Helsinki University Biomedical Dissertations No. 174

Automated Image Analysis of Cancer Tissue Adapted for Virtual Microscopy

Juho Konsti

Institute for Molecular Medicine Finland - FIMM
University of Helsinki
and
Department of Pathology
Haartman Institute
Faculty of Medicine
University of Helsinki

Academic dissertation

To be publicly discussed with the permission of the Faculty of Medicine of the University of Helsinki, at the Main Building of the University of Helsinki (Unioninkatu 34, 3rd floor), Auditorium XIV, on September 28th, 2012, at 12 o'clock noon.

Supervisor

Docent Johan Lundin

Institute for Molecular Medicine Finland - FIMM

University of Helsinki

Reviewers

Docent Pauliina Kronqvist

Department of Pathology

Faculty of Medicine

University of Turku

Acting Professor Jari Hannuksela

Department of Computer Science and Engineering

Faculty of Technology

University of Oulu

Opponent

Professor Janusz Szymas

Department of Clinical Pathology

University of Medical Sciences of Poznan

Poland

ISBN 978-952-10-8256-6 (paperback)

ISBN 978-952-10-8257-3 (PDF, http://ethesis.helsinki.fi)

ISSN 1457-8433

Helsinki University Print (Unigrafia)

Helsinki 2012

TABLE OF CONTENTS

TABLE OF CONTENTS **3**
LIST OF ORIGINAL PUBLICATIONS **4**
ABBREVIATIONS **5**
INTRODUCTION **7**
REVIEW OF LITERATURE **9**
BREAST AND COLORECTAL CANCER 9
VIRTUAL MICROSCOPY 9
VISUALIZATION METHODS 11
Fluorescence *in situ* hybridization 11
Immunohistochemistry 11
STUDIED BIOMARKERS 12
HER2 12
Ki-67 12
TISSUE TEXTURE FEATURES 13
AUTOMATED IMAGE ANALYSIS 13
HER2 and Fluorescence *in situ* hybridization 13
Ki-67 and immunohistochemistry 14
Automated cancer classification and local binary patterns 15
Image compression or scaling and automated image analysis of cancer tissue 16
AIMS OF THE CURRENT STUDY **18**
MATERIALS AND METHODS **19**
PATIENT SERIES 19
SAMPLE PREPARATION 20
SAMPLE DIGITIZATION 22
VIRTUAL MICROSCOPY PLATFORM 23
VISUAL ANALYSIS 24
AUTOMATED IMAGE ANALYSIS 25
FISH signal detection 26
Quantification of immunostaining 27
Texture features 27
Texture classifier 29
STATISTICAL METHODS 31
RESULTS **33**
DISCUSSION **46**
CONCLUSIONS **52**
ACKNOWLEDGEMENTS **53**
REFERENCES **55**

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications:

I. Konsti J, Lundin J, Jumppanen M, Lundin M, Viitanen A, Isola J. A public-domain image processing tool for automated quantification of fluorescence in situ hybridisation signals. J Clin Pathol. 2008 Mar;61(3):278-82.

II. Konsti J, Lundin M, Joensuu H, Lehtimäki T, Sihto H, Holli K, Turpeenniemi-Hujanen T, Kataja V, Sailas L, Isola J, Lundin J. Development and evaluation of a virtual microscopy application for automated assessment of Ki-67 expression in breast cancer. BMC Clin Pathol. 2011 Jan 25;11:3.

III. Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Ahonen T, Pietikäinen M, Lundin J. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. Diagn Pathol. 2012 Mar 2;7:22.

IV. Konsti J, Lundin M, Linder N, Haglund C, Blomqvist C, Nevanlinna H, Aaltonen K, Nordling S, Lundin J. Effect of image compression and scaling on automated scoring of immunohistochemical stainings and segmentation of tumor epithelium. Diagn Pathol. 2012 Mar 21;7(1):29.

ABBREVIATIONS

AUC	Area under the receiver operating characteristic curve
BMP	Bitmap image file format
c	Cost parameter value
CEP17	Chromosome 17 centromere
CI	Confidence interval
CK5	Cytokeratin 5
DAB	Diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DDFS	Distant disease-free survival
DFS	Disease-free survival
DNA	Deoxyribonucleic acid
DoG	Difference of Gaussians
EGFR	Epidermal growth factor receptor
ER	Estrogen receptor
FFPE	Formalin-fixed, paraffin-embedded
FISH	Fluorescence *in situ* hybridization
GB	Gigabyte
GHz	Gigahertz
H	Haematoxylin
H&E	Haematoxylin & Eosin
HCl	Hydrochloric acid
HER1	Human epidermal growth factor receptor
HER2	Human epidermal growth factor receptor 2
HRP	Horseradish peroxidase
IHC	Immunohistochemistry
JPEG	Joint Photographic Experts Group (compression format)
JPEG2000	JPEG 2000 compression format
Ki-67	Ki-67 protein
LBP	Local binary patterns
LBP/C	LBP and contrast hybrid feature
p53	Tumor protein 53

PR	Progesterone receptor
RAM	Random access memory
RGB	Red, green and blue color system
ROI	Region-of-interest
SD	Standard deviation
SVM	Support vector machine
TMA	Tissue microarray
TRIS	Tris(hydroxymethyl)aminomethane
VAR	Rotation invariant local variance

INTRODUCTION

Emerging large-scale digitization of microscopic tissue samples (i.e. virtual microscopy) in biomarker research and clinical pathology enables rapid, objective and repeatable computational analysis of the images. Automated image analysis is likely to be especially useful in personalized medicine, where high-throughput analysis is required for risk prediction, advanced diagnostics and targeted treatment of patients. Malignant tumors are profiled in detail to identify clinically actionable mutations and aberrant protein expression levels. Human observers are still predominantly interpreting visually the increasing number of biomarker assays with fluorescence *in situ* hybridization (FISH) and immunohistochemical (IHC) stainings. To aid in these quantification tasks, novel applications to automated image analysis of cancer tissues are developed, assessed and adapted to a virtual microscopy setting within this thesis.

The studies in this thesis are based on FISH and IHC stained cancer tissue samples. First, a whole-slide scanner acquires a digital representation of the entire specimen. The resulting gigapixel size image files are transferred to a virtual microscopy platform, where they can be administrated, viewed over computer networks, annotated and subjected to automated image analysis.

Computer vision methods evaluated in the current study include convolution with Difference of Gaussians (DoG) kernel, color deconvolution and Local Binary Patterns (LBP). Custom image processing algorithms based on these methods were built with freely available image processing software (ImageJ) and a commercial software package (MATLAB). In publication I, an algorithm for quantification of FISH signals in the analysis of HER-2 amplification in a breast cancer tissue is developed and assessed. In publication II, an algorithm for automated quantification of Ki-67 IHC staining in a second breast cancer series is introduced and evaluated. Publication III covers the development and validation of an automated texture classification tool for segmenting images of colorectal cancer tissue samples into epithelial and stromal compartments. In publication

IV, the effect of image compression and scaling on the performance of algorithms from publications II and III is studied.

In publication I, the comparison between the automated and visual methods for FISH quantification resulted in a high level of agreement (kappa value 0.82, AUC 0.97). The FISH quantification algorithm is freely available as an add-on to ImageJ, and also provided as a server-side application with a web interface.

In publication II, univariate survival analyses suggest that the image analysis based method (hazard ratios (HR) 1.77 and 2.34) for assessment of Ki-67 proliferation index is a statistically significant predictor of patient outcome, comparable to the human observer assessment (HR 1.41 and 2.58). In multivariate survival analyses, adjusted for other significant prognostic parameters, the automated but not the visual method was retained as significant outcome predictor (HR 1.62 and 1.73). Moderate agreement between the methods was detected (kappa value 0.57).

The colorectal cancer tissue texture classifier developed in publication III was able to distinguish epithelial and stromal compartments with a high level of agreement with visual classification (kappa value 0.93, AUC 0.995). An online database with all the images was released to other researchers.

Publication IV shows that images analyzed with algorithms from publications II and III can be reduced to one twenty-fifth of the original size by image scaling or compression while retaining reliable automated image analysis results.

This thesis suggests that computer assisted image analysis combined with virtual microscopy can be a valuable tool for biomarker research and personalized medicine. The results from the automated algorithms are comparable to visual assessments. The introduced algorithms are computationally efficient and they can be run in a batch mode to analyze thousands of samples in minutes. The algorithms are already utilized in high-throughput research settings, but may be adopted in clinical pathology routines after careful validation.

REVIEW OF LITERATURE

Breast and colorectal cancer

Breast cancer in females is the most common form of cancer worldwide and the leading cause of cancer death, whereas colorectal cancer is the third most commonly diagnosed cancer in males and the second in females (Jemal et al., 2011). In developed countries, the mortality rates of these diseases have been declining due to earlier detection and more advanced personalized treatment schemes (Jemal et al., 2011).

The foundations for personalized cancer treatment and prediction of patient prognosis are built on characterization of the tumor. Both the breast and colorectal cancers are divided into histopathological types based on the cells they originate from. The grade of the tumor describes the differentiation of the cancerous tissue i.e. how close to normal structures and cells their appearance is. The stage of a solid tumor is determined using TNM classification, where T stands for the size of the primary tumor, N for degree of tumor spread to regional lymph nodes and M for presence of metastasis to distant organs (Sobin et al., 2009). Molecular biomarkers are routinely used in tumor characterization. In breast cancer, these include estrogen receptor (ER), progesterone receptor (PR), cell proliferation (Ki-67) and HER2-oncogen amplification. Instead of, or in addition to, measuring the level of HER2 protein, the copy number of HER2 gene can be determined. After the tumor characterization, suitable personalized treatment for the patient can be selected.

Virtual microscopy

Virtual microscopy is a method for digitization of the entire microscopic sample on a glass slide, and viewing the resulting image over computer networks. Virtual microscopy has been preceded by various methods for telepathology, with the intent to perform remote diagnostics via communication networks (Weinstein et al., 1987). For example, dynamic telepathology comprises a remote

controlled physical microscope with a live video feed sent to the site of remote controlling. In contrast, static telepathology means sending of locally selected digital snapshots of regions-of-interest (ROIs) over telecommunication network to a remote pathologist (Sinard, 2005).

Virtual microscopy is an extension to static telepathology. First appearances of virtual microscopy in the literature were based on covering the specimen with low magnification images (Felten et al., 1999). Currently, the whole area of the specimen is digitized with high-power optics by a computer-controlled microscope setup, or a whole-slide scanner (Leong and McGee, 2001; Lundin et al., 2004a; Lundin et al., 2004b). A virtual slide is thus a full two-dimensional representation of a traditional microscopy glass slide. Optionally, tissue microarray (TMA) technology can be used with virtual microscopy in a high-throughput setting to digitize hundreds of samples on a single glass slide (Kononen et al., 1998). TMAs have been used with large sample series in biomarker validation (Joensuu et al., 2003; Hassan et al., 2008) and in translational research on associations between molecular changes and patient outcome (Torhorst et al., 2001; Sahu et al., 2011). The goal of translational cancer research is to link the function of the molecules measured as biomarkers with survival of the cancer patients, thus allowing creation of more effective treatments for cancer.

Commercial solutions for automated image analysis of virtual slides are available (Rojo et al., 2009), but usually the algorithms behind the user interface are hidden from the end user. This may hinder further development of the algorithms. In addition, the licenses for these software packages are often tied to a certain whole slide scanner brand, or may be expensive to purchase. Thus, there is a need for freely available automated image analysis solutions for virtual microscopy.

Visualization methods

Fluorescence *in situ* hybridization

Fluorescence *in situ* hybridization (FISH) is a method for detecting DNA sequences in chromosomes (Langersafer et al., 1982). In summary, a specific probe complementing the DNA sequence to be detected is linked with a fluorophore in the interphase or metaphase chromosomes. An epifluorescence microscope is used to excite the fluorophores and view the emitted light. The number of target DNA sequences can be quantified, as the fluorescent signals usually form distinct spots. Automated staining systems have reduced the manual preparation of FISH samples (Bankfalvi et al., 2004), but readout is still preformed by visual assessment (Levsky and Singer, 2003).
In situ hybridization is gold standard for detection of oncogene amplifications in human tumor samples and is used as "a second opinion" if immunohistochemical methods have given equivocal results (Wolff et al., 2007). The downside of FISH method is the fading of the fluorescence over time, which prevents the re-evaluation and archiving of the original FISH specimen.

Immunohistochemistry

Immunohistochemistry (IHC) is a method of detecting and visualizing the location of antibodies bound to antigens, usually proteins, in a tissue section (Coons et al., 1942). The visualization is usually achieved by conjugating either a fluorophore or an enzyme to the antibody. Fluorophores can be visualized with an epifluorescence microscope as previously described, and the enzymes that are used in IHC can act as catalytes in color producing reactions. The colors can be viewed with a brightfield microscope. In addition to immunohistochemical staining, a counterstain such as haematoxylin is usually applied to specimen to enhance the contrast and aid in navigating through the section.

Studied biomarkers

HER2

Human epidermal growth factor receptor 2 (HER2) is a protein, also known as ErbB-2, Neu, CD340 or p185. HER2 is encoded by the *ERBB2* (*HER2*) gene, located on the human chromosome 17 (17q21-q22). Over-expression of HER2 protein stands for excess production of the protein in cells, whereas amplification of *HER2* gene increases the number of gene copies in the DNA of the cell. The both phenomenon have been shown to activate molecular pathways in breast cancer leading to tumor progression and growth and associate with a less favorable survival. A monoclonal antibody trastuzumab is successfully used in treatment of patients with HER2 over-expression or *HER2* amplification (Joensuu et al., 2006). FISH analysis to assess *HER2* amplification in breast cancer patients has shown to be of clinical value and *in situ* hybridization methods are currently gold standard (Hicks and Tubbs, 2005; Wolff et al., 2007).

Ki-67

Ki-67 is a nuclear protein associated with cell proliferation. It is used as a marker for growth fraction of cancer cell populations, i.e. a higher number of Ki-67 positive cancer cells corresponds to a higher proliferation rate of the tumor. The prognostic value of Ki-67 has been shown in multiple cancer types, including carcinomas of the breast, prostate and brain, and most thoroughly studied in breast cancer (Querzoli et al., 1995; Domagala et al., 1996; Pietiläinen et al., 1996; Querzoli et al., 1996; Rudolph et al., 1999; de Azambuja et al., 2007; Stuart-Harris et al., 2008). The higher level of Ki-67 expression is associated with worse outcome of the patient (de Azambuja et al., 2007; Stuart-Harris et al., 2008). The standard staining method used for Ki-67 IHC assessments is diaminobenzidine (DAB) immunostaining combined with haematoxylin (H) counterstaining.

Tissue texture features

In an image, texture is built on the information of spatial arrangement of intensity or color values of the image. Pathologists examine the microscopic morphology of the tissue, i.e. the form of the cells and relationships between their structures. The morphology of a cancer tissue specimen can be considered as a texture of different types of cells and extracellular structures. Epithelium of cancer tissue can vary from organized cell structures of well differentiated (low grade) tumors to complete lack of organizational cell structures of poorly differentiated (high grade) tumors. Tissue texture is thus related to prognosis of the patient. Stromal tissue intervening the tumor epithelium has a different texture and is loose in arrangement and directionally organized (Eroschenko, 2005). Texture features are calculated with mathematical algorithms, and describe the texture content of an image. Various texture analysis methods include model-based methods such as Markov random fields or fractals, statistical methods based on co-occurrence matrices, and signal-processing methods based on local linear transforms, multichannel Gabor filtering or wavelets (Mirmehdi et al., 2008).

