Tampere University

HENRIK HELIN

Digital Histopathology of Cancer

Tampere University Dissertations 565

Tampere University Dissertations 565

HENRIK HELIN

Digital Histopathology of Cancer

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Medicine and Health Technology
of Tampere University,
for public discussion in the auditorium Syke
of the Finn-Medi 2 building, Biokatu 8, Tampere,
on 11 March 2022, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Medicine and Health Technology
Finland

Responsible supervisor and Custos	Professor Jorma Isola Tampere University Finland	
Pre-examiners	Professor Pekka Taimen University of Turku Finland	Docent Juha Väyrynen University of Oulu Finland
Opponent	Docent Tuomas Mirtti University of Helsinki Finland	

Cover design: Roihu Inc.

ISBN 978-952-03-2319-6 (print)
ISBN 978-952-03-2320-2 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2320-2

PunaMusta Oy – Yliopistopaino
Joensuu 2022

"We are here to add what we can to life, not to get what we can from it."

quote credited to Sir William Osler, renowned Canadian physician and student of Rudolf Virchow, German physician and founder of cellular pathology

ACKNOWLEDGEMENTS

This study was carried out at the Institute of Medical Technology, later BioMediTech, and currently the Faculty of Medicine and Health Technology, University of Tampere. The study was financially supported by research grants from the Cancer Society of Pirkanmaa, the Finnish Medical Society Duodecim, the Scientific Foundation of Instrumentarium Inc, the Medical Society of Finland, Biomedicum Foundation, the Life and Health Medical Fund, the Swedish Cultural Foundation in Finland, Finnish Cancer Foundation, and Sigrid Juselius Foundation, for which I express my gratitude.

I would like to offer my deepest appreciation to my supervisor professor Jorma Isola for his unwavering support as well as his ingenuity. I would also like to thank the reviewers of this thesis, professor Pekka Taimen and adjunct professor Juha Väyrynen, as well as the members of my thesis supervision committee, dean Tapio Visakorpi and professor Anne Kallioniemi. To my opponent, adjunct professor Tuomas Mirtti, I extend my utmost gratitude.

This work would not have come into existence without all of the wonderful and talented people I have had the pleasure to work with, both as a coworker and as a coauthor, and for this I am thankful. Professor Johan Lundin and Mikael Lundin, MD, deserve special thanks for the opportunity of being part of their team. I wish to thank Teemu Tolonen, PhD, for guidance on the practical aspects of digital pathology. I am grateful for the professional advice from Ms. Leena Juvonen on the spelling of the thesis.

I would like to thank my parents for their unconditional support and love which has carried me through this process as well as all of the other endeavors I have faced in my life. My dear friend Timo Uotinen, PhD, I would like to thank for his support and for all of the stimulating discussions we have had over the years. Lastly, my beautiful partner and my love, Emma, I am truly lucky to have you by my side.

ABSTRACT

Cancer is a significant and growing public health concern. According to the World Health Organisation's estimates it is – after cardiovascular diseases – the second leading cause of death worldwide. Excluding non-melanoma skin cancers the most common types of cancer are for women breast cancer and for men lung cancer followed by prostate cancer.

While the biological understanding of cancer has expanded, so too has the selection of available treatments. More than one fourth of all new medicines entering the market during 2010-2018 were for treating cancer. In order for a patient to benefit from the wide variety of cancer treatments, and avoid adverse effects, their unique cancer has to be matched with the appropriate treatment. For this the cancer needs to be both diagnosed accurately and classified in detail.

Although non-invasive imaging methods, such as magnetic resonance imaging, have evolved substantially in recent years, the basis of cancer diagnosis is still in histopathology, that is, the pathological evaluation of tissue removed through surgery or needle biopsy. The light microscope has remained the pathologist's main diagnostic tool for a century and a half allowing for the examination of tissue down to cellular – and even subcellular – level. Important adjuncts to routine histopathological staining of tissue, needed for light microscopy, are techniques allowing for the visualization of protein antigens and nucleic acid in the tissue. These techniques, among which are immunohistochemistry and in situ hybridization, respectively, can be used for instance in the molecular characterization of cancer.

There are challenges in meeting the need for accurate diagnosing and characterization of cancer. One such challenge is posed by the shortage of pathologists observed in Finland and elsewhere. Another challenge is the variability in the interpretation of the tumor growth pattern (grading, such as Gleason grading in prostate cancer) and in the interpretation of certain tissue staining patterns (such as the immunohistochemical staining of the HER2 molecule in breast cancer). This variability manifests itself both between pathologists (interobserver variation) and also in the same pathologist's work over time (intraobserver variation). A third

challenge is presented by the fact that the light microscope – although a reliable, cheap, and easy-to-use diagnostic tool – has shortcomings in the modern day pathology service.

Digital histopathology presents a new way of carrying out the central task of a pathologist in managing cancer patients, namely making the diagnosis and characterising the tumor in detail. Making the shift from a light microscope to a computer environment offers many benefits, some of which have been examined in this dissertation.

The present study was carried out with the purpose of developing and testing applications of digital pathology in order to improve the histopathological diagnosis of cancer. The individual studies looked at advancing the teaching and standardization of Gleason grading or prostate cancer, aiding in the interpretation of immunohistochemical staining of prostate and breast cancer, as well as facilitating the implementation of digital pathology by way of a novel whole slide image optimised image compression algorithm and mapping the determinants of an optimal imaging resolution for a whole slide scanner.

We demonstrated that whole slide images can be used to assess the Gleason grade of a prostate biopsy and that the use of an internet based platform can be beneficial in assessing interobserver variation in the grading and teaching and standardising the grading.

