* 2013, **8(Suppl 1):**S44

**Background:** The pathology is a discipline located between basic science and clinical practice. Tissue and/or cellular injury could be seen as a triptych with in upstream pathophysiological mechanism and in downstream the resulting symptoms.
In recent years, pathology, an essential step of clinical diagnosis, is being more complex integrating complementary technics (e.g. molecular biology), creating new classifications or new recommendations. So pathologists'specialization in one field is becoming a necessity. Unfortunately, since 1999, the French pathologists growth rate is negative and demographic previsions estimates that in 2050 pathologists' population should decrease by 50%. This crisis in vocations obliged pathologists to provide continuing education to their peers, share experiences but also to reform the teaching of the discipline during medical studies to make it more accessible and attractive. The development of digital imaging in this both areas has been a greatly helpful.
In France, a group of pathologists established in 1977 under the name ADICAP ("association for the development of informatics in pathology") helped the creation of adequate informatics tools [1]. Worldwide, the pathologists were among the first to use tools of telemedicine, especially in countries with difficult geographical conditions such as Norway or Canada. However, for 25 years, the difficulties of digital transmission and the static and selective state (selection bias in the sent fields) of numeric photography had left this technology underutilized [2]. In the 2000s, development of the virtual slides technology (virtual slides scanner) and of telepathology, allowed diversification and wider use of digital imaging in pathology. It had been now widely used in Canada to exchange views particularly in cases of diagnostic emergency as extemporaneous examinations [1]. In 2010, physicians in three Northern Ontario

communities had been virtually linked at all times to pathology specialists allowing frozen section examination on line [3]. It had been also used for teaching and universities, particularly in the U.S. and Switzerland, where virtual imaging have replaced conventional microscope. In France these technologies are progressively taking place both in the field of education and in the exchange of expertise. In Toulouse-Rangueil hospital, the pathology department has chosen to focus on education as the training provided to senior pathologists and to medical students.

**Methods:** Slides were scanned on « NDP Nanozoomer » Hamamatsu, in Rangueil University Hospital, in Toulouse. The virtual slides were loaded in to a computer server, in Toulouse-Paul Sabatier University. Hamamatsu provided the program patch with the NDP viewer, as a pilot program.

**For continuing education:** The Paul Sabatier University (TICE department, Direction of Technology and Information Systems) provided to the university community a web conferencing platform Adobe Connect Pro. This platform, web-based, used Flash technology present on 98% of computers. It offered not only a videoconferencing service but also a collaborative workspace (document sharing, screen or application sharing, instant messaging, whiteboard...) organized in modules. This type of system offered flexibility and brought a wealth of services in a nomadic environment. All medical data are anonymous.

Virtual meetings between pathologists: To access the platform, only an internet-connected computer, a webcam and a microphone was needed (Figure 1). The access was secured by a code.

**For theoretical teaching:** We integrated the virtual slide in presencial sessions in small numbers (30 students per group, promoting divided into 5 groups, 10 sessions). Each student had a microscope and glass slides with a representative lesion. Each slide was previously scanned. The virtual slide was projected and commented on by a teacher. The student had to the area of interest selected by the teacher at low magnification and then at higher magnification leading to the diagnosis. Questionnaire of satisfaction was distributed to students and teachers at the end of the first teaching session and at the end of the sessions cycle (10th session).

To enable students to visualize the virtual slide out of teaching session, a web-site is in process. It would allow the student to access to digital slides but also to recapitulative picture, with a brief explanation and a summary of what to remember. Normal tissue slides would also be available. Student would access the site through a custom code. It is not yet planned to grant access beyond the teaching year or to allow students external of our University to access the site. Thinking is still in progress on these points.

**Results of questionnaire of satisfaction: For continuing education:** This virtual organization had been ready for three months. To date, three meetings were organized. Members are satisfied with the ease of use and the gain of time. Some technical details are set up according to the settings of the individual participants, such as setting the webcam. Some participants, less in touch with computers, were a bit stressed and technical assistance was required.