Automated image analysis

HER2 and Fluorescence *in situ* hybridization

A series of studies have been published on development of automated methods for FISH quantification (Netten et al., 1997; Adiga and Chaudhuri, 1999; Klijanienko et al., 1999; Grigoryan et al., 2002a; Narath et al., 2004; Iourov et al., 2005; Raimondo et al., 2005; Kajtar et al., 2006). A top-hat spatial filter has been used in most studies that describe the applied image processing method (Netten et al., 1997; Adiga and Chaudhuri, 1999; Grigoryan et al., 2002a; Raimondo et al., 2005). None of the previous studies used a Difference of Gaussians (DoG) algorithm, which has been thought to mimic the neural signals rising from the retina of the mammalian eye (Enroth-Cugell and Robson, 1966).

The fluorescence signals may originate from different depths of the tissue sample that typically is 5-10 micrometers thick. The focus depth of the epifluorescence microscope is usually less than the thickness of the sample. Thus, information from different focal layers is needed. The extended depth-of-field image processing method generates in-focus microscopic images of three-dimensional objects by combining the signals from multiple focal layers, and has also been reported to improve FISH quantification (Grigoryan et al., 2002a; Grigoryan et al., 2002b).

Previous studies have presented only limited data on correlation and agreement between automated and visual methods in clinical samples. The focus of some of the studies has been on details of the image processing methods, or the reported agreement has been based on a cell level and not on a tissue sample level (Netten et al., 1997; Adiga and Chaudhuri, 1999; Raimondo et al., 2005). The automated count correlated with the visual counting in 89–95% of the analyzed cells in two of the studies with six and thirteen specimens, respectively (Netten et al., 1997; Adiga and Chaudhuri, 1999). Only few studies have compared the methods on a patient level, but agreement statistics needed for direct comparison with other studies have not been reported (Klijanienko et al., 1999; Grigoryan et al., 2002a). Perfect agreement on amplification status of five tissue specimens was reported in one of the studies (Grigoryan et al., 2002a). A correlation coefficient of 0.92 was reported in a study on 41 invasive breast carcinomas (Tubbs et al., 2006). A significant correlation (P=0.0001) was observed between automated and visual scoring in a study on 26 breast cancer cases, but correlation coefficients or agreement figures were not reported (Klijanienko et al., 1999).

Ki-67 and immunohistochemistry

Computer-assisted quantification of IHC stained protein expression has been studied extensively (Kolles et al., 1993; Pinder et al., 1995; Querzoli et al., 1995; Layfield et al., 1996; Pietiläinen et al., 1996; Querzoli et al., 1996; Corletto et al.,

1998; Lehr et al., 1999; Camp et al., 2002; Bloom and Harrington, 2004; Chen et al., 2004; Chung et al., 2007; Strömberg et al., 2007; Rexhepaj et al., 2008; Turbin et al., 2008; Sullivan et al., 2009; Turashvili et al., 2009; Tuominen et al., 2010; Dahlman et al., 2011). Also Ki-67 immunostaining has been shown to be evaluable with computer vision methods (Querzoli et al., 1995; Pietiläinen et al., 1996; Querzoli et al., 1996; Tuominen et al., 2010). DAB and H signals must be separated before the assessment of Ki-67 staining. A computational method, such as color deconvolution can be used for regular RGB images (Ruifrok and Johnston, 2001). Another option is to use multispectral imaging, where narrow bands of wavelengths can be captured separately, and a specific combination of bands can be used to construct images with only DAB or H signals (Rojo et al., 2009).

Two previous studies have found that semi-automated analysis of Ki-67 staining with image analysis can be used for prognostic assessment of patients with breast cancer (Pinder et al., 1995; Tuominen et al., 2010). The common approach for evaluating Ki-67 staining has been the segmentation of the nuclei in the sample and calculating the percentage of positively stained nuclei (Tuominen et al., 2010). This approach requires optimal levels of both DAB and H stainings, otherwise the staining in adjacent cells might fuse together, rendering the segmentation much more challenging. Moreover, also cytoplasmic and membraneous expressions of Ki-67 have shown to be prognostic factors in a recent study (Faratian et al., 2009), and these are not considered if only nuclei are segmented and analyzed.

Automated cancer classification and local binary patterns

In medical research, studies on automated segmentation of tissue images into cellular compartments have been published based on color space methods (Eramian et al., 2011), texture analysis (Hamilton et al., 1997; Diamond et al., 2004; Florea et al., 2005; Karacali and Tozeren, 2007; Sertel et al., 2009) or other morphology-based algorithms (Di Cataldo et al., 2010). A few automated cancer classification systems have been presented in the literature (Demir et al., 2005; Hwang et al., 2005; Petushi et al., 2006; Tahir and Bouridane, 2006; Doyle et al.,

2008; Signolle et al., 2010). Most of these methods rely on measurement of feature deviations on either a cellular-level or a tissue-level. The features can be divided into morphological, textural, fractal-based, topological and intensity-based groups. In this thesis, a texture-based method is covered in detail. Local binary patterns method (LBP) is a computationally simple and efficient approach to compute features for pattern classification (Ojala et al., 1996; Ojala et al., 2000; Ojala et al., 2002). LBP is grayscale and rotation invariant, which are desired properties in histological sample studies with unknown sample orientation, possibly variable staining intensities and illumination conditions. Partly due to these advantages, LBP has been used both in histology (Florea et al., 2005; Sertel et al., 2009) and in cytology (Nanni and Lumini, 2008b; Nanni et al., 2010). LBP has also been applied successfully to various other fields, such as face recognition (Ahonen et al., 2006), fingerprint matching (Nanni and Lumini, 2008a) and iris detection (Mäenpää, 2005). Other popular texture-based methods for pattern classification include Haralick features (Haralick et al., 1973) and Gabor filters (Manjunath and Ma, 1996).

Image compression or scaling and automated image analysis of cancer tissue

Uncompressed images produced by current whole-slide scanners can be as large as hundreds of gigabytes, which poses a challenge to data storage. Usually the digital slides need to be archived. The storage capacity needed for image archives and backups is rapidly increasing. To save space in the digital archive, the images can be compressed or scaled down.
In addition to challenges with regard to storage space, the bandwidth requirements for transferring the images to and from the digital archives are decreased by image compression and scaling. However, this comes at the expense of the higher processing power needed initially for image compression and scaling, as well as decompression of the images to be viewed or analyzed. The option for more objective, repeatable and less laborious automated quantification of biomarkers has recently become widely available (Mulrane et al., 2008; Kayser et al., 2009; Rojo et al., 2009). Downscaling reduces the

computational complexity of the automated quantification and classification algorithms, but very few studies have been published on the effects of image compression and scaling on the results of these algorithms.

Generally, uncompressed images without image scaling are recommended for image analysis (Gonzalez and Woods, 2002). By definition, in lossy image compression and downscaling of images some of the details in the original images are lost. It is however unclear whether these details are needed in making the clinical decision on the images. Published data suggests that image compression has only a minor effect on visually performed diagnostics (Kalinski et al., 2009) and computer assisted image analysis (Tengowski, 2004). The JPEG2000 format is considered as the most efficient way to store digital whole-slide images (Entwistle, 2003). The role of compression in automated quantification of IHC stainings has been assessed in a few previous publications with results showing only small effects on the classification and quantification accuracy (Lopez et al., 2008; Lopez et al., 2009; Lejeune et al., 2011). In texture classification of natural image series, the LBP algorithm provided reliable results in JPEG compressed images up to quality levels of 75 (Martens et al., 2010). Image scaling and visual image quality has been studied previously (Barker et al., 1998), but none of the publications has taken into account the effect of image scaling on the automated image analysis.

AIMS OF THE CURRENT STUDY

The objective of the current study was to develop and evaluate automated methods for high-throughput quantitative tissue analysis adapted to a virtual microscopy environment. More specifically, we studied:

-Automated quantification of FISH signals in HER2 amplification assessment of breast cancer tissue images.

-Automated quantification of IHC staining in Ki-67 proliferation assessment of breast cancer tissue images.

-Automated segmentation of colorectal cancer tissue images into epithelial and stromal tissue compartments.

-The effect of image compression and scaling on the performance of the previously mentioned automated methods for Ki-67 IHC assessment and cancer tissue segmentation.

MATERIALS AND METHODS

Patient series

For the automated analysis of FISH in publication I, tissue samples from 42 patients diagnosed with breast cancer at Tampere University Hospital or Seinäjoki Central Hospital during the years 2004 and 2005 were digitized. The samples were chosen to represent a distribution of HER2 non-amplified, amplified and equivocal border-line copy number cases similar to that observed in routine diagnostics.

For the analysis of automated Ki-67 quantification in publication II, 2842 breast cancer patients diagnosed in 1991 and 1992 within five hospital districts of Finland were chosen from the previously published FinProg series. The hospital districts cover half of the population of Finland and the cases represent 53% of breast cancers diagnosed during this period. Clinical data was extracted from the hospital case records, hospital registries, the Finnish Cancer Registry, and Statistics Finland. The data consisted of more than 50 clinicopathological factors, including the histological type and grade of breast cancer, the number of metastatic and examined lymph nodes, primary tumor size, tumor ER and PR content evaluated by IHC in the TMA samples, treatment details, and follow-up data. More than 50 pathologists participated in histological typing and grading of cancer at the time of diagnosis according to the World Health Organization guidelines. The median patient follow-up time was 9.5 years at the time of study. Patients with ductal or lobular carcinoma *in situ* were excluded from the study as well as those who had distant metastases at the time of the diagnosis, bilateral breast cancer, or other malignancy than breast cancer in history, except basal cell carcinoma or cervical carcinoma *in situ*. A patient was also excluded if no breast surgery was carried out. A single patient may have been excluded for one or more reasons.

For publication III on the segmentation of cancer tissue based on texture analysis, tissue samples from a previously studied series of 643 consecutive patients who underwent surgery for histologically verified colorectal cancer at the Helsinki University Central Hospital in 1989 to 1998 (Linder et al., 2009).

For analysis of compression and scaling in publication IV, two hundred patients were selected for the automated IHC analysis from a series of 570 consecutive patients with invasive non-metastatic breast cancer, treated in 1997 and 1998 at the Department of Oncology of the Helsinki University Central Hospital. Tumor samples were analyzed using the TMA technique. All patients underwent surgery and were treated according to standard guidelines for adjuvant chemotherapy, radiotherapy and endocrine treatment.
The automated tumor segmentation in publication IV was based on 144 randomly selected colorectal cancer tissue samples from the series used in publication III.

Sample preparation

In publication I, established protocols used in routine diagnostics for formalin fixation and processing into paraffin blocks were applied to the tumor samples. Representative tumor regions were defined and marked from haematoxylin and eosin (H&E) stained sections. One or multiple regions were visually selected to obtain the minimum number of 40 evaluable cancer cells for the FISH analysis. FISH was performed with a probe mixture of HER2 (SpectrumGreen), chromosome 17 centromere (SpectrumAqua), and Topoisomerase II-alpha (SpectrumOrange) based on the recommended protocol (Vysis, Abbott Laboratories, Illinois, USA). Enzymatic digestion was conducted at 37 degrees Celsius for 20–25 minutes with Digest-All III (Invitrogen, Carlsbad, California, USA) solution. The slides were post-fixed with 10% formaldehyde for 10 minutes and dehydrated in graded ethanols. Hybridization was carried out at 37 degrees Celsius overnight. The slides were mounted in Vectashield-DAPI (49,6-

diamidine-29-phenylindole dihydrochloride) (Vector Laboratories, Burlingame, California, USA).

In publication II, 1931 formalin-fixed, paraffin-embedded (FFPE) samples of the primary tumors were collected. For every patient 1-4 representative tissue cores (0.6 millimeters) were assembled into 23 tissue TMAs (Joensuu et al., 2004). Sections of 5 micrometers were cut and transferred to glass slides. A mouse monoclonal antibody (MM-1; Novocastra Laboratories; 1:1000 dilution) was used for Ki-67 immunostaining. The resulting TMAs had an evaluable Ki-67 staining for 1334 (69,1%) of the 1931 cases. 597 samples were lost due to tissue processing or with non-representative tissue spots.

H&E-stained tissue sections were used to locate representative colorectal cancer tumor regions in the FFPE samples in publications III and IV. Three punched cores from each of the 643 donor blocks were transferred to construct 27 TMA blocks with 10–180 tumor samples each. Two of the TMAs were randomly selected for publication IV. Four-micrometer sections were cut from the TMAs and transferred to glass slides. A Lab Vision Autostainer TM 480 (Lab Vision, Fremont, CA) was used for automated IHC according to standard procedures. A mouse monoclonal antibody (NCL-EGFR; Novo Castra, Newcastle upon Tyne, UK; 1:10 dilution) was used for EGFR immunostaining. The reaction products were visualized with diamino-benzidine chromogen (DAB) and counterstained with haematoxylin for 1 minute.

Representative breast cancer tumor regions in the FFPE samples were defined from the H&E-stained sections and marked for publication IV. Four cores from breast cancer specimens were punched from each donor block and transferred to the two TMA blocks, each containing 400 tumor samples. Sections of 3–4 μm were cut from the TMA blocks and transferred to glass slides. A mouse monoclonal antibody (Mib-1; Dako, Stockholm, Sweden; 1:100 dilution) was used for Ki-67 IHC in an automated immunostainer (Ventana Medical Systems Inc., Tucson, AZ, USA) with a DAB kit (Ventana). The slides were manually counterstained with haematoxylin.

Sample digitization

Samples in publication I were digitized with an Olympus BX61 microscope, equipped with epifluorescence optics, by an Olympus CC12 Soft Imaging System digital color camera (resolution 1376x1032). All FISH images were captured using an oil 60x objective with a numeric aperture of 1.25. Filters used for DAPI, HER2 and CEP17 were DAPI, SpectrumGreen and SpectrumAqua single-pass filters (Vysis), respectively. SpectrumOrange fluorescence (for Topoisomerase II-alpha) was ignored in this publication. For each channel (DAPI, HER2 and CEP17), maximum function between three different focal layers was calculated resulting in a single best-focus greyscale image. SpectrumGreen was pseudocolored as red, SpectrumAqua as green, and DAPI as blue, and channels were merged into an RGB image.

The TMA slides related to publication II and III were digitized with an automated whole-slide scanner (Mirax Scan, Zeiss, Göttingen, Germany), using a 20x objective with a numerical aperture of 0.75 and a Sony DFW-X710 camera (Sony, Tokyo, Japan). The pixel resolution of the scanner was 0.26 micrometers per pixel. The virtual slides were compressed with a ratio of 1:5 to a wavelet file format (Enhanced Compressed Wavelet, ECW, ER Mapper, Erdas Inc, Atlanta, Georgia).

In publication IV, the breast and colorectal cancer TMA slides were digitized with the same protocol as used in publications II and III, except that the images were initially stored in an uncompressed bitmap (BMP) format. Subsequently, the images were compressed to a publicly available ISO Standard JPEG2000 wavelet format with the freely available JVScomp software (Tuominen and Isola, 2009). The settings for the compression were: lossless, and lossy compression with ratios 1:12, 1:25 and 1:50. The compressed breast cancer tissue images were scaled down to 1:1, 1:2, 1:4, 1:8, 1:16, 1:32, 1:64 or 1:128. The LBP/C algorithm

used in the automated tumor segmentation method is scale dependent; therefore the colorectal cancer series images were not scaled down.