Besides Gleason grading another important aspect of prostate histopathology is the interpretation of immunohistochemistry. We created a method of viewing two whole slide images simultaneously and synchronously and tested this method in visualising the AMACR-p63 double stain along with normal hematoxylin and eosin staining of prostate biopsies. We showed that this technique can be used for histopathology education as well as in clinical diagnostics in selected cases.

A key issue in breast cancer diagnostics is defining the HER2 status of a tumor, that is, whether the tumor overexpresses the molecule and can then be treated with HER2 antibody based drugs. We studied the use of digital image analysis, using both photomicrographs and whole slide images, in aiding the pathologist in defining the HER2 status on a breast cancer surgical resection specimen. We showed that using a free and publicly available image analysis software can help to resolve cases otherwise deemed equivocal by conventional light microscopy.

The introduction of digital histopathology into routine diagnostic work is underway. One technical challenge is managing the large amounts of image data generated by whole slide images. When there is a need to store large numbers of whole slide images it is essential to strike a balance between image fidelity and file

size. To deal with this issue we studied the optimal imaging resolution of a whole slide scanner using a methodology that can be utilised for instance in comparing whole slide scanners before acquiring one. In addition we introduced a novel way of image compression suited for whole slide images in order to reduce the storage footprint, and cost, of whole slide images.

The first two studies in this dissertation represent the very beginnings of whole slide imaging in pathology, and the field has advanced since then, perhaps in small part due to the findings in these studies. Taken together, the findings in this dissertation can hopefully advance the use of digital pathology in cancer diagnostics and thereby improve the care of cancer patients.

TIIVISTELMÄ

Syöpä on merkittävä ja yleistyvä kansansairaus. Maailman terveysjärjestön mukaan syöpä on maailmanlaajuisesti toiseksi yleisin kuolinsyy sydän- ja verisuonitautien jälkeen. Jos ei-melanoottisia ihosyöpiä ei oteta huomioon, ovat tavallisimmat syöpätyypit naisilla rintasyöpä ja miehillä keuhkosyöpä ja eturauhassyöpä.

Sitä mukaa kun syövän biologisten syntymekanismien ymmärrys on lisääntynyt, ovat myös hoitovaihtoehdot lisääntyneet. Useampi kuin joka neljäs uusi lääke, joka lanseerattiin vuosina 2010-2018, oli tarkoitettu syövän hoitoon. Jotta potilas voisi hyötyä tarjolla olevasta laajasta syöpälääkevalikoimasta ja minimoida lääkkeiden haittavaikutukset, tulee hoito kohdistaa hänen yksilölliseen syöpäänsä. Tätä varten syöpä on sekä diagnosoitava luotettavasti että luokiteltava yksityiskohtaisesti.

Vaikka kajoamattomat kuvantamistutkimukset kuten magneettikuvaus ovat viime vuosina kehittyneet huomattavasti, on syöpädiagnostiikan perusta edelleen histopatologiassa eli leikkauksessa tai neulanäytteenotossa poistetun kudoksen mikroskooppisessa tutkimuksessa. Valomikroskooppi on pysynyt patologin pääasiallisena työvälineenä yli puolentoista vuosisadan ajan. Se on sallinut kudoksen tarkastelun aina solutasolle saakka ja jopa sitä pienempiin rakenteisiin. Tärkeitä lisätutkimuksia tavallisen valomikroskooppisen tutkimuksen lisäksi ovat proteiiniantigeenien osoittamistutkimukset, kuten immunohistokemia ja in situ -hybridisaatio, joita voidaan käyttää syöpäkudoksen luokittelemiseen.

Syövän tarkalla diagnosoimisella ja luokittelulla on haasteensa. Yksi sellainen on Suomessa ja ulkomailla vallitseva pula patologeista. Toinen haaste liittyy kasvainten välisen vaihtelun arviointiin, joka on tärkeää kasvainten kasvutaipumuksen luokittelussa (esim. eturauhassyövän Gleason-luokitus) ja tiettyjen värjäysten tulkinnassa (esim. rintasyövän HER2-värjäytyminen). Todellisen biologisen vaihtelun lisäksi vaihtelua esiintyy patologien välisissä arvioissa (interobserver variation) sekä saman patologin luokitteluissa eri ajan hetkellä (intraobserver variation). Kolmas haaste on itse valomikroskooppi. Vaikka se on luotettava, halpa ja helppokäyttöinen diagnostiikkalaite, on sillä omat puutteensa modernin patologian laboratorion työnkulussa.

Digitaalihistopatologia edustaa uutta tapaa toteuttaa patologin pääasiallinen työtehtävä syöpäpotilaan hoidossa: asettaa diagnoosi ja luokitella syöpä yksityiskohtaisesti. Siirtyminen valomikroskoopista tietokoneympäristöön tarjoaa monia etuja, joista muutamia on tutkittu tässä väitöskirjassa.

Tämän tutkimuksen tarkoituksena oli kehittää ja testata digitaalipatologian sovelluksia syöpädiagnostiikan parantamiseksi. Osatöissä tutkittiin eturauhassyövän Gleason-luokituksen opettamista ja standardointia, rinta- ja eturauhassyövän immunohistokemiallisten värjäysten tulkintaa, digitaalinäytteille kehitettyä kuvanpakkausmenetelmää, sekä näyteskannerin optimaalisen kuvausresoluution määrittämistä.