**For theoretical teaching:** No technical problem occurred during teaching session. All students answered to the satisfaction query. Students, at the end of the first session, found that visualization (ability in detecting and observing the interest area) and understanding was better (Figure 2a). At the end of 10 sessions, opinion upon the visualization of lesions had improved since 23% of students have found it excellent and 73% better. Understanding remained unchanged.

Teachers had same opinion (visualization enhanced 78% or excellent 22% and understanding better 78% or excellent 12%). The two less experienced teachers felt stressed by changing their teaching habits. It was interesting to note that students had not noticed this stress (95% of students found them perfectly relaxed).

At the end of the first session the majority of students (53%) would not give up the microscope. Nevertheless at the end of 10th session, they were only 25% (52% favorable to give up the microscope and 23% without opinion). All teachers though that virtual slides would be more beneficial than microscope. At the end of 10th over, 77% of students recommended to go on teaching with virtual slides however, 7% thought that it should be better to give up and 15% had no opinion (Figure 2b).

**Discussion:** Virtual microscopy is undeniable progress that has many applications in education.

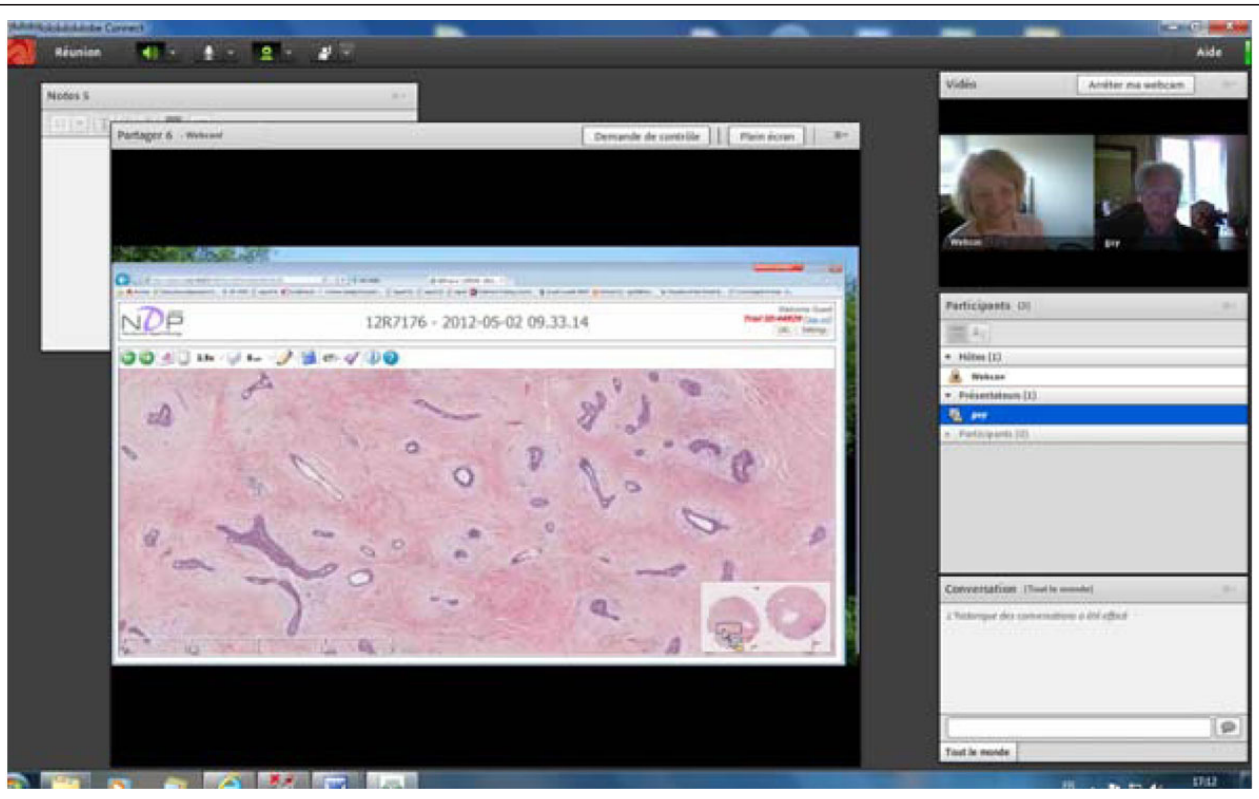



**Figure 1(abstract S44)** Virtual class visible by the senior pathologist in training.

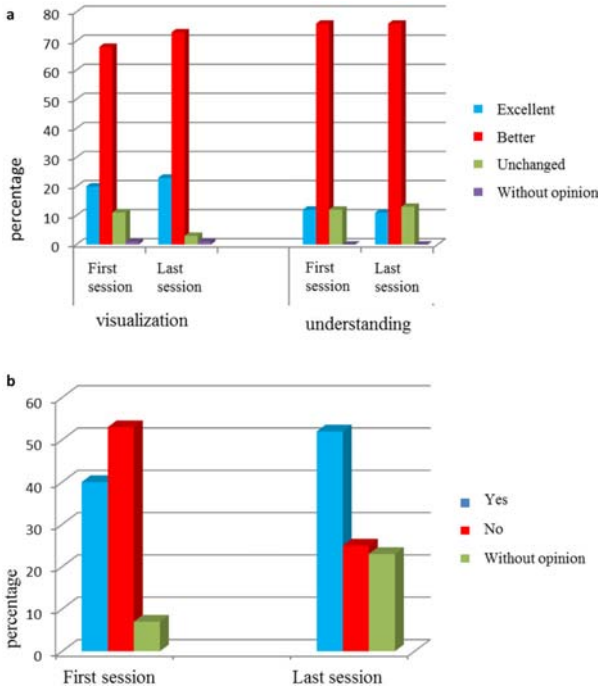



**Figure 2(abstract S44)** a) Students opinion on virtual slides for visualization and understanding at the first and last session. b) students opinion replacement of the microscope by the virtual slide at the first and last session.

This technology allowed organization of virtual group learning events for continuing education sound and images were travelling by web [4]. The virtual slides could be observed by everybody, to the extent that it was authorized, with the possibility for every one to speak, to remote on the virtual slide to show a particular area. A moderator gave the right to speak or to move on the slide. Everyone could see the same area of the slide at the same time. It was possible to take memos concomitantly in a specific area and to save them on the computer or send them by mail. Participants can view the slides prior to the meeting, provided they are received in sufficient time to allow the digitization and online procedure. In this