Virtual microscopy platform

In publication I, two FISH virtual slides were prepared to demonstrate the automated FISH quantification tool within a virtual microscopy framework. Sixteen separate images were stitched together into a single montage file a.k.a. virtual slide (Figure 1), which was compressed into a wavelet image file (JPEG2000) using the ER Mapper software (Earth Resource Mapping Pty, West Perth, Australia). The virtual slides were uploaded to our web server (http://www.webmicroscope.net) running the Image Web Server software (Earth Resource Mapping Pty). Virtual slides on the website can be viewed and analyzed within a standard web browser.

In publication II, the compressed virtual slides of the Ki-67 TMAs were uploaded to the web server described above. The server hardware in publication II consisted of two quad-core Intel Xeon processors and 16 GB of RAM, and allowed image analysis scripts to be run in a batch mode while also hosting the virtual slides. An annotation system in the graphical user interface (Webmicroscope) was used to define TMA spot locations on the virtual slides and link spots to the corresponding clinical data. The TMA spot images sized 1634x1634 pixels were extracted from the virtual slides as separate, losslessly compressed image files. The images were used as input to the computer vision algorithm, and stored on the server for later visual inspection and documentation.

In publications III and IV, the virtual slides of the TMAs were uploaded to the web server described in previous paragraphs. The server hardware had been updated to a 3.33 GHz six-core Intel Core i7 processor, and 24 GB of RAM. For publication IV, each of the breast cancer tissue cores in Ki-67 TMAs were annotated using an annotation tool in the graphical user interface (Webmicroscope) for subsequent image analysis.

Visual analysis

In publication I, every evaluable cell within each region of interest (ROI) was counted visually with the built-in cell-counter in ImageJ prior to the automated counting. The signals were counted separately for HER2 (red) and CEP17 (green) in the pseudocolored multi-color image, and the ratio of red to green signals was calculated for every ROI. The specimens were classified into three categories according to the calculated HER2/CEP17 ratio: less than 1.8 was considered non-amplified, 1.8–2.2 borderline and higher than 2.2 were scored as amplified with regard to HER2. The same criteria were used for automated signal enumeration. A binary cut-off of 2.0 was applied for calculation of specificity, sensitivity and accuracy of the automated counting, with the visual counting result as the reference.

For publication II, the visual analysis of 1292 of the cases had already been done as part of a previous study (Sihto et al., 2008). The extent of Ki-67 staining had been assessed by counting the number of positively stained tumor cells and classifying samples into negative, moderate or high expression categories with cut-off values at 0% and 20%. Immunostainings for ER, PR, HER1, HER2, p53 and CK5 had also been assessed previously and combined to form molecular subtype categories. The definition of molecular subtypes was following: luminal A (ER+ and/or PR+, HER2-), luminal B (ER+ and/or PR+, HER2+), basal-like (ER-, PR-, HER2-, CK5 +, and/or HER1+), HER2+/ER- (ER-, PR-, and HER2+), and five-marker negative (negative for all markers).

For training of the algorithm in publication III, representative epithelial (n=41) and stromal (n=39) regions were defined with the annotation tool in the graphical user interface (Webmicroscope, Figure 2). The training set images were discarded from subsequent analyses. A separate validation set (n=576) with 360 epithelial and 216 stromal images was annotated for optimization of the algorithm parameters. Finally, a separate test set (n=720) with 425 epithelial and 295 stromal images was defined for assessment of classifier accuracy. A pathologist verified all image annotations (SN). The dimensions of the

annotations varied between 93 and 2372 pixels in width, and between 94 and 2373 pixels in height. Pixel resolution was kept constant (0.26 micrometers per pixel) throughout the annotations. The images used in publication III were stored in a database and are available at http://fimm.webmicroscope.net/supplements/epistroma.

The colorectal cancer tissue images related to publication IV were annotated as described in the previous paragraph. The number of areas with stroma was 138 and with epithelium 269. The dimensions of the annotated areas varied in width between 168 and 1191 pixels and in height between 168 and 1190 pixels.

The visual scoring of the Ki-67 percentage in breast cancer TMAs related to publication IV was performed under the supervision of an experienced breast pathologist (R-MA) (Ahlin et al., 2007). The percentage of Ki-67 positive cells with unequivocal nuclear staining was evaluated in one high-power field (40x objective and a field-of-view with a diameter of 450 micrometers) in each of the four tissue cores on the TMAs. At least 200 cells were analyzed in each tumor. All statistical analyses were done using average values obtained by dividing all the positive cells from the four biopsy specimens by the total number of cells from the same specimens.

Automated image analysis

The automated image analysis algorithms in publications I and II were created with ImageJ (Rasband, 1997-2006), an open source image manipulation tool widely used in biomedical image processing. The algorithms were written with the macro language implemented in ImageJ, allowing the use of built-in functions and plug-ins developed by the ImageJ community.

FISH signal detection

In publication I, representative areas of background, nucleus (DAPI), red signal (HER2) and green signal (CEP17) were selected from the ROIs. The intensity values of these areas were analyzed and the results used in nucleus-background thresholding and in calculation of signal-to-noise ratio required for successful quantification of FISH signals.

The RGB images were split into red, green and blue color channels. The nuclei of the cells stained with DAPI were detected by applying the previously calculated nucleus-background threshold to the blue channel image. The blue channel image was binarised to create a nuclei-mask image for calculation of the CEP17 and HER2 signals.

The CEP17 and HER2 signals were assessed from green and red channels, respectively. First the image was smoothed with 3x3 pixel Gaussian kernel. Next, each pixel of the image was compared to surrounding background to determine if the pixel was a part of a potential signal or not by convolving the image with a 7x7 pixel DoG kernel. Background noise was removed by multiplying the image with the nuclei mask image. Remaining background noise from nuclear regions was removed by thresholding the image with a full 8-bit pixel value of 255.

A built-in function of the image processing software (Analyze Particles; ImageJ) was used for analysis and counting of detected signal spots. The function scans the image or selection until the edge of an object (spot) is found. The object is outlined by using the wand tool, measured, and filled with background color. The scanning was repeated until the end of the image. Clusters of HER2 signals were also approximated by the Analyze Particles function, by studying the circularity and size of detected spots. Found spots representing both the green and red signals were transparently overlaid with respective colors on top of the original blue channel image, creating the final result image.

The workstation used for image analysis in publication I was equipped with 2.8 GHz Intel Pentium 4 processor and 1 GB of RAM.

Quantification of immunostaining

The algorithm in publication II first splits the RGB images into separate color channels using a color deconvolution method (Ruifrok and Johnston, 2001). The ImageJ plugin for color deconvolution separated diaminobenzidine (DAB) and haematoxylin (H) stainings with a built in vector. Afterwards, DAB and H images were processed individually. Five random test samples stained for Ki-67 were used to calibrate the threshold levels for DAB and H. The thresholds were kept constant for the analysis of the main image dataset. Thresholding creates potentially overlapping binary masks of DAB and H positive areas. Binary masks were joined into a single result image, with the area of DAB-positive pixels pseudocolored red regardless of H-status, the area of H-positive and DAB-negative pixels pseudocolored green and the background displayed as white. The extent of staining was defined as the number of DAB-positive pixels divided by the union of the number of DAB-positive pixels and the number of H-positive pixels. The intensity of staining was defined as a mean pixel value of the original DAB image within the DAB-positive area. The mean intensity value was scaled to range from 0 to 100 percent.

Texture features

In LBP, the algorithm used in publication III, a pixel in a grayscale image is surrounded by a circularly symmetrical neighborhood of P equally spaced points on a circle of radius R. The gray values of points are estimated by interpolation, when a point is not situated in a center of a pixel. In order to achieve grayscale invariance, the gray value of the center pixel is subtracted from gray values of neighborhood points and the resulting values are binarised so that non-negative values are assigned with value 1 and negative values with value 0. When combined, these values can be thought as P-bit binary number. Rotation invariance is achieved by performing a circular bit-wise right shift P-1 times on the P-bit binary number and selecting minimum of these numbers. In practice, the occurrence frequencies of unique patterns vary greatly, which leads to poor

discrimination. By using uniform patterns to group more infrequently appearing patterns together, the discrimination can be enhanced. Uniform patterns are defined as patterns with at most two bitwise 0/1 changes. By definition, there are P+1 uniform binary patterns in a circularly symmetric neighborhood of P pixels. These uniform patterns are denoted with numbers from 0 to P. Patterns with three or more bitwise changes are grouped together and denoted with number P+1. A simplified demonstration of LBP algorithm is shown in figure 3. Multiresolution analysis of texture can be achieved by combining multiple LBP operators with varying P and R values. By using Gaussian low-pass filtering, problems induced by sparse sampling in multiresolution analysis can be defeated (Mäenpää and Pietikäinen, 2003).
In the classification phase, the dissimilarity of sample and model histograms is evaluated as a test of goodness-of-fit, which is measured with a statistical test (Ojala et al., 2002). A test sample can then be assigned to a specific class of the model according to the dissimilarity measure.

In publication III, images were first scaled down by a factor of 0.5. The RGB image was converted to grayscale with a following weighted equation: 0.2989 * R + 0.5870 * G + 0.1140 * B. A binary mask separating foreground and background was created with a threshold value of 240. Structures in the binary mask were smoothed with morphological closing and eroding operators (Gonzalez and Woods, 2002).
The classification of texture was performed in independent image blocks. An 80 x 80 pixel square window was slid row-by-row 40 pixels at a time through the image to create the 50% overlapping blocks. A single block was further processed, if it contained at least 50% tissue, determined by the aforementioned binary mask.

To extract the LBP/C texture features, each block was processed by using two discrete joint distributions: the first was a combination of rotation invariant, uniform LBP with eight pixels and a radius of one, and an eight pixel rotation invariant local variance (VAR) with a radius of one (Ojala et al., 2002). The second joint distribution was a combination of rotation invariant, uniform LBP

with sixteen pixels and a radius of two, and a sixteen pixel VAR with a radius of two. The histograms were concatenated to a single feature vector with a length of (8 + 2) x 8 + (16 + 2) x 8 = 224 bins. The Euclidean norm of the feature vector was normalized to one. The MATLAB implementations for the LBP and VAR operators used in publication III are available at http://www.cse.oulu.fi/MVG/Downloads.

In publication III, the results from the LBP/C features were compared to the results from Haralick texture features (Haralick et al., 1973) and Gabor features (Manjunath and Ma, 1996). To calculate the Haralick texture features, pixel values from grayscale images were linearly quantized to 8 levels, defining the size of the co- occurrence matrix to 8 x 8. Three symmetrical co-occurrence matrices were used to describe second-order statistics. The metrics computed from the matrices and used as inputs for the classifier were autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference normalized, and inverse difference moment normalized (Haralick et al., 1973; Soh and Tsatsoulis, 1999; Clausi, 2001).

The Gabor features were computed from the filter bank defined by using six orientations and four scales. For each parameter combination a unique Gabor transformation was defined, and the mean and the standard deviation of the magnitude of the transformation coefficients were used for classification purposes. The parameter settings resulted in a 2x24 component feature vector used as an input for the classifier.

Texture classifier

The classification of the image blocks extracted from the input images was performed with a linear support vector machine (SVM). The SVM classified the

image blocks based on a model that it has learned from the LBP/C, Haralick and Gabor features of the training set images. The classes of the training set were separated with the largest possible margin by a hyperplane described by the model. The validation set images were used to optimize the parameters of the trained classifier, and the independent test set images were used to test the performance of the classifier. A linear capacity constant SVM (C-SVM) was implemented by a MATLAB library for large linear classification (LIBLINEAR) (Fan et al., 2008).

A varying number of blocks per image were classified by the SVM due to variability in pixel dimensions of the analyzed images. The test image was assigned to either epithelium or stroma class according to the average SVM score of all the blocks in the image. The absolute value of the classification score was the distance between the feature vector and the decision hyperplane, and the sign of the classification indicated whether the feature vector belonged to the epithelium or stroma class. Thus, the feature vectors near the hyperplane were more likely erroneous than the ones further from it, i.e. the decision value was a measure of the prediction accuracy. Images with an SVM score higher than 1 or lower than -1 were regarded as stronger candidates for the respective classes, whereas those with an SVM score between -1 and 1 were considered as weaker candidates.

In the result images the pixels belonging to the block in the original image are pseudo-colored according to the decision value of the particular block with a heat map: The most probable epithelial regions with large positive values were colored with dark red and the most probable stromal regions with large negative values with dark-blue. The colors between the extremes changed from light blue and turquoise to light green, and from light green to yellow and orange. Light green color represented the most uncertain areas with values near zero. Averages of the overlapping decision values were used for the overlapping block areas.

The automated IHC algorithm from publication II was also used in publication IV with the breast cancer specimens, and the automated texture analysis algorithm

from publication III was also used in publication IV with the colorectal cancer specimens.

Statistical methods

In publications I, II, III and IV, percentage agreement and kappa statistics were used to study the level of agreement between the visual and automated methods. In publication IV, percentage agreement and kappa statistics were also used in comparisons of results from differently compressed and scaled images, regarding the results from losslessly compressed and non-scaled images as a reference. Kappa values were categorized as suggested previously in the literature: <0 as disagreement, 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement (Landis and Koch, 1977).

The automated countings in publications I and III were also evaluated by calculating the accuracy and the area under the receiver operating characteristic curve (AUC). In publication I, correlation coefficients were calculated using the Pearson product-moment correlation method.

For statistical analyses in publication II, the highest extent or intensity of Ki-67 staining was selected if multiple tissue cores were evaluable for the same patient. Continuous Ki-67 extent and intensity values generated by the computer vision algorithm were grouped according to tertiles. The automated results were also split into similar proportions of low, moderate and high expression as obtained based on the visual results, to evaluate the agreement between the visual and automated assessment of Ki-67. The chi-square test was used for frequency table analysis. Since the categories of the variables are ordered, a linearly weighted kappa value was used for evaluation of agreement between the visual and automated methods (Cohen, 1968). The Kaplan-Meier method was used in calculation of life tables. Distant disease-free survival was calculated from the date of the diagnosis to the date of detection of metastases outside of

the locoregional area or to the date of death from breast cancer, if a patient died of breast cancer without distant metastases. Patients who died from an intercurrent disease were censored on the date of death. The log-rank test was used for survival curve comparison. Multivariate survival analyses were performed with the Cox proportional hazards model with following covariates: automated assessment of Ki-67 extent of staining or visually assessed Ki-67 proliferation index, method of tumor detection, tumor size in centimeters, number of metastatic lymph nodes, histological grade, and age at diagnosis. The assumption of proportional hazards was ascertained by assessment of log minus log survival plots. All P values are two tailed. In publication IV, the continuous visual and automated Ki-67 IHC assessment results were dichotomized with a seventh decile cut-off (Ahlin et al., 2007).

RESULTS

Publication I

Automated counting was performed on the 42 cases, covering an average of 65 cells per image (median 64, range 21–112) and an average of 113 cells per case (median 94, range 47–254). The number of ROI images per case was one in 21 of the cases, two in 12 of the cases, three in eight of the cases and four in one of the cases. The algorithm developed in publication I for automated FISH signal counting processed a single 1376x1032 pixel image in ten seconds with the described hardware. On average, 500 spots (combined amount of HER2 and CEP17 signals, median 382, range 86–2565) per image or 868 spots (median 657, range 261–4478) per case were detected. According to the automated algorithm, the HER2 gene was amplified (average HER2/CEP17 ratio >2.2) in 16 (38%) out of the 42 informative cases. The range of the HER2/CEP17 ratio was from 0.42 to 8.55, with a mean of 2.56 and a standard deviation (SD) of 1.70.
In the visual counting the HER2 gene was amplified in 17 cases (40%). The range of the visual ratio was from 0.45 to 8.11, with a mean of 2.68 and a SD of 1.80.
Considering the visual counting result as a reference with the cut-off for amplification set at 2.0, the sensitivity of the automated method was 89%, the specificity 100%, and the accuracy 95%. The corresponding AUC was 0.97 (95% confidence interval from 0.92 to 1.00). The correlation coefficient of the automated and visual HER2/CEP17 ratios was 0.98.
When the ratios were analyzed in three categories (non-amplified, borderline and amplified), the percentage agreement was 90% and the Cohen's kappa value was 0.82 (Table 1).
The algorithm has been made available as a server-side application at the Webmicroscope website (http://www.webmicroscope.net/fishj). A view of the user interface is provided in figure 4. Alternatively, the algorithm can be downloaded from the website as an add-on macro for ImageJ.