Väitöskirjassa osoitetaan, että digitaalinäytteitä voi käyttää eturauhaskoepalan Gleason-luokituksen tekemiseen ja että internet-pohjainen ohjelma voi edistää tulkitsijoiden välisen vaihtelun määrittämistä sekä Gleason-luokituksen opettamista ja standardisointia.

Gleason-luokituksen ohella toinen tärkeä osa eturauhassyövän histopatologiaa on immunohistokemiallisten värjäysten tulkinta. Tässä väitöskirjassa esitetään menetelmä, jolla kahta digitaalinäytettä voidaan tutkia yhtäaikaisesti ja synkronoidusti. Menetelmää testattiin eturauhassyövän immunohistokemiallisella AMACR–p63-kaksoisvärjäyksellä yhdessä rutiininomaisen hematoksyliini–eosiini-värjäyksen kanssa. Tutkimuksessa osoitettiin, että tekniikkaa voidaan käyttää hyväksi histopatologian opetuksessa ja valikoiduissa tapauksissa kliinisessä diagnostiikassa.

Keskeinen asia rintasyövän diagnostiikassa on HER2-statuksen tutkiminen, koska kasvaimia, joissa HER2 on yli-ilmentynyt, voidaan hoitaa anti-HER2-lääkkeillä. Yhdessä osatöistä tutkittiin digitaalisen kuva-analyysin käyttöä niin valomikroskooppikuvilla kuin digitaalinäytteillä tarkoituksena auttaa patologia määrittämään kirurgisesti poistetun kasvainkudoksen HER2-status. Työssä osoitettiin, että ilmaista ja kaikille avointa ohjelmistoa käyttämällä voitiin vähentää HER2-statuksen suhteen vaikeatulkintaisten tapausten määrää.

Digitaalihistopatologian käyttöönotto rutiinidiagnostiikkaan on laajentumassa nopeasti. Yksi tekninen haaste on digitaalinäytteiden vaatiman suuren tallennuskapasiteetin hallinta. Tarve tallentaa suuria määriä tietoa edellyttää digitaalinäytteiden kuvanlaadun ja tiedostokoon yhteensovittamista. Yhdessä tämän väitöskirjan osatöistä tutkittiin skannerimikroskoopin optimaalisen kuvausresoluution määrittämistä. Menetelmää voidaan hyödyntää esimerkiksi vertailtaessa skannereita ennen hankintaa. Toisessa osatyössä esiteltiin uusi kuvanpakkausmenetelmä, joka suunniteltiin varta vasten histopatologisia

digitaalinäytteitä varten niiden tiedostokoon minimoimiseksi ja siten tallennuskustannusten pienentämiseksi.

Tämän väitöskirjan kaksi ensimmäistä osatyötä edustavat digitaalipatologian alkutaivalta ja tutkimuskenttä on kehittynyt sittemmin, mahdollisesti pieneltä osin edellä mainittujen tutkimusten löydösten myötä. Yhteenvetona osatyöt toivottavasti vievät digitaalipatologian alaa eteenpäin ja siten edesauttavat syöpäpotilaiden hoitoa.

SAMMANFATTNING

Cancer är en växande folksjukdom. Enligt Världshälsoorganisationens uppskattningar är cancer den näst vanligaste dödsorsaken i världen efter hjärt- och kärlsjukdomar. Med undantag för icke-melanom hudcancer är de vanligaste cancerformerna för kvinnor bröstcancer och för män lungcancer efterföljt av prostatacancer.

I takt med att den biologiska förståelsen för cancer har ökat har också urvalet av cancerbehandlingar vuxit. Mer än var fjärde av alla nya läkemedel som lanserades under 2010-2018 var till för att behandla cancer. För att en patient ska dra nytta av det stora utbudet av cancerbehandlingar och undvika biverkningar måste hennes unika cancer paras ihop med rätt behandling. Detta kräver både korrekt diagnos och noggran klassifikation av cancern i fråga.

Även om icke-invasiv bilddiagnostik, såsom magnetkameraundersökning, har utvecklats påtagligt under senare år utgörs fortfarande grunden i cancerdiagnostik av histopatologi, det vill säga den mikroskopiska undersökningen av vävnad avlägsnad kirurgiskt eller med nålprovtagning. Ljusmikroskopet har kvarstått som patologens huvudsakliga diagnostiska instrument under ett och ett halvt sekel. Det möjliggör undersökningen av vävnad ända till cellulär och även subcellular nivå. Förutom morfologisk undersökning av färgade vävnadssnitt kan man använda sig av metoder som visualiserar proteinantigen och nukleinsyra i vävnaden. Dessa tekniker, inklusive immunhistokemi och in situ hybridisering, kan användas för molekylär profilering av cancer.

Det finns utmaningar i att uppfylla kravet för korrekt och detaljerad karaktärisering av cancer. En sådan är bristen på patologer som konstaterats i Finland och annanstans. En annan utmaning är den variabilitet som existerar i hur man bedömer olika egenskaper i vävnaden, till exempel differentieringsgraden av en tumör (till exempel med Gleasonsystemet när det gäller prostatacancer) eller vissa färgningsmönster (till exempel den immunhistokemiska färgningen av HER2-molekylen i bröstcancerpreparat). Denna variabilitet manifesterar sig både mellan olika patologer (så kallad interobservervariaton) och i hur en och samma patolog bedömer preparat över tid (intraobservervariation). En tredje utmaning utgörs av

det faktum att även om ljusmikroskopet är ett pålitligt, billigt och lättanvändligt diagnostiskt verktyg har den vissa tillkortakommanden i dagens moderna patologilaboratorium.