context of penury of pathologists but also in an economic standpoint, this virtual meeting, provides continuing education for seniors, lifting the geographical barrier, thus saving travel time. It improves practices. It allowed too the construction of a virtual slide box, ensuring deepening and maintenance of the latest knowledge. So, it allows the participation of people discouraged by the duration of transport. In our department, requests for such meeting are increasing.

In theorical education, as in our data, virtual slides have been regarded favorably by students improving visualization and understanding [5-7]. It eliminates the three most frequent complaints: the focus, the binocular vision difficulty and the identification of the area of interest (virtual slides can be annotated). In universities which have introduced this technique for many years, particularly in the U.S., teachers objectified better success in practical exams [8]. Most studies have observed that students preferred the virtual slides to the microscope [5-7]. The computer is perfectly mastered by the young generations, integrated into their world and do not necessitate special competence. In our study, teachers believed that the technique would be acclaimed, students needed to be convinced. They think that knowing how to use a microscope belongs to their formation. This skepticism was found in other studies in the literature, where it could take 2-3 years before a massive acceptance (95%), even if the technic was gradually introduced [8-10]. Digital slides require no particular training for teachers. In our study, teachers were experienced and inexperienced, someone had limited computer competency. No technical problem occurred. It is interesting to note that most experienced teachers were also more motivated and more enthusiastic. Young teachers have considered, at the first session, the virtual slide as additional stress before they feel it as comfortable. Our learning website is still under development and no evaluation is yet available.

**Conclusion:** Virtual slides in pathology teaching were particularly welcomed by students as by teachers. Highly praised, the virtual slide, or virtual microscope, improves understanding and visualization and has educational application. It is powerful for senior pathologist continuing education e-learning in association with webconference. The virtual slides can be viewed on any computer, via a website, provide support for an examination or creation of a database of images ...