Publication II

According to the computer vision analysis, the range of extent of staining for Ki-67 was from 0 to 95.3% with a mean value of 8.8% and a median value of 4.3%. For automated intensity assessment, the range was 33.9-67.0%, the mean intensity 52.5% and median intensity 53.2%. Two hundred fifty-one patients had duplicate, 21 patients triplicate and seven patients quadruplicate cores available for analysis. The automated analysis of 1648 TMA cores from the 1334 patients took 107 minutes. On average, a single TMA spot was analyzed in 3.9 seconds. The figures 5 and 6 show the graphical user interface used in this publication, and samples of original and result images.

According to the visual assessment of Ki-67 staining, 7.7% of the patients were assigned to the negative expression group, 55.7% to the moderate expression group and 36.7% to the high expression group.

The percentage agreement between the visual and the automated methods was 87% and weighted kappa value 0.57 (Table 1), when the results of the automated assessment of Ki-67 extent of staining were divided into similarly sized groups as the visual Ki-67 results. Example images of different Ki-67 expression levels are shown in figure 7. There were two types of discrepancies between automated and visual assessment of Ki-67 expression: either the results of the automated assessment were lower or higher than the results of the visual assessment. The discrepant cases were reviewed visually. The observed variability occurred mainly between adjacent groups, only two patients with a negative visual score were in the automatically assessed high extent group. These two cases were caused by a partially folded TMA spot and a falsely dyed spot. Generally, the shift from a visually higher group to a lower automatically assessed group was caused by the automated method underestimating the extent of staining in samples with larger stromal areas or with a few strongly positive tumor nuclei. The main causes for shift from a visually lower group to a higher automatically assessed group were out-of-focus samples, debris on the glass slide or positive staining of the tumor cytoplasm.

For analysis of the relationship between the automated assessment of Ki-67 expression and clinicopathological variables (Table 2), the patient series was divided into tertile groups according to the automated assessment of Ki-67: In the low extent group the extent of staining was from 0 to 2.3 percent, in the moderate extent group from 2.4 to 6.3 percent, and in the high extent group from 6.4 to 100 percent. In the low intensity group the intensity range was from 0 to 49.6 percent, in the moderate intensity group from 49.7 to 56.6 percent, and in the high intensity group from 56.7 to 100 percent.

Patients younger than 35 years at the time of diagnosis had a higher automated assessment of Ki-67 extent than those older than 35 years at diagnosis ($P = 0.0008$). Cancers detected within mammography screening had a significantly lower extent of Ki-67 staining than those detected outside of screening ($P = 0.0001$). A higher Ki-67 extent of staining was significantly associated with a larger primary tumor ($P < 0.0001$), and a higher histological grade (a less differentiated tumor) ($P < 0.0001$). A higher Ki-67 extent of staining was also associated with a higher number of metastatic axillary lymph nodes ($P = 0.01$). The distribution of Ki-67 extent of staining was significantly associated with histological type of breast cancer ($P < 0.0001$): higher extent was seen in ductal carcinomas, whereas lower extent of Ki-67 staining was more frequent in lobular carcinomas. Higher Ki-67 extent was significantly associated with molecular markers (all with $P < 0.0001$): negative ER and PR expression, positive HER2 amplification and expression, and high p53 expression were associated with higher extent of Ki-67 staining. In the analysis of molecular Luminal A was associated with lower Ki-67 extent ($P < 0.0001$).

Higher extent of Ki-67 staining was significantly associated with a decrease in distant disease-free survival (DDFS) ($P < 0.0001$). Compared to patients in the low extent group, patients in the moderate extent group had a hazard ratio of 1.77 (95% confidence interval (CI) 1.31-2.37) for distant recurrence, and those in the high extent group had a hazard ratio of 2.34 (95% CI 1.76-3.10). The 5- and 10-year DDFS for the low extent group was 89% and 81%, for the moderate extent group 77% and 71%, and for the high extent group 69% and 64%, respectively.

In a subgroup analysis according to histological grade, patients with grade 1 tumors (well differentiated) in the moderate extent group had a hazard ratio of 1.05 (95% CI 0.35-3.22) and patients in the high extent group a hazard ratio of 4.63 (95% CI 1.83-11.75), as compared to patients in the low extent group. For patients with grade 2 tumors (moderately differentiated), the corresponding figures were 1.51 (95% CI 0.95-2.42) and 1.99 (95% CI 1.26-3.15), and in the grade 3 tumor (poorly differentiated) subgroup, the results were 1.37 (95% CI 0.69- 2.73) and 0.97 (95% CI 0.51-1.84). In the ductal breast cancer subgroup, the corresponding hazard ratios were 1.83 (95% CI 1.30-2.58) and 2.30 (95% CI 1.66-3.18), and in the lobular carcinoma subgroup, the figures were 1.40 (95% CI 0.67-2.93) and 2.29 (95% CI 1.06-4.95).
When extent of Ki-67 staining was determined visually, patients with a moderate Ki-67 expression had a hazard ratio of 1.41 (95% CI 0.83-2.39) for distant recurrence, and those with a high Ki-67 expression had a hazard ratio of 2.58 (95% CI 1.52-4.37), as compared to patients with a negative Ki-67 expression.

There were no statistically significant differences between the low and moderate intensity groups according to Ki-67 staining. Compared to the low intensity group, the high intensity group was associated with a less favourable distant disease-free survival (HR = 1.34, 95% CI 1.04-1.73).
In a multivariate survival analysis, adjusted for tumor size, the number of positive lymph nodes, histological grade, method of detection and age at diagnosis, patients with a moderate Ki-67 extent of staining had a hazard ratio of 1.62 (95% CI 1.10-2.39) and those in the high extent group a hazard ratio of 1.73 (95% CI 1.19-2.51), compared to patients in the low extent group.
The visually determined Ki-67 expression and the automated Ki-67 intensity assessment were not retained as significant prognostic factors in multivariate analysis.

Publication III

The LBP/C, Haralick and Gabor texture classifiers were optimized on the validation set by computing the AUC over a set of cost parameter values (c) for the linear SVM classifier for each of the LBP/C-, Haralick- and Gabor descriptors. The selected c values were: 300 for the LBP/C, 2048 for the Haralick and 2 for the Gabor descriptors.

Significant differences between the accuracies for the classifiers based on the LBP/C features and both the Haralick features as well as Gabor filtered images were observed in the testing phase. The AUC for assigning the correct class to the test images for the LBP/C was 0.995 (95% CI 0.991-0.998) (Table 1), for the Haralick 0.976 (95% CI 0.966-0-986) and for the Gabor descriptors 0.981 (95% CI 0.973-0.990). The analysis time for the 720 images in the test set was 99 seconds for the LBP/C algorithm, 47 seconds for the Haralick features, and 145 seconds for the Gabor filtering.

The results for the best performing LBP/C classifier were studied in more detail. In the test set, the sensitivity of the LBP/C classifier was 99% (95% CI 98%-99%) and the specificity was 94% (95% CI 92%-95%). The percentage agreement between the human observer and the LBP/C texture classifier was 97% (kappa value 0.93, $P<0.0001$, Table 1). The average SVM score in the epithelial images was 1.73 (SD 0.89, range from -2.3 to 3.8) and in the stromal images -2.37 (standard deviation (SD) 1.16, range from -5.6 to 1.3). Out of the 425 epithelial images, 364 were strongly and 42 weakly assigned to the epithelium category, and 19 images were wrongly classified as stroma. Out of the 295 stromal images, 263 were strongly and 28 weakly assigned to the correct class, and 4 images were wrongly classified as epithelium.

To visualize the result of the LBP/C texture analysis method when processing larger areas of tissue, a whole TMA with 73 colorectal tumor tissue spots was analyzed and can be viewed at http://fimm.webmicroscope.net/supplements/epistroma.

A sample with original and result images is provided in figure 8.

Publication IV

Using the losslessly compressed and non-scaled images, the percentage agreement between the visual and automated methods for Ki-67 proliferation index assessment was 85 % and the kappa value 0.64. For the colorectal cancer series, the percentage agreement between the visual annotations and automated segmentation method was 97 % and the kappa value 0.93.

Both the compression and scaling reduced file sizes significantly. Sample images of the effect of image compression and scaling on observed image quality are presented in figures 9 and 10.

In the automated IHC quantification, over 98% of percentage agreement was observed with combined compression ratios up to 1:50 and scaling down to 1:8, with the results from losslessly compressed and non-scaled images as a reference. Corresponding kappa values exceed 0.96. The results suggest a high level of agreement between these compression and scaling levels.

The corresponding results for the automated tumor segmentation showed percentage agreements of over 97% and kappa values of over 0.93 with compression ratios up to 1:25. However, the compression ratio 1:50 resulted in a less favorable percentage agreement of 86% and a kappa value of 0.71.

Table 1: Summary of Kappa values, AUC, image sizes and analysis times from publications I-III.

Publication	I (HER2 FISH)	II (Ki-67 IHC)	III (Texture)
Kappa value*	0.82	0.57	0.93
AUC*	0.97	NA**	0.995
Image size in pixels	1376x1023	1634x1634	524x523***
Average analysis time per image	10 s	3.9 s	0.1 s

* Visual versus automated analysis
** AUC not available for publication II due to original tripartite visual classification
*** Images in publication III varied in size, median size reported.

Table 2: Association of automated assessment of Ki-67 extent of staining with clinicopathological characteristics (green denotes association with low extent, red with high extent)

Characteristic	Value	P
Age at diagnosis (y)	< 35	0.0008
Method of detection	Mammography screening	0.0001
Primary tumor diameter (cm)	< 1 > 2	< 0.0001
No. of positive axillary nodes	>= 1	0.0142
Histological grade	1 2-3	< 0.0001
Histological type	Lobular Ductal	< 0.0001
ER expression	Positive Negative	< 0.0001
PR expression	Positive Negative	< 0.0001
HER2 amplification	Positive	< 0.0001
HER2 expression	Positive	< 0.0001
p53 expression	Low Moderate-High	< 0.0001
Molecular subtype	Luminal A Basal HER2+/ER-/PR- Luminal B Five-marker negative Unclassified	< 0.0001

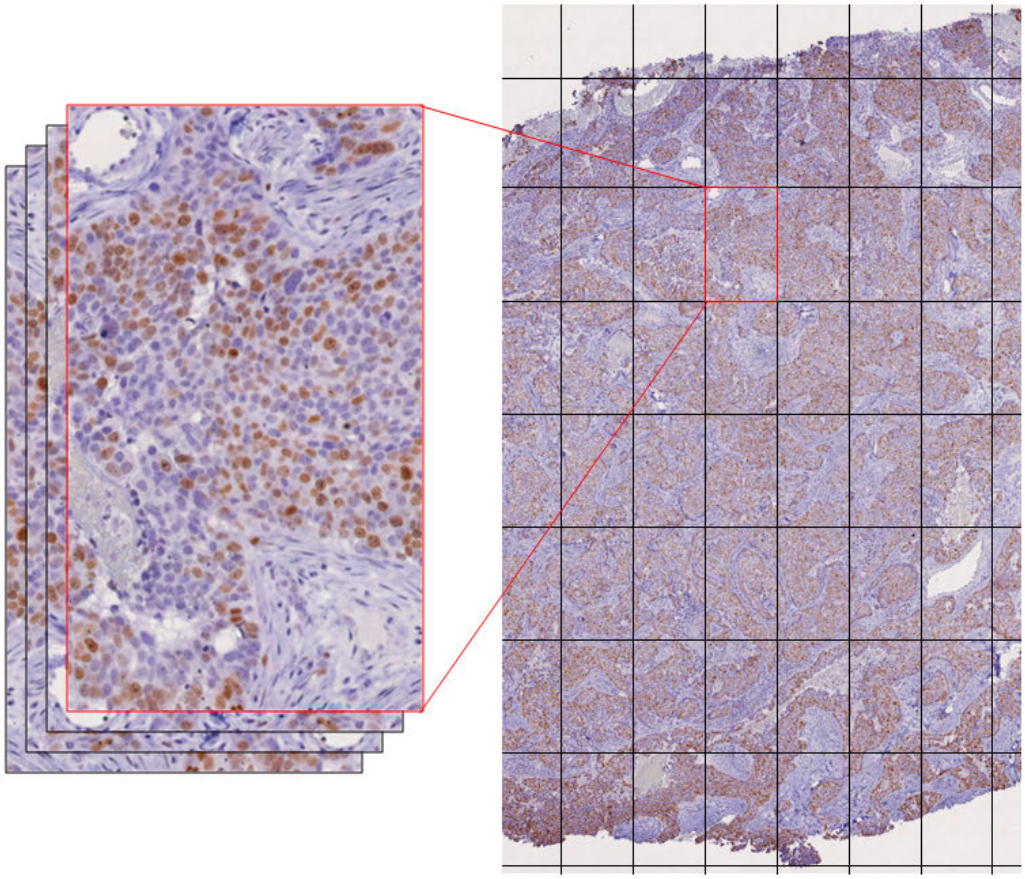

Figure 1: The single snapshot images (on the left) are stitched together to form a virtual slide (on the right)

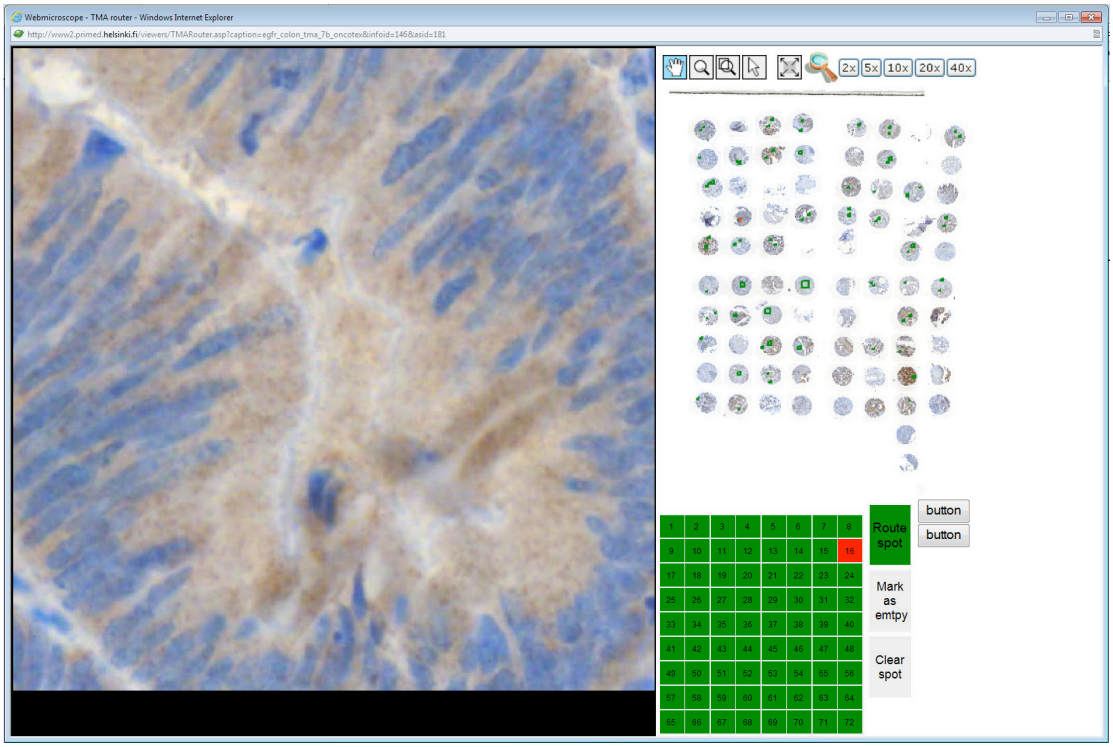

Figure 2: Annotation tool with annotated areas shown as green rectangles on the overview image (top right)