Digital histopatologi representerar ett nytt sätt att genomföra patologens centrala roll i handläggningen av cancerpatienter, nämligen diagnosticeringen och karaktäriseringen av tumören. Att byta från ljusmikroskopi till digital mikroskopi medför många fördelar varav några har undersökts i denna avhandling.

Målet med den här avhandlingen var att utveckla och testa tillämpningar av digital patologi i syfte att förbättra den histopatologiska cancerdiagnostiken. De individuella studierna behandlar dels bröst- och prostatacancer (undervisningen och standardiseringen av Gleason-gradering för prostatacancer samt tolkningen av immunohistokemi vid bröst- och prostatacancer) och dels digital patologi i allmänhet (utvecklingen av en ny bildkompressionsmetod optimerad för digitala histopatologiska preparat samt utvecklingen av en metod för att kartlägga den optimala digitaliseringsupplösningen för preparatscanners).

Denna avhandling påvisade att digitala preparat kan användas för Gleason-graderingen av prostatabiopsier och att användningen av ett internetbaserat program kan uttnyttjas i studerandet av interobservervariation samt i undervisningen och standardiseringen av Gleason-gradering.

Förutom Gleason-graderingen är en annan viktig aspekt av prostatahistopatologi tolkningen av immunhistokemi. I en av delstudierna beskrivs en metod för att visualisera två digitala preprat samtidigt och synkront. Denna metod testades med AMACR–p63-dubbelfärgning samt vanlig hematoloxylin–eosin-färgning. Den här dubbelvisualiseringstekniken kan användas för undervisning samt för rutindiagnostik i utvalda fall.

En viktig komponent i bröstcancerdiagnostiken är bestämmandet av om tumören överuttrycker molekylen HER2 eftersom det har en direkt koppling till vare sig man kan använda anti-HER2-läkedemel mot cancern i fråga. I denna avhandling undersöktes användningen av digital bildanalys, med både mikrofotografi och hela digitala preparat, som stöd för patologens tolkning av HER2-status i operationspreprat från bröstcancrar. Det visade sig att användandet av ett fritt och allmänt tillgängligt datorprogram för bildanalys kan underlätta för patologen att lösa annars problemtiska fall med avseende på HER2-status.

Införandet av digital patologi i rutindiagnostik är i antågande. Ett tekniskt hinder är hur man handskas med den stora mängd bilddata som genereras av hela digitaliserade histopatologiska preparat. Ju fler digitala preparat som man måste lagra desto större är kravet att balansera bildkvalitén med filstorleken. För att

hantera denna situation undersöktes hur man bestämmer den optimala scanningupplösningen för en preparatscanner. Den metod som utvecklades kan användas i framtiden för till exempel jämförandet av olika preparatscanners. Dessutom utvecklades en ny bildkompressionsmetod avsedd för hela digitala histopatologipreparat med avsikt att reducera filstorleken och på så sätt behovet och kostnaden av lagringsutrymme.

De två första delstudierna i denna avhandling representerar själva början av digital patologi och fältet har utvecklats sedan dess, möjligen dels tack vare fynden i nämnda studier. Tillsammans kan resultaten av denna avhandling förhoppningsvis hjälpa till att utöka användandet av digital patologi i cancerdiagnostiken och på så sätt befrämja vården av patienter med cancer.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

A/D	analog-to-digital
AI	artificial intelligence
AMACR	alpha-methylacyl-coenzyme-A racemase
ASAP	atypical small acinar proliferation (of the prostate)
ASCO	The American Society of Clinical Oncology
CAP	The College of American Pathologists
CCD	charge-coupled device
CI	confidence interval
CISH	chromogenic in situ hybridization
CMOS	complementary metal oxide semiconductor
CMYK	cyan, magenta, yellow, key
DAB	diaminobenzidine
DCT	discrete cosine transform
DIA	digital image analysis
DICOM	Digital Imaging and Communications in Medicine
DNA	deoxyribonucleic acid
DP	digital pathology
DPA	The Digital Pathology Association
DWT	discrete wavelet transform
ER	estrogen receptor
FFPE	formalin-fixed and paraffin-embedded
FISH	fluorescence in situ hybridization
GB	gigabyte
GUPS	The Genitourinary Pathology Society
HDD	hard disk drive
HER2	human epidermal growth factor receptor 2 (ERBB2)
H&E	hematoxylin and eosin (staining)
IEC	International Electrotechnical Commission
IHC	immunohistochemistry

ISH	in situ hybridization
ISUP	The International Society of Urological Pathology
ITU	International Telecommunication Union
ISO	International Organization for Standardization
JP2	JPEG 2000 image compression algorithm
JPEG	Joint Photography Experts Group
LCD	liquid-crystal display
LED	light-emitting diode
LIS	laboratory information system
M	molar (mol/l)
mM	millimolar (mmol/l)
MB	megabyte
MeSH	Medical Subject Headings
MOSFET	metal-oxide-semiconductor field-effect-transistors
NA	numerical aperture
OLED	organic light-emitting diode
PIN	prostatic intraepithelial neoplasia
PR	progesterone receptor
PSA	prostate-specific antigen
RGB	red-green-blue
RNA	ribonucleic acid
ROI	region of interest
SSD	solid state drive
TB	terabyte
TFT-LCD	thin-film-transistor liquid-crystal display
tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
WHO	The World Health Organisation
WSI	whole slide image/whole slide imaging

ORIGINAL PUBLICATIONS

Publication I Web-based virtual microscopy in teaching and standardizing Gleason grading. Helin H, Lundin M, Lundin J, Martikainen P, Tammela T, Helin H, van der Kwast T, Isola J. Hum Pathol. 2005 Apr;36(4):381-6. doi:10.1016/j.humpath.2005.01.020