**Competing interests:** The authors declare that they have no competing interests.

**Authors' contributions:** M-L R: corresponding author, in charge of continuing education

CG-F: in charge of theoretical education for medical students. Contribution of this author is equally to corresponding author.

MJ: technical assistance and creation of web-site.

PG: gestion of server.

EU-C: participation of elaboration of virtual environment and pedagogical support.

YN: technical support and scanning slides.

M-DB: participation of elaboration of virtual environment and pedagogical support.

**Acknowledgements:** Authors want to thank Hamamatsu for its technical assistance, particularly Julien Aupetit et Antoine Disher.

**References**

1. Martin E, Hémet J, Christel Le Bozec C, Cordier JC: **Use of digital media in the field of pathology: present and future through the Adicap experience.** *Revue Française des Laboratoires* 2004, **367**:21-6.
2. Viellefond A, Staroz F, Fabre M, Bedossa P, Martin-Pop V, Martin E, Got C, Franc B: **Reliability of the anatomopathological diagnosis by static image transfer.** *Arch Anat Cytol Pathol* 1995, **43(4)**:246-50.
3. **Le UHN exploite le premier système de télépathologie en Ontario.** [http://www.eurekalert.org/pub_releases_ml/2010-07/aaft-t072110.php].
4. **La Pitié-Salpêtrière teste la télépathologie en temps réel.** [http://www.ticsante.com/show.php?page=story&id=767].
5. Kumar RK, Velan GM, Korelle SO, Kandara M, dee FR, Wakefield D: **Virtual microscopy for learning and assessment in pathology.** *J Pathol* 2004, **204**:613-18.
6. Dee FR: **Virtual microscopy in pathology education.** *Hum Path* 2009, **40**:1112-21.
7. Raja S: **Virtual microscopy as a teaching tool adjuvant to traditional microscopy.** *Med Educ* 2010, **44**:1126.
8. Krippendorf BB, Lough J: **Complete and Rapide Switch From Light Microscopy to Virtual Microscopy for Teaching Medical Histology.** *Anat Rec* 2005, **285B**:19-25.
9. Blake CA, Lavoie HA, Millette CF: **Teaching Medical Histology at the University of South Carolina School of Medicine : Transition to Virtual Slides and Virtual Microscopes.** *Anat Rec* 2003, **275B**:196-206.
10. Szymas J, Lundin M: **Five years of experience teaching pathology to dental students using the WebMicroscope.** *Diag Pathol* 2011, **6**:S13-S19.

## S45

### Preliminary slide scanner throughput evaluation in a intensive digitization facility setting

Vincenzo Della Mea[1*†], Giampiero Duglio[2†], Filippo Crivelli[3†], Pierluigi Banfi[2†], Giancarlo Chiovini[2†]

[1]Dept. of Mathematics and Computer Science, University of Udine, Italy; [2]Cloud Pathology Group, Milano, Italy; [3]Department of Pathology, Azienda Ospedaliera Ospedale di Circolo di Busto Arsizio, Busto Arsizio, Italy

E-mail: dellamea@uniud.it

*Diagnostic Pathology* 2013, **8(Suppl 1):**S45

**Background:** Whole Slide Imaging (WSI), called also Digital Microscopy, is the most current approach to digitization of histological information [1,2]. It allows for transferring a whole histological slide into digital form, thus enabling any kind of digital treatment from storage and transmission, to telediagnosis, to automatic image analysis. WSI technologies developed only recently, and thus most uses described in literature are coming from research and teaching applications [3]. However, one acknowledged

potential use of this technology is also aimed at dematerializing slide archives, by bringing them in digital form inside a so-called PACS (Picture Archiving and Communication System) [4]. This application would provide a major boost in the adoption of WSI in the routine work of a clinical pathology laboratory.

Two main differences can be recognised between academic applications like research and teaching, and routine application in the pathology laboratory: slide volume, and diagnostic reliability.