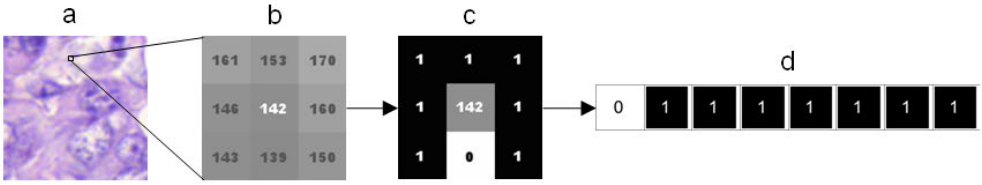

Figure 3: The LBP algorithm (original image (a), 8 pixel neighborhood with pixel values (b), binarised neighborhood (c) and final binary code (d))

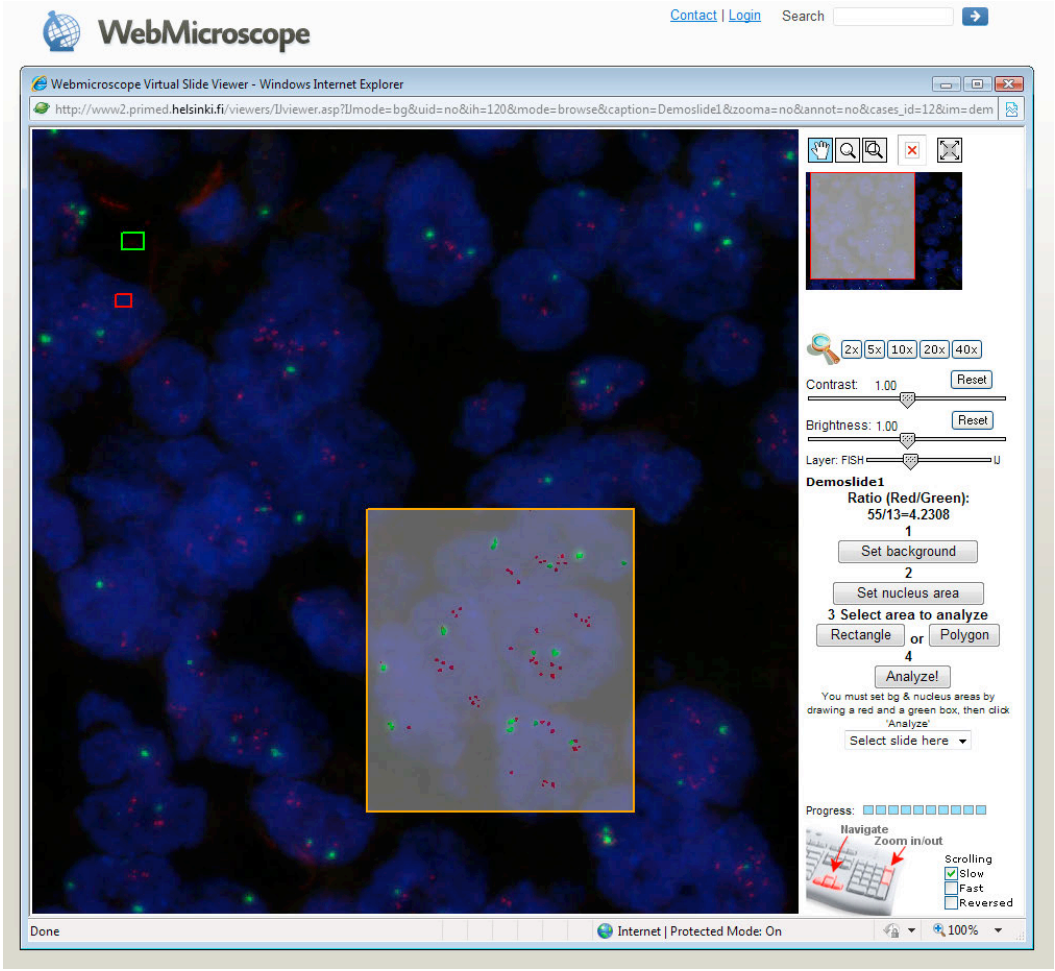

Figure 4: A graphical user interface of the automated FISH algorithm developed in publication I with a HER2 amplified breast cancer sample

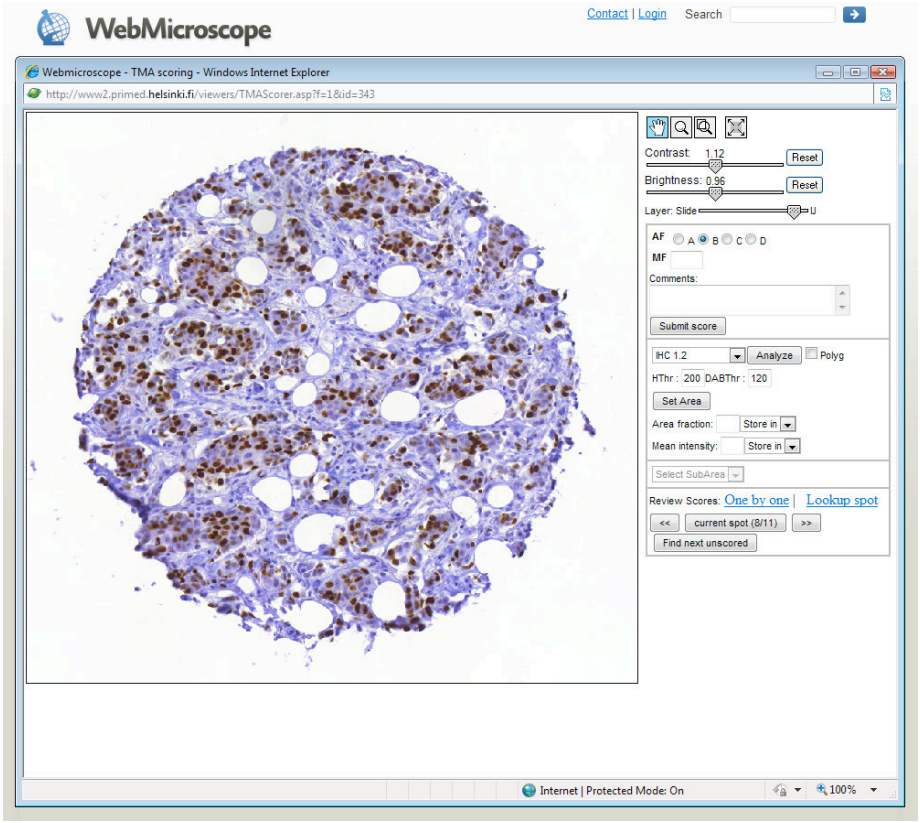

Figure 5: A graphical user interface of the automated IHC algorithm developed in publication II, original Ki-67 breast cancer TMA spot

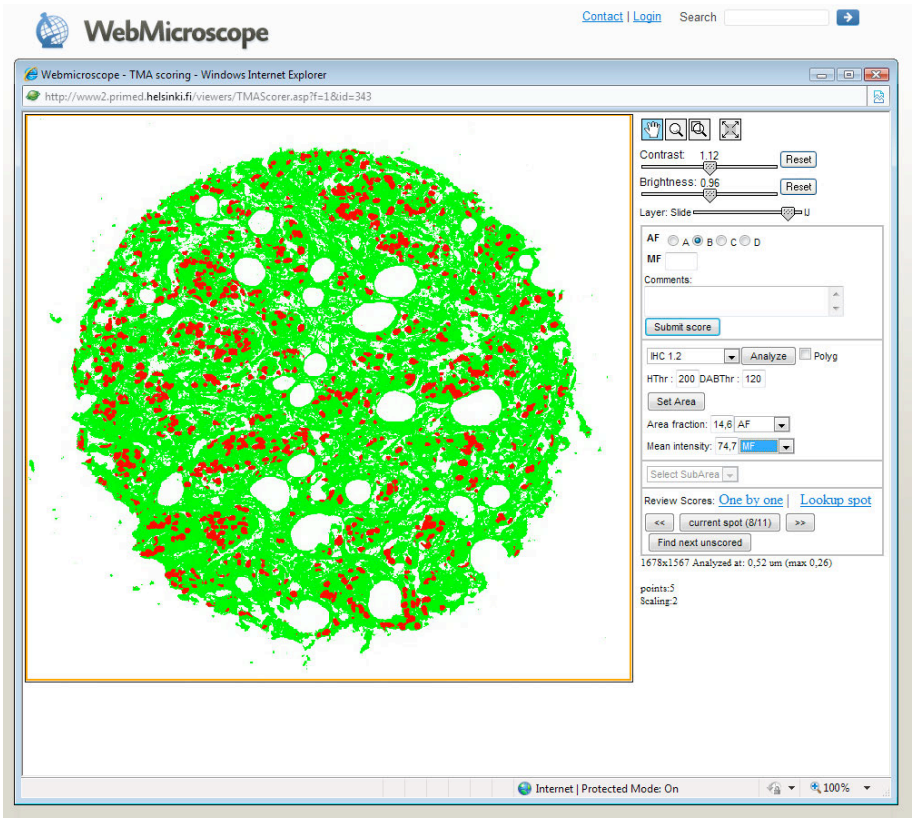

Figure 6: A sample result image of the spot shown in figure 5

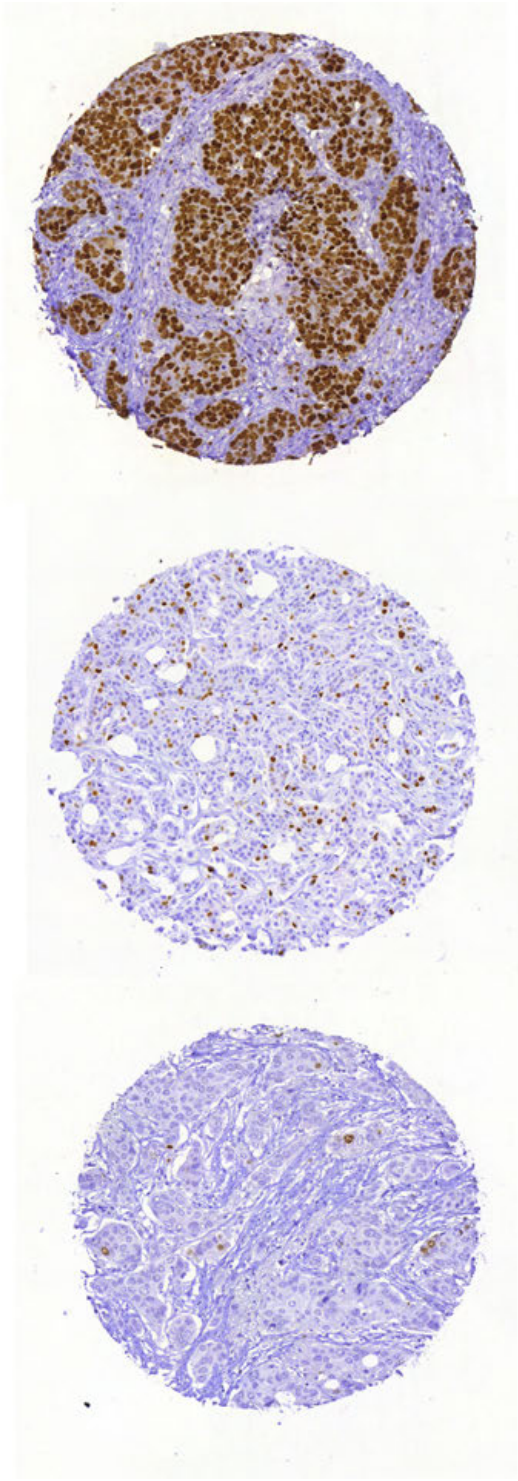

Figure 7: Sample images of different Ki-67 expression levels (high extent on the top, moderate extent on the middle and low extent on the bottom)

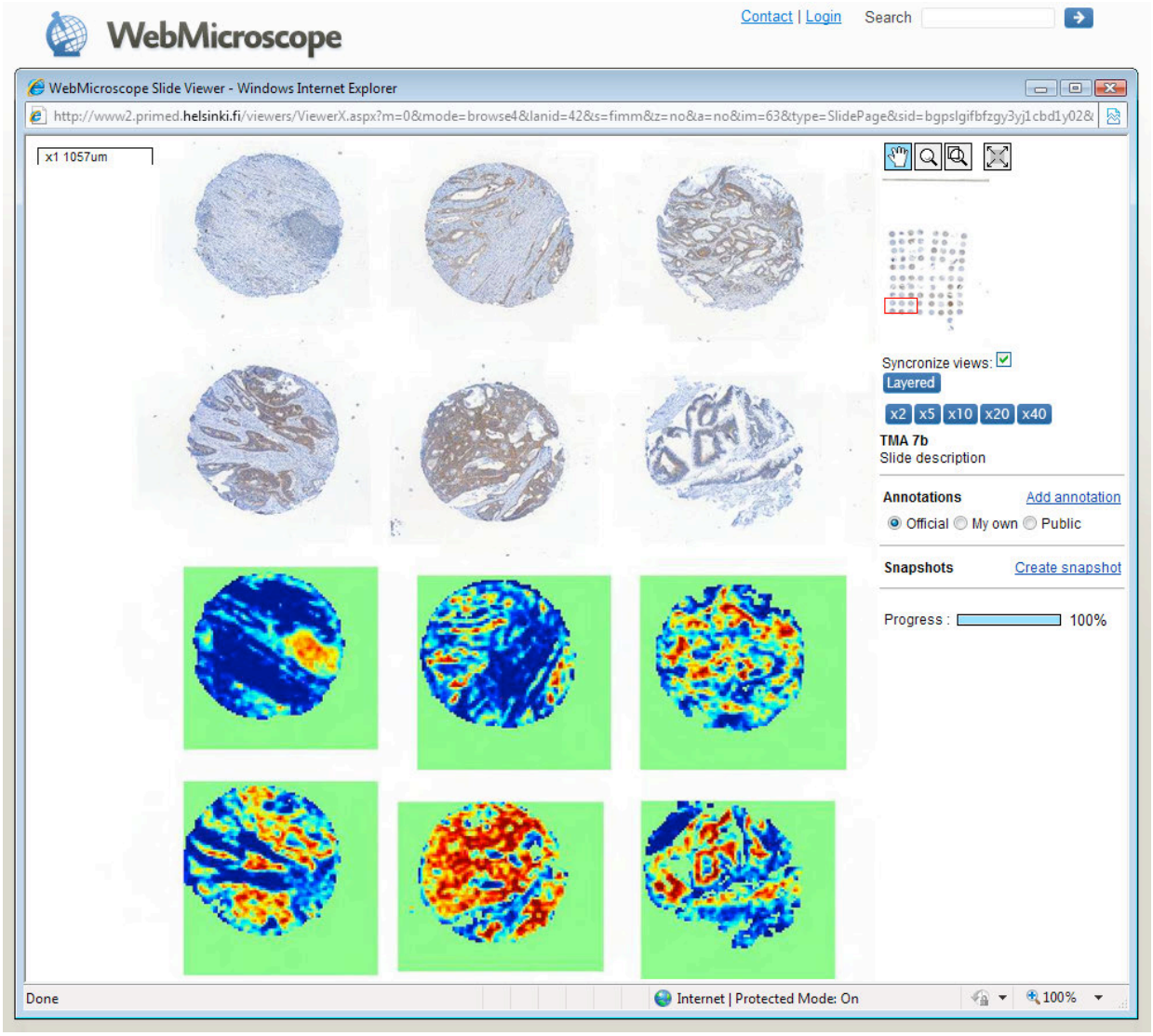

Figure 8: A sample of original and result images from publication III, showing automated segmentation of epithelial (red) and stromal (blue) tissue compartments of colorectal cancer TMA spots