Publication II Virtual microscopy in prostate histopathology: simultaneous viewing of biopsies stained sequentially with hematoxylin and eosin, and alpha-methylacyl-coenzyme A racemase/p63 immunohistochemistry. Helin HO, Lundin ME, Laakso M, Lundin J, Helin HJ, Isola J. J Urol. 2006 Feb;175(2):495-9. doi:10.1016/S0022-5347(05)00164-3

Publication III Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry. Helin HO, Tuominen VJ, Ylinen O, Helin HJ, Isola J. Virchows Arch. 2015 Oct 22. [Epub ahead of print] doi:10.1007/s00428-015-1868-7

Publication IV Optimized JPEG 2000 Compression for Efficient Storage of Histopathological Whole slide Images.Helin H, Tolonen T, Ylinen O, Tolonen P, Näpänkangas J, Isola J.J Pathol Inform. 2018 May 25;9:20. doi: 10.4103/jpi.jpi_69_17. ECollection 2018.

In addition, previously unpublished data (Study V) is presented in chapters 4.3.5, 5.5, and 6.4.

Study V Determining the optimal resolution in a whole slide scanner. Helin H, Ylinen O, Isola J.

AUTHOR'S CONTRIBUTION

The author designed the studies I-IV together with Jorma Isola. In addition the author

Publication I collected the data and performed the statistical analysis independently, and wrote the manuscript with input from the coauthors

Publication II performed the histopathological staining, immunohistochemistry and whole slide scanning independently; and wrote the manuscript with input from the coauthors

Publication III collected and analysed the data independently; carried out the statistical analysis independently; and wrote the manuscript with input from the coauthors

Publication IV collected the data together with the coauthors; carried out the statistical analysis independently; and wrote the manuscript with input from the coauthors

Study V designed and carried out the study together with the coauthors.

1 INTRODUCTION

Cancer is a significant and growing public health concern. According to the World Health Organisation's estimates (2018a) it is – after cardiovascular diseases – the second leading cause of death worldwide. In 2018 there were 18 million new cancer cases and 9.5 million cancer-related deaths in the world (International Agency for Research on Cancer, 2020). By the year 2040 the number of new cancer cases is projected to rise to 29.5 million globally. The reasons for rising cancer rates are thought to include population growth and ageing as well as social and economic development (International Agency for Research on Cancer, 2018).

Worldwide the cancer types with the highest recorded incidences are for women breast cancer (estimated 24% of all cancers in women in 2018) and for men lung cancer followed by prostate cancer (estimated 14.5% and 13.5%, respectively, of all cancers in men in 2018). Breast cancer is globally the leading cause of cancer death in women (estimated 15% of all cancer deaths in women). For men, globally, prostate cancer comes fourth (estimated 6.7% of cancer deaths in men) after lung, liver and stomach cancers (accounting for 22%, 10.2% and 9.5% of cancer deaths in men, respectively) (International Agency for Research on Cancer, 2020). Prostate cancer occurs only in men whereas breast cancer predominantly in women; the incidence of breast cancer in men is 1% of that in women (Lester, 2015).

These incidence and mortality figures probably underestimate the number of non-melanoma skin cancers because of difficulties in obtaining reliable data (Lucas et al., 2006). Non-melanoma skin cancer is mainly made up of basal cell cancer (roughly 80%) and squamous cell cancer (roughly 20%) whose global incidences in 2000 were estimated as 10.5 million and 2.9 million, respectively. These can be compared with the global incidences of lung (2.1 million), breast (2.1 million) and prostate cancer (1.3 million) in 2018 (International Agency for Research on Cancer, 2020).

Cancer affects countries and peoples with high and low living standards alike. Whereas developed countries have traditionally had high incidence rates, developing populations are affected particularly severely: about 70% of deaths

from cancer occur in low- and middle-income countries (World Health Organization, 2018b).

Instead of a single disease entity, cancer is better understood as a group of diseases affecting different organs in different ways, although sharing common pathophysiological features. While the biological understanding of cancer has expanded, so too has the selection of available treatments. More than one fourth of new medicines entering the market during 2010-2018 were for treating cancer (Tufts Center for the Study of Drug Development, 2019).

To choose the best therapy for a patient the treating doctor has to know the specific type of cancer the patient has. With the great variety of cancer treatments currently available the need for a detailed and accurate diagnosis and classification of cancer is greater than ever before. A challenge in meeting the need for an accurate diagnosis without undue delay is presented by pathologist shortages reported in the United States (Metter et al., 2019) and the United Kingdom (The Royal College of Pathologists, 2018) as well as Finland (Mälkiä, 2015; Hänninen & Liljeberg, 2019). Many developing countries are of course persistently underserved with regard to pathology services (Tsang & Kovarik, 2010).

Although non-invasive radiological imaging methods, such as magnetic resonance imaging, have evolved substantially in recent years (Tonttila, 2020), the basis of cancer diagnosis is still in the pathological evaluation of tissue removed through surgery or needle biopsy, that is, histopathology (Gress et al., 2017). In histopathology light microscopy is used to visualize tissue architecture and cellular structures using visible light and special optics. To be able to discern microscopic details, the tissue specimen must be fixed (i.e. preserved, usually using formalin, a solution of formaldehyde and water), embedded (usually in paraffin wax), cut into thin sections, mounted on a glass slide and stained, so as to allow light to pass through the tissue and permit examination of the tissue details (Muskett, 2012).