The former difference is related to the number of involved glass slides and the time span on which scanning occurs. Teaching in particular, but also research applications, foresee the acquisition of a limited number of slides, in terms of either total number or scanning needs per time unit. In fact, teaching with digital slides usually involves slides accumulating into a teaching archive that may slowly grow, in years and years, but with no massive amounts of slides involved. Research trials instead might involve the scanning of a large number of slides but in a limited time frame, related to the life span of the research project, and that can be archived offline at the end of the project.

Both cases differ from the routine acquisition in a clinical pathology laboratory, where there is the need for a sustained acquisition of a fraction (or all) the glass slides daily produced by the laboratory, to be made available to pathologists when needed. This means that the scanning procedure should be as efficient as possible, and in particular able to perform a sustained scanning as quick as glass slide production is in the specific laboratory. This means also that there is need for personnel (e.g., laboratory technicians) that feed the scanner with glass slides, start the scanning procedure, check associated patient data, verify results, unload slides, etc. Any technical hitch occuring in those phases (e.g., software bugs, slide loading difficulties, etc) is likely to decrease the overall thoughput of the scanner.

The latter difference is related to how digital slides are used. Scanning is not a process without errors: loss of information is always present, and derives in part from the process itself, in part from specific features and pitfalls of the scanning device, in part from the preparation technical quality of the source glass slide. In teaching applications, slides are selected for their educational meaning, so there is a selection of the acquired material, that allows for recognising scanning errors or simply missing information. Such selection is also made possible by the low number of slides acquired at each scanning session. The same can be considered true for any research usage, since it is done as part of the research project. On the other side, routine scanning is aimed at providing slides for the diagnostic work of the pathologist, either for primary diagnosis or for giving access to previous slides of the same patient when diagnosing a new histologic exam. Thus, diagnostic reliability of the digital slide should be guaranteed, and this means that acquired slides should be as good as possible as they come out from the scanner. Since every device may fail in acquisition of some slide, the least they fail, the better is.

The present paper aims at evaluating the first of the two issues, namely throughput, by means of a trial carried out by submitting to three different slide scanners a large number of glass slides coming from six different Pathology laboratories, to obtain a continuous one-month scanning session. A preliminary evaluation of the second issue, namely diagnostic reliability, has been carried out on a subset of the slides. The experimentation has been carried out to provide a technical basis for the selection of a number of scanners to be applied in a newcoming company that provides outsourced scanning and storage services. Research questions include:

- Do scanners support continous acquisitions on 24/7, in terms of software and hardware behaviour?
- Are manual operations related to scanners repeatable and hassle-free?
- How many slides per day can be truly acquired, taking into account all the operations needed?
- How many slides need to be acquired again due to quality issues?
- Is there any connection between clinical pathology laboratories procedures and glass quality, that influence acquisition quality or speed?
- Is digital diagnosis equivalent to microscope-based diagnosis?

**Material and methods: Slide scanners:** Three scanners (A, B, C) have been provided by three manufacturers by means of their national distributors, which also provided training and technical support during the experimentation. Among the different models in the companies listings, the scanners were chosen among those aimed at high throughput, i.e., with slide loaders ablle to host hundreds of slides.

**Cases and slides:** Glass slides have been provided by six Pathology laboratories from Italian hospitals. Glass slides were aimed at representing the average production of those labs, so they have been chosen consecutively from lab archives among biopsies and surgical samples. 1200 slides per lab have been requested.

**Data analysis:** The following variables have been recorded for each acquisition:

- acquisition speed
- acquisition success
- barcode acquisition success
- digital slide size
- diagnosis

Some more information has been collected regarding the scanning sessions:

- scanner downtime during experimentation
- accidental events (slide jams, etc)

For these variables, average, minimum, maximum, totals have been calculated depending on the variable, with data aggregated by scanner, by hospital and by both.

For a first preliminary evaluation of diagnostic performance, about 10% of cases will be randomnly extracted from the whole set and examined by 3 pathologist for each case and for each scanner. This way, every case will receive a total of 9 diagnoses to be compared with the gold standard microscope diagnosis. For comparison, a senior pathologist (FC) will examine all diagnoses and categorize differences on a 0-3 scale (diagnosis not possible; wrong diagnosis; incomplete or inaccurate diagnosis; correct diagnosis).