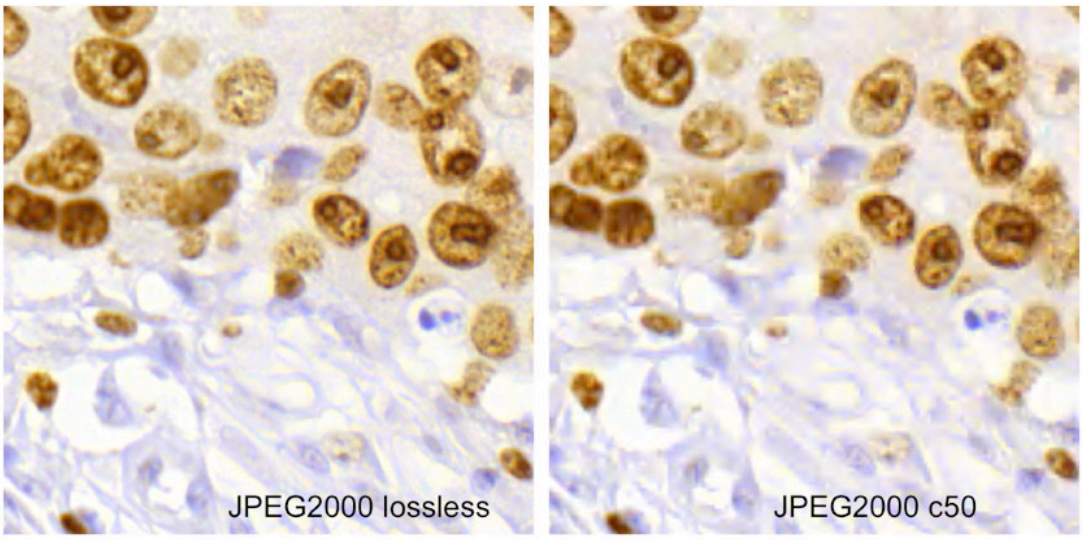

Figure 9: Sample images of the effect of image compression used in publication IV on visual image quality, Ki-67 breast cancer sample

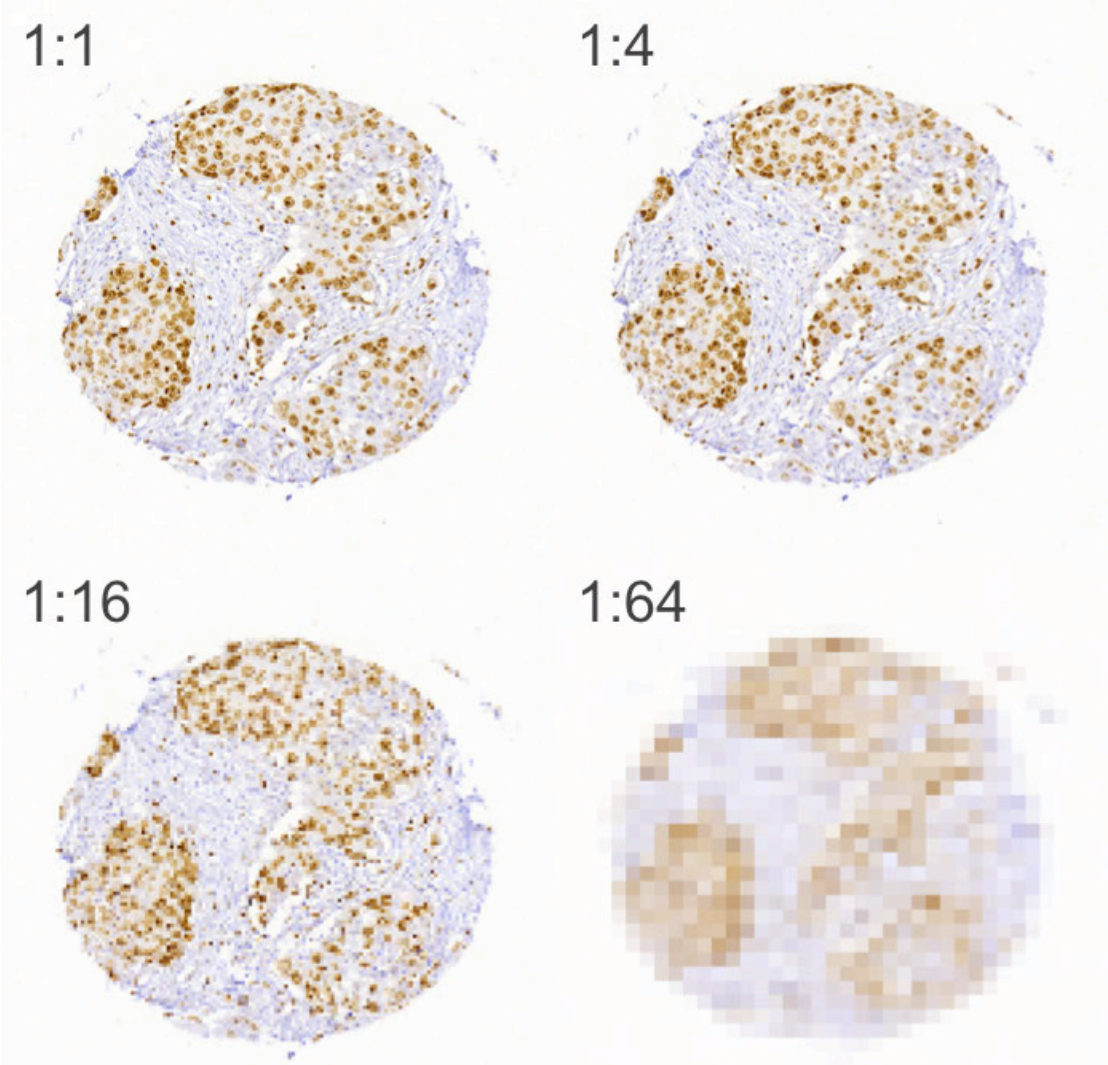

Figure 10: Sample images of the effect of image scaling used in publication IV on visual image quality, Ki-67 breast cancer sample

DISCUSSION

The automation of high-throughput quantitative analysis of digitized microscopy images of cancer tissue proved to be feasible and the results are comparable to the results from the conventional visual analyses. The biomarker research community already uses the released algorithms.

In publication I, the kappa value was 0.82, corresponding to almost perfect agreement between visual and automated assessment of HER2 gene copy number by FISH. Comparable statistics from other publications related to this issue are scarce. A correlation coefficient (r=0.92) was reported in a study with 41 breast cancer patients, but agreement figures were not given (Tubbs et al., 2006). The reported correlation coefficient is lower than the correlation coefficient 0.98 in publication I. In 1–2% of breast cancer cases, spatially heterogeneous HER2 amplification has been reported (Bartlett et al., 2001). In these cases, FISH signals need to be evaluated in a larger number of microscopic fields (Hicks and Tubbs, 2005). Compared to conventional epifluorescence microscopy, proper selection of ROIs may be easier using whole-slide imaging and virtual microscopy, since an overview of the whole specimen is available. Moreover, re-scoring of specimens is possible without the fading of fluorescence using this approach. The developed automated assessment method could be applied to other FISH studies than HER2 gene copy number enumeration in a straightforward manner. Modifications could be made to cover also other *in situ* hybridization methods such as silver and chromogenic *in situ* hybridization.

Considering publication II, there is no consensus for Ki-67 cut-off values in the literature, some studies use arbitrary values, some the median, and some divide data according to tertiles (de Azambuja et al., 2007; Stuart-Harris et al., 2008). In a publication on 265 breast cancer patients, univariate survival analysis was carried out with different Ki-67 thresholds to aid in selection of the cut-off value (Jalava et al., 2006). With a selected cut-off point of 15%, subsequent multivariate analysis resulted in a hazard ratio of 2.02 for visually assessed Ki-67 proliferation index. This figure is in 95% confidence interval of the results from

automated Ki-67 assessment in our study, although the multivariate model was slightly different and the patient series was divided into two groups instead of tertiles. Also the association of the automated assessment of Ki-67 extent of staining and clinicopathological factors was in line with previous reports for visual assessment (Ahlin et al., 2007; Sihto et al., 2008).

The automated and visual methods for Ki-67 IHC assessment in publication II showed only moderate agreement with a kappa value of 0.57, but both were significant predictors of distant disease-free survival in univariate analyses with hazard ratios similar to the previous meta-analyses, where hazard ratios of 1.93-2.18 for disease-free survival (DFS) have been observed (de Azambuja et al., 2007; Stuart-Harris et al., 2008). However, in multivariate survival analyses, only the automated method remained as a significant predictor of patient outcome. These observations appear to contradict the custom of using the visual method as gold standard.

Visual interpretation has shown to result in variability due to human errors (Camp et al., 2002; Cregger et al., 2006). To reduce the variation, a consensus of two or more experts with a critical review of discrepancies would be needed as gold standard in most of the studies. In high throughput research setting with thousands of samples, lack of resources often limits the possibility to use consensus of multiple experts. Instead of trying to mimic experts' results, methods could be rated at how they can predict some clinically valid endpoint, such as patient outcome.

The prognostic value of an automated image analysis algorithm is especially relevant in routine breast cancer diagnostic pathology, where results are subsequently used in clinical decisions on the treatment of patients. TMAs with samples from dozens of different patients on a single glass slide were used to study Ki-67 expression in publication II, whereas whole slide sections are predominantly used in routine diagnostic pathology. The algorithm used in publication II can be extended to selected ROIs from whole slide sections, or even applied to the whole section itself. The usage of multicore TMA slides in routine assessment of prognostic factors in breast cancer has also been suggested with results comparable to whole slide diagnostics (Sapino et al., 2006).

There are a few drawbacks in the automated Ki-67 assessment algorithm described in publication II. The first weakness is the manual adjustment of threshold levels needed prior to the batch analysis, where multiple TMA spots are analyzed automatically in series. However, when all the studied specimens were processed and stained with same protocols, a constant threshold was used successfully. Optionally, a method for automated threshold selection could be developed to cover also the possible variations in the sample staining protocols. The achieved results should always be visually reviewed after the analysis as a quality control for the selected threshold.

The automatically assessed intensity of Ki-67 staining was of only limited prognostic value. The previously reported non-linear relationship between the amount of antigen and the DAB staining intensity could explain this observation (Fritz et al., 1995). Another weakness of the current computer vision approach is that the unsegmented image of the tissue specimen is analyzed and possible stromal components are included in the automated assessment. The tumor grade and histological type can affect the ratio of stroma to epithelium in the TMA cores. This could have affected the results in publication II. Therefore a subgroup analysis according to histological grade and type was performed, as well as adjustment for these possible confounding factors in a multivariate survival model. The automated assessment of Ki-67 extent of staining was a significant prognostic factor in all subgroups, except in the poorly differentiated (i.e. grade 3) tumors. This lack of prognostic value of Ki-67 in grade 3 tumors has been previously reported (Klintman et al., 2010). The stromal part of the tumors could be excluded by using a preprocessing step with a tissue classifier, such as the one described in publication III. Commercial image analysis software systems have already incorporated similar preprocessors for stromal exclusion (Turbin et al., 2008; Turashvili et al., 2009). Also the stromal expression of proteins has shown to be of clinical value (Finak et al., 2008; Farmer et al., 2009; Pietras and Ostman, 2010). Thus, in addition to exclusion of stroma, a tissue classifier can be used to analyze the stroma compartment separately. A tissue classifier could also aid in selection of cells and tissue areas for laser capture microdissection (Emmert-Buck et al., 1996).

In publication III, the algorithm for texture classification of colorectal cancer specimens based on LBP/C features showed significantly higher classification accuracy as compared to the algorithms based on either Haralick or Gabor features. In addition to the rotation and grayscale invariance of the LBP/C features, which may explain partially this phenomenon, Gabor features may have suffered from previously reported tendency to favor low-frequency components over high-frequency components abundant in natural images (Field, 1987). This low-pass filtering smoothens the details in the images, and small but possibly important structures for the correct classification of the cancer tissue may be lost. The direct comparison of the algorithm from publication III with results from previous studies is challenging, mostly due to different tissue materials used in the studies. To enable other researchers to compare their algorithms with the one we constructed, all the images used in publication III were published as a database available online. Also the selection of the features and the classifier, and the selected parameters, makes a direct comparison difficult. The closest comparable study using LBP was based on 43 whole-slide neuroblastoma samples (Sertel et al., 2009). A binary classification was used to segment the samples into either stroma-rich or stroma-poor classes. The reported overall classification accuracy was 88%, compared to over 99% according to our results.

The classification in our study was based on square image blocks of approximately 40 micrometers, and the algorithm may have overlooked a few cancerous cells scattered between stromal structures. The algorithm was trained with colorectal cancer samples stained for analysis of EGFR expression. Other types of tumors and stainings can be used for algorithm training, and it would be of interest to see whether algorithms for segmentation of tumor tissue into stromal and epithelial compartments can have more general use in histology as previously suggested (Kayser et al., 2006). The prognostic value of this algorithm remains to be further studied; promising results from similar algorithms have already been reported (Beck et al., 2011). Texture based algorithms have also been applied in quality comparison between different microscopic imaging solutions (Walkowski and Szymas, 2011).

The publication IV studying the effect of image size reduction on the results of the automated image analysis algorithms suggest that significant amount of data storage space can be saved by using image compression and scaling without compromising the automated assessment results. With the studied algorithms, images can be reduced to less than four percent of the original image sizes. These results are in line with a study where maximum JPEG2000 compression levels (1:46) resulted in good analysis results in low complexity images, but images with higher complexity tolerated only medium JPEG2000 compression levels (1:23) (Lejeune et al., 2011).

As described in the literature (Mulrane et al., 2008; Kayser et al., 2009), the workflow of the pathology laboratories will move towards an increasingly digital environment similar to the digitization that already has occurred in radiology. The pieces of software developed within this thesis are tools that in combination with virtual microscopy have the potential to provide faster, more objective and less laborious sample analysis in histological pathology. Also, inter- and intra-observer reproducibility of the results will be enhanced. Automated quantification of FISH signals can be of clinical value in determining which breast cancer patients would be eligible for HER2 targeted monoclonal antibody therapy. By applying this therapy only to patients with HER2 gene amplification, the treatment can be targeted to patients that are expected to benefit and adverse effects in HER2 negative patients can be avoided. Rational selection of patients for therapy will also have considerable economical effects.