Immunohistochemistry (IHC) is an adjunct to the morphological evaluation of tissue enabling the visualization of protein structures called antigens through binding with specific antibodies labeled with color bearing chromogens (Orchard & Javed, 2012). In situ hybridization (ISH) is similar to IHC with the exception of visualizing nucleic acid (DNA or RNA) by hybridization with specific probes. Two common methods of visualising in situ hybridization is by fluorescent dye (fluorescence in situ hybridization, FISH) and chromogens (chromogenic in situ hybridization, CISH) (Cullen et al., 2012).

The light microscope has remained the histopathologist's main diagnostic tool for over a century and a half (van den Tweel & Taylor, 2010). It permits the

examination of tissue down to cellular – and even subcellular – level, but it has important shortcomings regarding its usability in the modern healthcare setting. The light microscopic diagnosis requires that the microscope, the slide and the pathologist are all in one place. The microscopic slide is unique and can break or be misplaced. The standard microscope furthermore only permits one viewer to examine the slide at a time with the exception of multi-viewer microscopes which are nor available in every pathology department. Lastly, quantitative microscopic evaluation, that is, carrying out measurements on the tissue, is cumbersome.

Advances in information and communication technology, including digital photography, high speed internet access, computational power and digital data storage have enabled the development of digital microscopy systems (Sucaet & Waelput, 2014a). These systems have permitted addressing the usability issues of conventional light microscopy and extending its diagnostic power. This thesis examines the development and evaluation of digital pathology applications in the histopathological diagnostics of cancer, in particular breast and prostate cancer.

2 REVIEW OF THE LITERATURE

2.1 A brief history of light microscopic imaging

The light microscope is an instrument that uses visible light and lenses made from specially ground glass to produce magnified images of objects (Abramowitz, 2003). Light microscopes are divided into simple and compound ones, where the former have a single lens and the latter two or more lenses or lens systems. A magnifying glass is a form of a simple microscope whereas modern microscopes are all compound ones by design.

The compound light microscope was invented in the late sixteenth or early seventeenth century (Ball, 1966; Abramowitz, 2003). The Dutch lens-makers Hans and Zacharias Jansen, father and son, are often cited as being first in producing this kind of a microscope. Some other pioneers of light microscopy are the Dutch scientist Antonie van Leeuwenhoek, and the English scientist Robert Hooke. Leeuwenhoek advanced the production of high quality microscope lenses and made scientific discoveries of microorganisms while Hooke was the first person to visualize a biological cell using a microscope.

The introduction of the microscope in pathology, and the foundation of histopathology, occurred in the middle of the nineteenth century, and can be attributed in large part to the German scientists Johannes Müller and Rudolph Virchow (van den Tweel & Taylor, 2010; Hajdu, 2005). Müller was Virchow's mentor and encouraged Virchow to study the pathological changes in tissues microscopically. Virchow eventually founded cellular pathology, in which disease is conceptualized on the cellular level instead of the level of organs, thus laying the ground for modern scientific pathology.

At the end of the nineteenth century microscope optics had been improved by the introduction of so called achromatic objectives, a means to manufacture them precisely, and the principle of standardized so called Köhler illumination (Abramowitz, 2003). Achromatic microscope objectives correct for some chromatic aberration, that is, the optical defect in the microscopic image resulting from different wavelengths of white light being refracted differently.

While the optics of the light microscope have gone through further development, such as the introduction of apochromatic objectives (producing even higher quality images with even less chromatic aberrations than achromatic objectives), pioneered by the German physicist Ernst Abbe, it can be argued that no fundamental improvements have been made to the design of the instrument in over a hundred years (Abramowitz, 2003; Nature, n.d.).

Photomicrography, or microscopic photography, emerged in the nineteenth century (Morrison & Gardner, 2015; Overney & Overney, 2011; Jardim & Peres, 2014). The English scientist William Henry Fox Talbot is often cited as the first person to have successfully taken microscopic photographs. Initially the images, which lacked color, were formed on bulky slides and plates and the photographic process was both laborious and difficult to master. A major breakthrough in both ease-of-use and image quality came in the 1930s when the first high quality three-colored multilayered photographic films were introduced into the commercial market.

In the 1950s and 1960s new imaging techniques called video microscopy and television microscopy started to emerge so as to allow making histopathological diagnoses over a distance (Krupinski et al., 2016). These techniques, together forming the beginnings of telepathology, are imaging methods in which the analog video image from a microscope is sent to a remote location for visual analysis with the microscope still being operated on-site.

The term telepathology was first used in 1986 by pathologist Ronald Weinstein (Weinstein, 1986) who also invented the first dynamic-robotic telepathology system in which the viewer controls the microscope remotely (Weinstein et al., 1987). In parallel with the introduction of dynamic-robotic telepathology another method, called static telepathology, was being developed. In this approach individual static images, captured by a pathologist on-site, are transmitted to an off-site pathologist for analysis. Telepathology is nowadays defined in the medical subject headings database (MeSH) as ”transmission and interpretation of tissue specimens via remote telecommunication, generally for the purpose of diagnosis or consultation but may also be used for continuing education” (U.S. National Library of Medicine, n.d.).

The invention of the virtual microscope in the 1980s and its first commercial design in 1994 are credited to the scientist and inventor James Bacus (Weinstein et al., 2012; Pantanowitz, 2017). The first mentions of virtual microscopy in the MEDLINE research database are from 1997 (Ferreira et al., 1997; Grimes et al., 1997). The development of virtual microscopy was made possible by a number of

technological advances, the most important of which are the large scale development of filmless digital cameras in the 1980s and the invention of the World Wide Web in the early 1990s (Sucaet & Waelput, 2014). Additional important enabling factors were the rapidly advancing processing power of computers, the growing hard drive storage capacity and the widespread introduction of high speed Internet access – the first two of these phenomena being sometimes referred to as Moore's law (Peercy, 2000) and Kryder's law (Walter, 2005), respectively. Moore's law states that the number of transistors on a microchip doubles approximately every two years whereas Kryder's law asserts that the hard disk storage density doubles approximately every thirteen months.