**Results and discussion: Glass slides:** Each hospital provided at least 1200 glass slides as requested. However, the first scanned batch resulted to have 1244 slides, which have been scanned without a preliminary counting, done thereafter for the other hospitals. Thus, the total amount of scanned slides is 7244.

Cases were representative of the lab production. The case set was composed by an average on 75% biopsies (range: 61%-88%) versus surgical samples; each case included an average of 3.65 slides (range:2.67-5.04), for an average total of 354 cases per lab (range: 238-449). Table 1 shows details.

**Scanner throughput:** Table 2 shows details on scanner throughput.

Scanning succeeded on slightly less than 98% of acquisitions (range: 95.65%-99.19%), with an average time of 4'28" per slide (range: 3'46"-5'14"), higher than values declared by manufacturers. Barcode acquisition failed on a very low number of slides (1%), with most errors in the very initial phases of the experimentation, due to software problems quickly solved by manufacturers programmers. This might reveal that barcode acquisition until now has not been as usual as supposed to be.

All scanners experienced some downtime, averaging at about one hour and half on the whole experimentation, one month long.

When aggregated by laboratory, the same data show a partially unexpected variability, as shown in table 3.

In particular, acquisition time per slide, averaged on the three scanners, ranges from 3'27" to 5'23" – a higher variability than the one we obtained by aggregating per scanner. While a reason could be a different percentage of biopsies versus surgical samples, at first glance this does not seem to explain all the difference, and thus needs further investigation. A candidate reason seems to be the quality of histological preparation, which might not be influencing human vision, but indeed might influence automatic scanning.

**Table 1(abstract S45) Case set features**

| Lab | slides | cases | %biopsies | slides/case |
|---|---|---|---|---|
| 1 | 1200 | 449 | 77.95% | 2,67 |
| 2 | 1200 | 238 | 78,57% | 5.04 |
| 3 | 1200 | 250 | 60,79% | 4.80 |
| 4 | 1200 | n.a. | n.a. | n.a. |
| 5 | 1200 | 433 | 88.22% | 2.77 |
| 6 | 1244 | 402 | 69.15% | 2.99 |
| | 7244 | 354 | 74.94% | 3.65 |

**Table 2(abstract S45) Scanner throughput details**

| Scanner | slides | % success | total time | time/slide | barcode failure % | downtime (h) |
|---|---|---|---|---|---|---|
| A | 7332 | 99,19% | 72:47 | 3:46 | 1,86% | 1,54 |
| B | 7200 | 95,65% | 100:38 | 5:14 | 1,19% | 2,04 |
| C | 7210 | 98,31% | 92:43 | 4:25 | 0,00% | 1,00 |
| | 21742 | 97,72% | 89:27 | 4:28 | 1,02% | 1,53 |

**Table 3(abstract S45) Acquisition success**

| Lab | total slides | % success | time/slide | barcode failure % | % biopsies |
|---|---|---|---|---|---|
| 1 | 3600 | 97,86% | 5:23 | 0,55% | 77.95% |
| 2 | 3600 | 98,25% | 4:52 | 1,06% | 78,57% |
| 3 | 3610 | 97,34% | 3:27 | 0,69% | 60,79% |
| 4 | 3600 | 98,44% | 3:37 | 0,85% | n.a. |
| 5 | 3600 | 97,03% | 4:52 | 3,47% | 88.22% |
| 6 | 3732 | 97,39% | 4:37 | 7,14% | 69.15% |
| | 21742 | 97,72% | 4:28 | 1,02% | 74.94% |