A group of users that could take advantage of high throughput automated image analysis is researchers studying tissue biomarkers. The number of samples is usually high and there is a need for automated solutions for fast translational analysis of clinical correlates. As an example, a slightly modified version of the algorithm created in publication II was already used in automated analysis of more than two thousand patient samples (Sahu et al., 2011). The readout of an IHC staining was available for statistical analysis within 24 hours. Similarly, the automated scoring of HER2 amplification of 36456 FISH spots was performed in 15 minutes, whereas the corresponding visual assessment required an equal

amount of mouse clicks for the publication I with only 42 patients involved. This visual scoring was performed by the author in multiple sessions over a period of two weeks.

Table 1 summarizes key figures from publications I-III. Although not directly comparable due to different algorithms, image sizes and hardware used, the analysis time per image has drastically reduced from 10 seconds in publication I to 0.1 seconds in publication III. The processing power of personal computers continues to improve, but also other options such as grid and cloud computing are emerging. In the future, this trend will probably enable even more complicated algorithms to be run real-time on gigapixel size virtual microscopy images.

CONCLUSIONS

In this thesis, tools for high-throughput biomarker research in a digital microscopy environment were created and evaluated. In publication I, the algorithm for automated quantitative assessment of FISH signals to determine the HER2 gene amplification status in breast cancer tissue images proved to be comparable to visual scoring. In publication II, the extent of Ki-67 staining determined in breast cancer tissue images by the automated algorithm was a significant predictor of patient outcome in both uni- and multivariate analyses. In publication III, the automated segmentation tool divided the colorectal cancer images into epithelial and stromal histological classes with high accuracy. In publication IV, image compression and scaling led to significant reductions in image sizes without compromising the results of the algorithms introduced in publications II and III.

The already released algorithms developed in this thesis are freely accessible to be used by the research community, facilitating also the external validation of the algorithms. After further validation studies, the algorithms can potentially be applied in clinical pathology – especially within risk prediction, diagnostics and targeted treatment of cancer patients in a personalized medicine setting.

ACKNOWLEDGEMENTS

This work was carried out at the Folkhälsan Research Center during the years 2004-2009, and at the Institute for Molecular Medicine Finland (FIMM) during the years 2009-2012. I thank the Directors of Folkhälsan Research Center and FIMM, Prof. Anna-Elina Lehesjoki and Prof. Olli Kallioniemi for the superb environments to conduct scientific research, and all the staff at FIMM for creating an exciting work place. Financial support was also provided by Helsinki Biomedical Graduate School, the Finnish Cancer Society, the Instrumentarium Foundation for Science, Finska Läkaresällskapet, the Sigrid Juselius Foundation, Medicinska Understödsföreningen Liv och Hälsa, and the Dorothea Olivia, Karl Walter, and Jarl Walter Perklén Foundation.

Completing this thesis has been a challenging but rewarding task. It would not have been possible without the help and advice from my teachers, co-workers, collaborators, friends and family.

First, I thank my supervisor, Docent Johan Lundin, for accepting me into his lab, for having the patience to teach me, and for allowing me to take part into many innovative scientific projects. Johan is also extremely flexible making it possible to successfully combine research, medical studies and clinical work. Although being a busy man as a Research Director of FIMM, Johan is always available for discussions and consultation, and his knowledge and fresh ideas are a constant source of motivation to me.

Second, my sincere gratitude should also be directed to Docent Pauliina Kronqvist and Acting Professor Jari Hannuksela for rapid review of the thesis, and for Professor Janusz Szymas for accepting the invitation to be the opponent.

The scientific research would be nearly impossible without collaborators. I wish to thank especially Prof. Jorma Isola from the University of Tampere for his remarkable contributions to my work. Vilppu Tuominen, Mervi Jumppanen, Arttu Viitanen, Heikki Joensuu, Harri Sihto, Kaija Holli, Taina Turpeenniemi-Hujanen, Vesa Kataja, Liisa Sailas, Esa Rahtu, Stig Nordling, Caj Haglund, Timo Ahonen, Matti Pietikäinen, Carl Blomqvist, Heli Nevanlinna and Kirsimari Aaltonen are also thanked for their collaboration.

I have been privileged to work in a multidisciplinary and multicultural research group set up by Johan. Mikael Lundin is the person responsible for setting up and maintaining Webmicroscope platform, Nina Linder has been kind enough to guide me through the practicalities of this thesis project, and Tiina Lehtimäki has been a wonderful peer in the process. Riku Turkki and Ville Ojansivu have brought their technical knowledge and engineering point-of-view into our previously medically oriented group, and the new members of our group, Margarita Walliander and Anne Grote, have already succeeded to influence the way our group functions. I wish to thank all the group members for their support and great discussions.

Outside the scientific community, I have been fortunate to have numerous friends with various backgrounds: the Olari high school posse, the physicists from Otaniemi, the physicians from Meilahti and the guys at Mendor Ltd. I would like to thank all my friends for helping me to relax and to take my mind off the research. Of my friends, Matti Alvesalo, né Saarinen, deserves to be mentioned by name. Apart from being former flatmates, Matti and I studied in the same schools all the way from Olari high school to the medical faculty at Meilahti. Thanks Matti for all the great times we have spent, and many more to come!

I am thankful to my parents Jorma and Seija, my sister Anna, my grandparents Pentti and Riitta, and all my other relatives. I would not be here without your support.

Finally, my deepest gratitude goes to my loving wife Emma and to my beautiful daughters Elsa and Aino. Thank you for tolerating my long hours of work writing this thesis. I love you all!

Helsinki, September 2012

Juho Konsti

REFERENCES

Adiga PS and Chaudhuri BB (1999). *Efficient cell segmentation tool for confocal microscopy tissue images and quantitative evaluation of FISH signals*. Microsc Res Tech **44**(1): 49-68.

Ahlin C, Aaltonen K, Amini RM, Nevanlinna H, Fjällskog ML and Blomqvist C (2007). *Ki67 and cyclin A as prognostic factors in early breast cancer. What are the optimal cut-off values?* Histopathology **51**(4): 491-498.

Ahonen T, Hadid A and Pietikäinen M (2006). *Face Description with Local Binary Patterns: Application to Face Recognition*. IEEE Trans. Pattern Analysis and Machine Intelligence **28**(12): 2037-2041.

Bankfalvi A, Boecker W and Reiner A (2004). *Comparison of automated and manual determination of HER2 status in breast cancer for diagnostic use: a comparative methodological study using the Ventana BenchMark automated staining system and manual tests*. Int J Oncol **25**(4): 929-935.

Barker NJ, Zahurak M, Olson JL, Nadasdy T, Racusen LC and Hruban RH (1998). *Digital imaging of black and white photomicrographs - Impact of file size*. Am. J. Surg. Pathol. **22**(11): 1411-1416.

Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P, Richmond J, Donald B, Ferrier R and Cooke TG (2001). *Evaluating HER2 amplification and overexpression in breast cancer*. J Pathol **195**(4): 422-428.

Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, van de Vijver MJ, West RB, van de Rijn M and Koller D (2011). *Systematic analysis of breast cancer morphology uncovers stromal features associated with survival*. Science translational medicine **3**(108): 108ra113.

Bloom K and Harrington D (2004). *Enhanced accuracy and reliability of HER-2/neu immunohistochemical scoring using digital microscopy*. Am J Clin Pathol **121**(5): 620-630.

Camp RL, Chung GG and Rimm DL (2002). *Automated subcellular localization and quantification of protein expression in tissue microarrays*. Nat Med **8**(11): 1323-1327.

Chen W, Reiss M and Foran DJ (2004). *A prototype for unsupervised analysis of tissue microarrays for cancer research and diagnostics*. IEEE Trans Inf Technol Biomed **8**(2): 89-96.

Chung GG, Zerkowski MP, Ghosh S, Camp RL and Rimm DL (2007). *Quantitative analysis of estrogen receptor heterogeneity in breast cancer*. Lab Invest **87**(7): 662-669.

Clausi DA (2001). *Improved texture recognition of SAR sea ice imagery by data fusion of MRF features with traditional methods*. Int Geosci Remote Se: 1170-1172.

Cohen J (1968). *Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit*. Psychol Bull **70**(4): 213-220.

Coons AH, Creech HJ, Jones RN and Berliner E (1942). *The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody*. J Immunol **45**: 159-170.

Corletto V, Verderio P, Giardini R, Cipriani S, Di Palma S and Rilke F (1998). *Evaluation of residual cellularity and proliferation on preoperatively treated breast cancer: a comparison between image analysis and light microscopy analysis*. Anal Cell Pathol **16**(2): 83-93.

Cregger M, Berger AJ and Rimm DL (2006). *Immunohistochemistry and quantitative analysis of protein expression*. Arch Pathol Lab Med **130**(7): 1026-1030.

Dahlman A, Rexhepaj E, Brennan DJ, Gallagher WM, Gaber A, Lindgren A, Jirström K and Bjartell A (2011). *Evaluation of the prognostic significance of MSMB and CRISP3 in prostate cancer using automated image analysis*. Mod Pathol.

de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, Sotiriou C, Larsimont D, Piccart-Gebhart MJ and Paesmans M (2007). *Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients*. Br J Cancer **96**(10): 1504-1513.

Demir C, Gultekin SH and Yener B (2005). *Augmented cell-graphs for automated cancer diagnosis*. Bioinformatics **21 Suppl 2**: ii7-12.

Di Cataldo S, Ficarra E, Acquaviva A and Macii E (2010). *Achieving the way for automated segmentation of nuclei in cancer tissue images through morphology-based approach: A quantitative evaluation*. Comput Med Imag Grap **34**(6): 453-461.

Diamond J, Anderson NH, Bartels PH, Montironi R and Hamilton PW (2004). *The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia*. Hum Pathol **35**(9): 1121-1131.

Domagala W, Markiewski M, Harezga B, Dukowicz A and Osborn M (1996). *Prognostic significance of tumor cell proliferation rate as determined by the MIB-1 antibody in breast carcinoma: its relationship with vimentin and p53 protein*. Clin Cancer Res **2**(1): 147-154.

Doyle S, Agner S, Madabhushi A, Feldman M and Tomaszewski J (2008). *Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features*. IEEE Xplore: 496-499.

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA and Liotta LA (1996). *Laser capture microdissection*. Science **274**(5289): 998-1001.

Enroth-Cugell C and Robson JG (1966). *The Contrast Sensitivity of Retinal Ganglion Cells of the Cat*. J Physiol-London **187**(3): 517-552.

Entwistle A (2003). *Formats of image data files that can be used in routine digital light micrography. Part two*. Biotech Histochem **78**(2): 91-99.

Eramian M, Daley M, Neilson D and Daley T (2011). *Segmentation of epithelium in H&E stained odontogenic cysts*. J Microsc **244**(3): 273-292.

Eroschenko VP (2005). Atlas of histology. Baltimore, MD 21210, Lippincott Williams & Wilkins.

Fan RE, Chang KW, Hsieh CJ, Wang XR and Lin CJ (2008). *LIBLINEAR: A library for large linear classification.* The Journal of Machine Learning Research **9**: 1871–1874.

Faratian D, Munro A, Twelves C and Bartlett JM (2009). *Membranous and cytoplasmic staining of Ki67 is associated with HER2 and ER status in invasive breast carcinoma*. Histopathology **54**(2): 254-257.

Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, Andre S, Piccart M, Campone M, Brain E, Macgrogan G, Petit T, Jassem J, Bibeau F, Blot E, Bogaerts J, Aguet M, Bergh J, Iggo R and Delorenzi M (2009). *A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer*. Nat Med **15**(1): 68-74.

Felten CL, Strauss JS, Okada DH and Marchevsky AM (1999). *Virtual microscopy: High resolution digital photomicrography as a tool for light microscopy simulation*. Hum Pathol **30**(4): 477-483.

Field DJ (1987). *Relations between the Statistics of Natural Images and the Response Properties of Cortical-Cells*. J Opt Soc Am A **4**(12): 2379-2394.

Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M and Park M (2008). *Stromal gene expression predicts clinical outcome in breast cancer*. Nat Med **14**(5): 518-527.

Florea F, Vertan C, Rogozan A, Bensrhair A and Darmoni S (2005). *Comparison of histogram-based feature sets for meical image modality categorization*. IEEE: 47-50.

Fritz P, Wu X, Tuczek H, Multhaupt H and Schwarzmann P (1995). *Quantitation in immunohistochemistry. A research method or a diagnostic tool in surgical pathology?* Pathologica **87**(3): 300-309.

Gonzalez RC and Woods RE (2002). Digital image processing. Upper Saddle River, New Jersey, Prentice-Hall, Inc.

Grigoryan AM, Dougherty ER, Kononen J, Bubendorf L, Hostetter G and Kallioniemi O (2002a). *Morphological spot counting from stacked images for automated analysis of gene copy numbers by fluorescence in situ hybridization*. J Biomed Opt **7**(1): 109-122.

Grigoryan AM, Hostetter G, Kallioniemi O and Dougherty ER (2002b). *Simulation toolbox for 3d-fish spot-counting algorithms*. Real-Time Imaging **8**: 203-212.

Hamilton PW, Bartels PH, Thompson D, Anderson NH, Montironi R and Sloan JM (1997). *Automated location of dysplastic fields in colorectal histology using image texture analysis*. J Pathol **182**(1): 68-75.

Haralick R, Shanmugam K and Dinstein I (1973). *Textural features for image classification*. IEEE transactions on systems, man and cybernetics **3**: 610-621.

Hassan S, Ferrario C, Mamo A and Basik M (2008). *Tissue microarrays: emerging standard for biomarker validation*. Curr Opin Biotechnol **19**(1): 19-25.

Hicks DG and Tubbs RR (2005). *Assessment of the HER2 status in breast cancer by fluorescence in situ hybridization: a technical review with interpretive guidelines*. Hum Pathol **36**(3): 250-261.

Hwang HG, Choi HJ, Lee BI, Yoon HK, Nam SH and Choi HK (2005). *Multi-resolution wavelet-transformed image analysis of histological sections of breast carcinomas*. Cell Oncol **27**(4): 237-244.

Iourov IY, Soloviev IV, Vorsanova SG, Monakhov VV and Yurov YB (2005). *An approach for quantitative assessment of fluorescence in situ hybridization (FISH) signals for applied human molecular cytogenetics*. J Histochem Cytochem **53**(3): 401-408.

Jalava P, Kuopio T, Juntti-Pätinen L, Kotkansalo T, Kronqvist P and Collan Y (2006). *Ki67 immunohistochemistry: a valuable marker in prognostication but with a risk of misclassification: proliferation subgroups formed based on Ki67 immunoreactivity and standardized mitotic index*. Histopathology **48**(6): 674-682.

Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D (2011). *Global cancer statistics*. CA Cancer J Clin **61**(2): 69-90.

Joensuu H, Isola J, Lundin M, Salminen T, Holli K, Kataja V, Pylkkänen L, Turpeenniemi-Hujanen T, von Smitten K and Lundin J (2003). *Amplification of erbB2 and erbB2 expression are superior to estrogen receptor status as risk factors for distant recurrence in pT1N0M0 breast cancer: a nationwide population-based study*. Clin Cancer Res **9**(3): 923-930.

Joensuu H, Kellokumpu-Lehtinen PL, Bono P, Alanko T, Kataja V, Asola R, Utriainen T, Kokko R, Hemminki A, Tarkkanen M, Turpeenniemi-Hujanen T, Jyrkkiö S, Flander M, Helle L, Ingalsuo S, Johansson K, Jääskeläinen AS, Pajunen M, Rauhala M, Kaleva-Kerola J, Salminen T, Leinonen M, Elomaa I and Isola J (2006). *Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer*. N Engl J Med **354**(8): 809-820.