Virtual microscopy is defined by The digital pathology association (DPA) as ”viewing of microscope slides on a computer screen over a network” and is sometimes used interchangeably with the term digital pathology (Digital Pathology Association, n.d). Virtual microscopy as a term is being increasingly replaced by digital pathology: web search engine results via Google give ”digital pathology” 931 000 results, ”virtual microscopy” 258 000 results and ”virtual microscope” 131 000 results (web search conducted 6 January 2022 on https://www.google.com/).

Digital pathology is defined by the DPA as ”a dynamic, image-based environment that enables the acquisition, management and interpretation of pathology information generated from a digitized glass slide” (Digital Pathology Association, n.d). The first research papers using the term ”digital pathology” in the MEDLINE database are from the year 2000 (Saltz, 2000; Barbareschi, 2000; Danielsen, 2000). Digital pathology can be thought of as an integral part of pathology informatics, which is defined by the Association for Pathology Informatics as ”collecting, examining, reporting, and storing large complex sets of data derived from tests performed in clinical laboratories, anatomic pathology laboratories, or research laboratories in order to improve patient care and enhance our understanding of disease-related processes” (Association for Pathology Informatics, n.d.).

Whole slide imaging is according to the DPA ”the acquisition process of creating a virtual slide or whole slide image on a slide scanner” (Digital Pathology Association, n.d). The terms whole slide image, digital slide and virtual slide are used interchangeably to refer to a complete digitized microscopic slide produced by a whole slide scanner. A whole slide scanner is in essence a robotic microscope fitted with a digital camera that photographs a microscopic specimen and uses special software to stitch the individual photomicrographs (microscopic

photographs) into a single whole slide image which can be navigated freely and viewed using different magnifications.

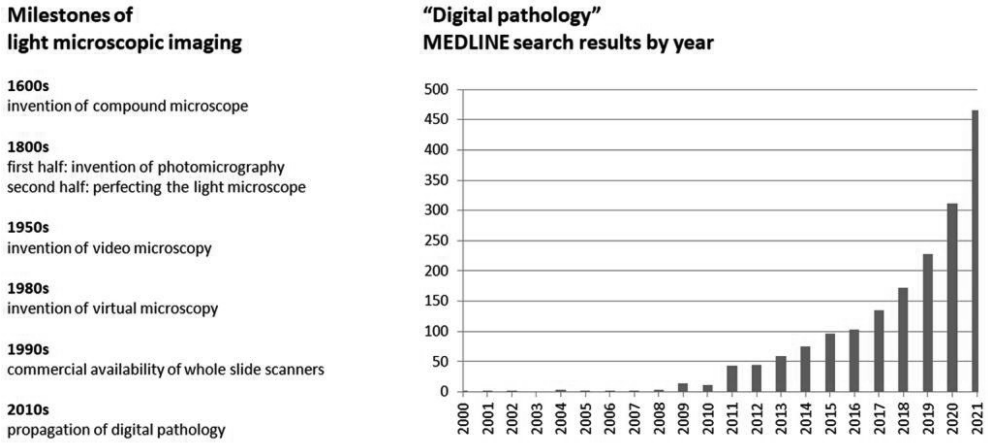
An overview of the use of digital pathology related terminology is presented in Table 1. It is worth noting that from the viewpoint of pathological imaging, digital pathology and telepathology are wider concepts encompassing both microscopic and gross (macroscopic) imaging, whereas virtual microscopy and whole slide imaging comprise only microscopic imaging. Figure 1 presents some milestones in the history of light microscopic imaging as well as the rate of publications on digital pathology.

Table 1. An overview of the use of digital pathology related terminology

	Telepathology	Virtual microscopy/ virtual microscope	Digital pathology	Whole slide imaging/ whole slide imaging
Appeareance of term in MEDLINE*	1986	1997	2000	2006
MEDLINE* term search results 6 January 2022	1198	408/83	1597	590
Google term search results 6 January 2022**	139 000	258 000/131 000	931 000	102 000

* https://www.ncbi.nlm.nih.gov/pubmed/
** https://www.google.com/

Figure 1. Milestones of light microscopic imaging, and digital pathology publications timeline