**Table 4(abstract S45) slide sizes**

| Avg. Slide Size | | | (GB) | | |
|---|---|---|---|---|---|
| | A | B | C | *avg* | *% biopsies* |
| 1 | 0,56 | 0,52 | 0,70 | 0,59 | *77,95%* |
| 2 | 0,50 | 0,53 | 0,89 | 0,64 | *78,57%* |
| 3 | 0,33 | 0,30 | 0,52 | 0,38 | *60,79%* |
| 4 | 0,32 | 0,30 | 0,48 | 0,36 | |
| 5 | 0,36 | 0,29 | 0,43 | 0,36 | *88,22%* |
| 6 | 0,41 | 0,35 | 0,53 | 0,43 | *69,15%* |
| *avg* | 0,41 | 0,38 | 0,59 | 0,46 | |

**Table 5(abstract S45) diagnostic agreement**

| | Scanner A | Scanner B | Scanner C | TOTAL |
|---|---|---|---|---|
| **No. of cases** | 116 | 116 | 116 | 348 |
| **Avg. Agreement level** | **2.655** | **2.681** | **2.310** | **2.548** |
| **Level 3** | 81% | 83% | 65% | 76% |
| **Level 2** | 9% | 8% | 14% | 10% |
| **Level 1** | 5% | 4% | 9% | 6% |
| **Level 0** | 5% | 5% | 12% | 7% |

**Slide size:** A total of about 10 Terabytes have been acquired during the experimentation, which, due to the total number of slides, corresponds roughly to the output of an Italian surgical pathology laboratory (histology only). The experimentation took about 1600 hours of scanning in one month. Table 4 shows a summary of slide sizes by scanner and by laboratory. Even here a partially unexpected variability is apparent.

**Diagnostic performance:** The preliminary diagnostic performance evaluation has been carried out on 116 cases, equivalent to 8.89% of the total number of cases. Each case has been acquired with 3 scanners, so the total number of digital cases has been 348. Cases were attributed at random to pathologists, chosen in order to have each case reviewed by one pathologist from the originating hospital and two pathologists from other hospitals.

Digital diagnosis has been categorized as follows: (0) diagnosis not possible; (1) wrong diagnosis; (2) incomplete or inaccurate diagnosis; (3) correct diagnosis. Table 5 shows details.

Results, though preliminary and in need of further investigation, seems substantially in line with other similar experimentations [5,6].

**Conclusions:** The present study provides novel insights on current slide scanners from the point of view of their massive application in a slide scanning service. As far as we know, no other study attempted yet the same kind of intensive evaluation, though relevant evaluations have been done in the European Scanner Contest series [7]. The present paper illustrates some preliminary findings on scanner throughput and reliability.

Real world scanning time seems higher than declared by manufacturers. However, it seems also dependent on slide preparation quality, thus suggesting that preparing for scanning is more crucial than preparing for microscope and human eye. From this point of view, when in need of routine, massive digital slide scanning, preparation guidelines or standards should be provided for a better and quicker overall operation.

Scanner reliability has been proven to be high and scanning success too, but both are not 100%, so, even if scanning is automatic, it is not possible to do it in a non supervisioned way. This means that personnel should take care of all the steps: loading, scanning, trouble management, informatics and networks issues, etc. From this experience, we can tell that expertise needed is both in laboratory techniques and Information technology. Figures with both expertises are rare, in particular in Italian surgical pathology laboratories.

A peculiar issue is the slide trays role: the least trays are manipulated, the least errors, glass breakings, misidentifications, and time lost. This means that in an ideal situation, they could be used not only to load the scanner , but also as a transport medium and/or as a definitive glass slide storage medium too. Furthermore, their mechanics should be reliable in intensive scanning setups, and misplacement free. The trays of the examined scanners were not equal from this point of view, and suggested that maybe they should be reengineered from the point of view of the user, in this case an outsourced service provider willing to scan hundreds of thousands slides per year.

Acquisition time and slide size varied not only per scanner but also per laboratory: further investigation is needed to obtain better knowledge on this phenomenon, though it seems related to histological preparation quality. From this point of view, guidelines should be provided for a preparation more adequate for scanning.