Joensuu H, Lehtimäki T, Holli K, Elomaa L, Turpeenniemi-Hujanen T, Kataja V, Anttila A, Lundin M, Isola J and Lundin J (2004). *Risk for distant recurrence of breast cancer detected by mammography screening or other methods*. JAMA **292**(9): 1064-1073.

Kajtar B, Mehes G, Lorch T, Deak L, Kneifne M, Alpar D and Pajor L (2006). *Automated fluorescent in situ hybridization (FISH) analysis of t(9;22)(q34;q11) in interphase nuclei*. Cytometry A **69**(6): 506-514.

Kalinski T, Zwonitzer R, Grabellus F, Sheu SY, Sel S, Hofmann H, Bernarding J and Roessner A (2009). *Lossy compression in diagnostic virtual 3-dimensional microscopy-where is the limit?* Hum Pathol **40**(7): 998-1005.

Karacali B and Tozeren A (2007). *Automated detection of regions of interest for tissue microarray experiments: an image texture analysis*. BMC Med Imaging **7**: 2.

Kayser K, Gortler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E and Kayser G (2009). *AI (artificial intelligence) in histopathology--from image analysis to automated diagnosis*. Folia Histochem Cytobiol **47**(3): 355-361.

Kayser K, Radziszowski D, Bzdyl P, Sommer R and Kayser G (2006). *Towards an automated virtual slide screening: theoretical considerations and practical experiences of automated tissue-based virtual diagnosis to be implemented in the Internet*. Diagn Pathol **1**: 10.

Klijanienko J, Couturier J, Galut M, El-Naggar AK, Maciorowski Z, Padoy E, Mosseri V and Vielh P (1999). *Detection and quantitation by fluorescence in situ hybridization (FISH) and image analysis of HER-2/neu gene amplification in breast cancer fine-needle samples*. Cancer **87**(5): 312-318.

Klintman M, Bendahl PO, Grabau D, Lovgren K, Malmstrom P and Ferno M (2010). *The prognostic value of Ki67 is dependent on estrogen receptor status and histological grade in premenopausal patients with node-negative breast cancer*. Mod Pathol **23**(2): 251-259.

Kolles H, Forderer W, Bock R and Feiden W (1993). *Combined Ki-67 and Feulgen stain for morphometric determination of the Ki-67 labelling index*. Histochemistry **100**(4): 293-296.

Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G and Kallioniemi OP (1998). *Tissue microarrays for high-throughput molecular profiling of tumor specimens*. Nat Med **4**(7): 844-847.

Landis JR and Koch GG (1977). *Measurement of Observer Agreement for Categorical Data*. Biometrics **33**(1): 159-174.

Langersafer PR, Levine M and Ward DC (1982). *Immunological Method for Mapping Genes on Drosophila Polytene Chromosomes*. P Natl Acad Sci-Biol **79**(14): 4381-4385.

Layfield LJ, Saria EA, Conlon DH and Kerns BJ (1996). *Estrogen and progesterone receptor status determined by the Ventana ES 320 automated immunohistochemical stainer and the CAS 200 image analyzer in 236 early-stage breast carcinomas: prognostic significance*. J Surg Oncol **61**(3): 177-184.

Lehr HA, Hansen DA, Kussick S, Li M, Hwang H, Krummenauer F, Trouet S and Gown AM (1999). *Assessment of proliferative activity in breast cancer: MIB-1 immunohistochemistry versus mitotic figure count*. Hum Pathol **30**(11): 1314-1320.

Lejeune M, Lopez C, Bosch R, Korzynska A, Salvado MT, Garcia-Rojo M, Neuman U, Witkowski L, Baucells J and Jaen J (2011). *JPEG2000 for automated quantification of immunohistochemically stained cell nuclei: a comparative study with standard JPEG format*. Virchows Arch **458**(2): 237-245.

Leong FJWM and McGee J (2001). *Automated complete slide digitization: a medium for simultaneous viewing by multiple pathologists*. J Pathol **195**(4): 508-514.

Levsky JM and Singer RH (2003). *Fluorescence in situ hybridization: past, present and future*. J Cell Sci **116**(Pt 14): 2833-2838.

Linder N, Martelin E, Lundin M, Louhimo J, Nordling S, Haglund C and Lundin J (2009). *Xanthine oxidoreductase - clinical significance in colorectal cancer and in vitro expression of the protein in human colon cancer cells*. Eur J Cancer **45**(4): 648-655.

Lopez C, Lejeune M, Escriva P, Bosch R, Salvado MT, Pions LE, Baugells J, Cugat X, Alvaro R and Jaen J (2008). *Effects of Image Compression on Automatic Count*

of Immunohistochemically Stained Nuclei in Digital Images. J. Am. Med. Inf. Assoc. **15**(6): 794-798.

Lopez C, Martinez JJ, Lejeune M, Escriva P, Salvado MT, Pons LE, Alvaro T, Baucells J, Garcia-Rojo M, Cugat X and Bosch R (2009). *Roundness variation in JPEG images affects the automated process of nuclear immunohistochemical quantification: correction with a linear regression model*. Histochem. Cell Biol. **132**(4): 469-477.

Lundin M, Lundin J, Helin H and Isola J (2004a). *A digital atlas of breast histopathology: an application of web based virtual microscopy*. J Clin Pathol **57**(12): 1288-1291.

Lundin M, Lundin J and Isola J (2004b). *Virtual microscopy*. J Clin Pathol **57**(12): 1250-1251.

Mäenpää T (2005). *An iterative algorithm for fast iris detection*. Advances in Biometric Person Authentication, IWBRS 2005 Proceedings, Lect Notes Comput Sc **3781**: 127-134.

Mäenpää T and Pietikäinen M (2003). *Multi-scale binary patterns for texture analysis*. Lect Notes Comput Sc **2749**: 885-892.

Manjunath BS and Ma WY (1996). *Texture features for browsing and retrieval of image data*. Ieee T Pattern Anal **18**(8): 837-842.

Martens G, Poppe C, Lambert P and Van de Walle R (2010). *Noise- and compression-robust biological features for texture classification*. Visual Comput **26**(6-8): 915-922.

Mirmehdi M, Xie X and Suri J (2008). Handbook of texture analysis. London, Uk, Imperial College Press.

Mulrane L, Rexhepaj E, Penney S, Callanan JJ and Gallagher WM (2008). *Automated image analysis in histopathology: a valuable tool in medical diagnostics*. Expert Rev Mol Diagn **8**(6): 707-725.

Nanni L and Lumini A (2008a). *Local binary patterns for a hybrid fingerprint matcher*. Pattern Recogn **41**: 3461-3466.

Nanni L and Lumini A (2008b). *A reliable method for cell phenotype image classifation*. Artificial intelligence in medicine **43**: 87-97.

Nanni L, Lumini A and Brahnan S (2010). *Local binary pattern variants as texture descriptors for medical image analysis*. Artificial intelligence in medicine **49**: 117-125.

Narath R, Lorch T, Rudas M and Ambros PF (2004). *Automatic quantification of gene amplification in clinical samples by IQ-FISH*. Cytometry B Clin Cytom **57**(1): 15-22.

Netten H, Young IT, van Vliet LJ, Tanke HJ, Vroljik H and Sloos WC (1997). *FISH and chips: automation of fluorescent dot counting in interphase cell nuclei*. Cytometry **28**(1): 1-10.

Ojala T, Pietikäinen M and Harwood D (1996). *A comparative study of texture measures with classification based on featured distributions.* Pattern Recogn **29(1)**: 51–59.

Ojala T, Pietikäinen M and Mäenpää T (2000). *Gray scale and rotation invariant texture classification with local binary patterns*. Computer Vision - Eccv 2000, Pt I, Proceedings **1842**: 404-420.

Ojala T, Pietikäinen M and Mäenpää T (2002). *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. Ieee T Pattern Anal **24**(7): 971-987.

Petushi S, Garcia FU, Haber MM, Katsinis C and Tozeren A (2006). *Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer*. BMC Med Imaging **6**: 14.

Pietiläinen T, Lipponen P, Aaltomaa S, Eskelinen M, Kosma VM and Syrjänen K (1996). *The important prognostic value of Ki-67 expression as determined by image analysis in breast cancer*. J Cancer Res Clin **122**(11): 687-692.

Pietras K and Ostman A (2010). *Hallmarks of cancer: Interactions with the tumor stroma*. Exp Cell Res **316**(8): 1324-1331.

Pinder SE, Wencyk P, Sibbering DM, Bell JA, Elston CW, Nicholson R, Robertson JF, Blamey RW and Ellis IO (1995). *Assessment of the new proliferation marker MIB1 in breast carcinoma using image analysis: associations with other prognostic factors and survival*. Br J Cancer **71**(1): 146-149.

Querzoli P, Albonico G, Ferretti S, Rinaldi R, Magri E, Indelli M and Nenci I (1996). *MIB-1 proliferative activity in invasive breast cancer measured by image analysis*. Journal of Clinical Pathology **49**(11): 926-930.

Querzoli P, Ferretti S, Albonico G, Magri E, Scapoli D, Indelli M and Nenci I (1995). *Application of Quantitative-Analysis to Biologic Profile Evaluation in Breast-Cancer*. Cancer **76**(12): 2510-2517.

Raimondo F, Gavrielides MA, Karayannopoulou G, Lyroudia K, Pitas I and Kostopoulos I (2005). *Automated evaluation of Her-2/neu status in breast tissue from fluorescent in situ hybridization images*. IEEE Trans Image Process **14**(9): 1288-1299.

Rasband WS (1997-2006). ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA.

Rexhepaj E, Brennan DJ, Holloway P, Kay EW, McCann AH, Landberg G, Duffy MJ, Jirström K and Gallagher WM (2008). *Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer*. Breast Cancer Res **10**(5): R89.

Rojo MG, Bueno G and Slodkowska J (2009). *Review of imaging solutions for integrated quantitative immunohistochemistry in the Pathology daily practice*. Folia Histochem Cytobiol **47**(3): 349-354.

Rudolph P, Olsson H, Bonatz G, Ratjen V, Bolte H, Baldetorp B, Ferno M, Parwaresch R and Alm P (1999). *Correlation between p53, c-erbB-2, and*

topoisomerase II alpha expression, DNA ploidy, hormonal receptor status and proliferation in 356 node-negative breast carcinomas: Prognostic implications. J Pathol **187**(2): 207-216.

Ruifrok AC and Johnston DA (2001). *Quantification of histochemical staining by color deconvolution*. Anal Quant Cytol Histol **23**(4): 291-299.

Sahu B, Laakso M, Ovaska K, Mirtti T, Lundin J, Rannikko A, Sankila A, Turunen JP, Lundin M, Konsti J, Vesterinen T, Nordling S, Kallioniemi O, Hautaniemi S and Jänne OA (2011). *Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer*. Embo J **30**(19): 3962-3976.

Sapino A, Marchio C, Senetta R, Castellano I, Macri L, Cassoni P, Ghisolfl G, Cerrato M, D'Ambrosio E and Bussolati G (2006). *Routine assessment of prognostic factors in breast cancer using a multicore tissue microarray procedure*. Virchows Arch **449**(3): 288-296.

Sertel O, Kong J, Shimada H, Catalyurek UV, Saltz JH and Gurcan MN (2009). *Computer-aided Prognosis of Neuroblastoma on Whole-slide Images: Classification of Stromal Development*. Pattern Recognit **42**(6): 1093-1103.

Signolle N, Revenu M, Plancoulaine B and Herlin P (2010). *Wavelet-based multiscale texture segmentation: Application to stromal compartment characterization on virtual slides*. Signal Process **90**: 2412-2422.

Sihto H, Lundin J, Lehtimäki T, Sarlomo-Rikala M, Butzow R, Holli K, Sailas L, Kataja V, Lundin M, Turpeenniemi-Hujanen T, Isola J, Heikkilä P and Joensuu H (2008). *Molecular subtypes of breast cancers detected in mammography screening and outside of screening*. Clin Cancer Res **14**(13): 4103-4110.

Sinard JH (2005). Practical Pathology Informatics: Demystifying informatics for the practicing anatomic pathologist. New Haven, CT, Springer Science + Business Media, Inc.

Sobin LH, Gospodarowicz MK and Wittekind CH (2009). TNM Classification of Malignant Tumors. Oxford, Wiley-Blackwell.

Soh LK and Tsatsoulis C (1999). *Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices*. Ieee T Geosci Remote **37**(2): 780-795.

Strömberg S, Björklund MG, Asplund C, Skollermo A, Persson A, Wester K, Kampf C, Nilsson P, Andersson AC, Uhlen M, Kononen J, Ponten F and Asplund A (2007). *A high-throughput strategy for protein profiling in cell microarrays using automated image analysis*. Proteomics **7**(13): 2142-2150.

Stuart-Harris R, Caldas C, Pinder SE and Pharoah P (2008). *Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients*. Breast **17**(4): 323-334.

Sullivan CA, Ghosh S, Ocal IT, Camp RL, Rimm DL and Chung GG (2009). *Microvessel area using automated image analysis is reproducible and is associated with prognosis in breast cancer*. Hum Pathol **40**(2): 156-165.

Tahir MA and Bouridane A (2006). *Novel round-robin tabu search algorithm for prostate cancer classification and diagnosis using multispectral imagery*. IEEE Trans Inf Technol Biomed **10**(4): 782-793.

Tengowski MW (2004). *Image compression in morphometry studies requiring 21 CFR Part 11 compliance: procedure is key with TIFFs and various JPEG compression strengths*. Toxicol Pathol **32**(2): 258-263.

Torhorst J, Bucher C, Kononen J, Haas P, Zuber M, Kochli OR, Mross F, Dieterich H, Moch H, Mihatsch M, Kallioniemi OP and Sauter G (2001). *Tissue microarrays for rapid linking of molecular changes to clinical endpoints*. Am J Pathol **159**(6): 2249-2256.

Tubbs RR, Pettay JD, Swain E, Roche PC, Powell W, Hicks DG and Grogan T (2006). *Automation of manual components and image quantification of direct dual label fluorescence in situ hybridization (FISH) for HER2 gene*

amplification: A feasibility study. Appl Immunohistochem Mol Morphol **14**(4): 436-440.

Tuominen VJ and Isola J (2009). *The application of JPEG2000 in virtual microscopy*. J Digit Imaging **22**(3): 250-258.

Tuominen VJ, Ruotoistenmäki S, Viitanen A, Jumppanen M and Isola J (2010). *ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67*. Breast Cancer Res **12**(4): R56.

Turashvili G, Leung S, Turbin D, Montgomery K, Gilks B, West R, Carrier M, Huntsman D and Aparicio S (2009). *Inter-observer reproducibility of HER2 immunohistochemical assessment and concordance with fluorescent in situ hybridization (FISH): pathologist assessment compared to quantitative image analysis*. BMC Cancer **9**: 165.

Turbin DA, Leung S, Cheang MC, Kennecke HA, Montgomery KD, McKinney S, Treaba DO, Boyd N, Goldstein LC, Badve S, Gown AM, van de Rijn M, Nielsen TO, Gilks CB and Huntsman DG (2008). *Automated quantitative analysis of estrogen receptor expression in breast carcinoma does not differ from expert pathologist scoring: a tissue microarray study of 3,484 cases*. Breast Cancer Res Treat **110**(3): 417-426.

Walkowski S and Szymas J (2011). *Quality evaluation of virtual slides using methods based on comparing common image areas*. Diagn Pathol **6**(Suppl 1): S14.

Weinstein RS, Bloom KJ and Rozek LS (1987). *Telepathology and the networking of pathology diagnostic services*. Arch Pathol Lab Med **111**(7): 646-652.

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 and Hayes DF (2007). *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 **131**(1): 18.