As a final remark, the infrastructure needed for Digital Pathology is not just a scanner on a spare table in corner of the lab, like most often until now. This approach aimed at the enthusiast pathologist – that scans a limited number of slides, not in a hurry - seems to have guided some design choice by scanner manufacturers, including trays design and partially incomplete barcode scanning software. Routine massive scanning of slides needs solid infrastructure and personnel able to deal with a number of interdisciplinary issues like:

- backup, power supply, network sizing and management,
- memory sizing, resizing, and management,
- software management and upgrade;
- slide loading, verification, unloading, archival.

At present, it is not clear who may take care of all of this in the current surgical pathology laboratory.

The step beyond is to redesign some features having in mind industrial usage of scanners, in a regular workflow, with standards-based processes, including imaging standards [8-10].

**Competing interests:** VDM has received research funding from CloudPathology. The remaining authors declare that they have no competing interests.

**Authors' contributions:** GD and VDM designed the study. GD, PB and GC supervised the overall experimentation. FC did the diagnostic performance evaluation. GD, VDM, FC analysed data. VDM wrote the first draft of the paper. All Authors reviewed and modified the paper.

**Acknowledgements:** The Hospitals directly or indirectly participating in the study: Bologna Maggiore, Napoli Pascale, Milano Niguarda, Busto Arsizio, Saronno, Trento, Vimercate, Novi Ligure, Cremona. Scanner manufacturers and distributors tha provided the devices and assistance during the experimentation: Leica, Hamamatsu, Aperio/Nikon Italia.

**References**

1. Brachtel E, Yagi Y: **Digital imaging in pathology–current applications and challenges.** *J Biophotonics* 2012, **5**:327-35.
2. García Rojo M: **State of the art and trends for digital pathology.** *Stud Health Technol Inform* 2012, **179**:15-28.
3. Hamilton PW, Wang Y, McCullough SJ: **Virtual microscopy and digital pathology in training and education.** *APMIS* 2012, **120**:305-15.
4. Demichelis F, Della Mea V, Forti S, Dalla Palma P, Beltrami CA: **Digital storage of glass slides for quality assurance in histopathology and cytopathology.** *J Telemed Telecare* 2002, **8**:138-42.
5. Wilbur DC, Madi K, Colvin RB, Duncan LM, Faquin WC, Ferry JA, Frosch MP, Houser SL, Kradin RL, Lauwers GY, Louis DN, Mark EJ, Mino-Kenudson M, Misdraji J, Nielsen GP, Pitman MB, Rosenberg AE, Smith RN, Sohani AR, Stone JR, Tambouret RH, Wu CL, Young RH, Zembowicz A, Klietmann W: **Whole-slide imaging digital pathology as a platform for teleconsultation: a pilot study using paired subspecialist correlations.** *Arch Pathol Lab Med* 2009, **133**:1949-53.
6. Fónyad L, Krenács T, Nagy P, Zalatnai A, Csomor J, Sápi Z, Pápay J, Schönléber J, Diczházi C, Molnár B: **Validation of diagnostic accuracy using digital slides in routine histopathology.** *Diagn Pathol* 2012, **7**:35.
7. **European Scanner Contest.** [http://scanner-contest.charite.de/].
8. Daniel C, García Rojo M, Bourquard K, Henin D, Schrader T, Della Mea V, Gilbertson J, Beckwith BA: **Standards to support information systems integration in anatomic pathology.** *Arch Pathol Lab Med* 2009, **133**:1841-9.
9. Daniel C, Rojo MG, Klossa J, Della Mea V, Booker D, Beckwith BA, Schrader T: **Standardizing the use of whole slide images in digital pathology.** *Comput Med Imaging Graph* 2011, **35**:496-505.
10. Daniel C, Booker D, Beckwith B, Della Mea V, García-Rojo M, Havener L, Kennedy M, Klossa J, Laurinavicius A, Macary F, Punys V, Scharber W, Schrader T: **Standards and specifications in pathology: image management, report management and terminology.** *Stud Health Technol Inform* 2012, **179**:105-22.

**Cite abstracts in this supplement using the relevant abstract number, e.g.:** Della Mea *et al.*: **Preliminary slide scanner throughput evaluation in a intensive digitization facility setting.** *Diagnostic Pathology* 2013, **8(Suppl 1)**:S45