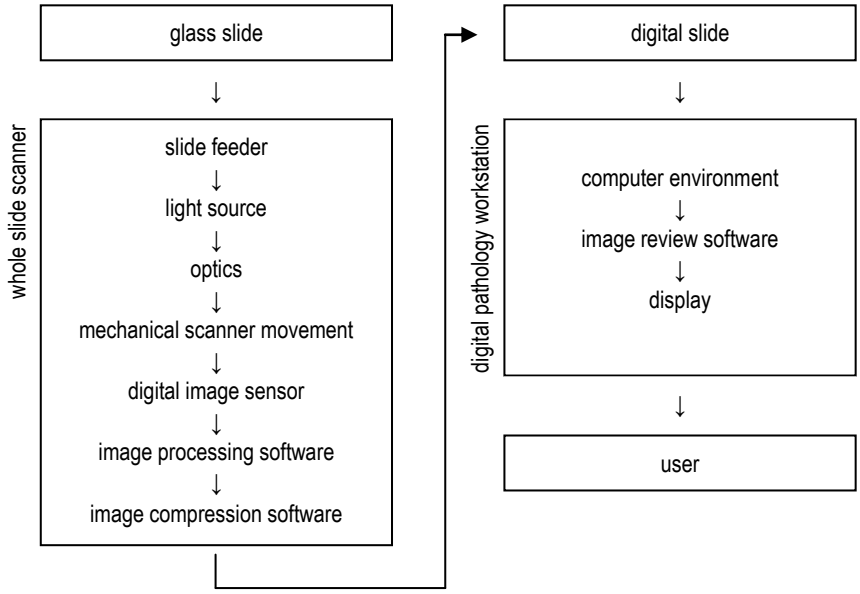
2.2 Whole slide scanning

Whole slide imaging in pathology is the process of producing entire digitized representations of glass slides and viewing them. Producing digital slides (whole slide scanning) is carried out using a whole slide scanner (or a digital slide scanner) whereas viewing them is done using a digital pathology workstation.

Digitizing a glass slide using a slide scanner can be conceptualized as a sequence of events together comprising the part of the imaging chain that represents the flow of information from a glass slide into a digital slide (U.S. Food and Drug Administration, 2016; Fiete, 2010a). A common digital slide scanner's imaging chain has the following components: glass slide – slide feeder – light source – optics – mechanical scanner movement – digital image sensor – image processing software – image composition software – digital slide (Figure 2). The rest of the imaging chain is made up of the digital pathology workstation.

Whole slide scanners are in essence robotic microscopes that capture individual images of tissue on a glass slide and then use special software to stitch the individual images seamlessly together into a whole slide image. The resulting digital slide can then be viewed using a digital pathology workstation. Whole slide scanners can be fully automated high-throughput laboratory instruments or standard light microscopes fitted with a digital camera and computer software to operate the scanning process. An overview of some commercially available scanners suitable for digital pathology is presented by Pantanowitz, 2017. There are also portable whole slide scanners (Meyer Instruments, n.d; Grundium, n.d; Holmström, et al., 2019).

Figure 2. An overview of whole slide imaging: whole slide scanner and digital pathology workstation



2.2.1 Glass slide

Standard histopathology glass slides are by convention 1 mm thick and 25 x 75 mm in dimensions, and have the stained tissue section, usually 3-4 micrometers thick, placed between the slide and a thin (usually 0.17 mm) coverslip. A standard in histopathology is to use formalin-fixed and paraffin-embedded (FFPE) tissue that is cut into thin sections with a microtome, stained (most commonly using hematoxylin and eosin stains), and then mounted onto the slide. The transparent glass slide and coverslip, and the thinly sectioned and stained tissue, make it possible to visualize the tissue in a light microscope (Muskett, 2012).

Although digital slide scanners normally employ standard histopathological glass slides, the use of digital imaging has permitted histopathological imaging without conventional slides, using techniques such as microscopy with UV (ultraviolet) surface excitation (Fereidouni et al, 2017).

2.2.2 Slide feeder

The glass slide is fed into the scanner either manually or using a slide feeder. The feeder allows transferring a batch of slides into the scanner without having to manually insert the slides one by one. Careless preparation of the slides can lead to problems in the scanning process (Pantanowitz, 2017). Air bubbles under the coverslip, for instance, can disturb the focusing of the images to be captured. Excess mounting media on the slide or protruding slide labels, in turn, can make the slide get stuck in the feeder.

These things can be thought of both as a weakness and a strength of whole slide imaging. Whole slide scanners are vulnerable to simple defects on the slide, for instance air bubbles under the coverslip, which a pathologist can work their way around using a normal light microscope. On the other hand, if the use of whole slide imaging puts more focus on careful preparation of slides in the pathology laboratory and leads to better quality slides, it can be seen as advantageous.

In order to be used in a pathology department with a high caseload, the slide feeder has to function reliably and have a high capacity. The options of using larger 50 x 75 mm whole mount slides and the ability to bypass individual slides in the feeder are favorable in a clinical setting (Pantanowitz, 2017).

2.2.3 Light source

The light source of a whole slide scanner has a lamp and a microscope condenser as its main components. The light source most commonly employs either LED (light-emitting diode) or halogen light bulbs that produce white light which is then gathered by the condenser and concentrated onto the tissue specimen. The light source and condenser follow the same optical principles as in a conventional light microscope (Abramowitz, 2003).

An exception to this is strobe illumination which is sometimes used together with the mechanical scanner movement in order to remove motion blur in capturing sequential images of the tissue (Varga et al., 2012). Strobe illumination means lighting the specimen for a short period of time – enough to expose the image adequately – and then shutting off the light, so as not to illuminate the specimen while it is moving in the camera field of view thereby eliminating motion blur.

Using normal visible white light to illuminate the microscopy specimen is called brightfield microscopy and is by far the most common method of lighting used in histopathology. Some whole slide scanners have the option of using fluorescent illumination and thus permit the use of immunofluorescence and fluorescence in situ hybridization (FISH).

2.2.4 Optics

The imaging optics of a digital scanner consist of a microscope objective lens and eventually other secondary lenses, such as a photo tube lens, that project an enlarged image of the tissue onto the digital image sensor. A modern microscope objective is a system of individual optical elements, or lenses, that together form the image (Abramowitz, 2003).

The system of different types of lenses is designed to produce a high-quality image and correct for many optical aberrations, or distortions in the image, produced by single lenses. Chromatic aberration, for instance, is the process of light of different wavelengths being refracted slightly differently by the lens. The three most common types of objectives, with increasing degrees of chromatic aberration correction, are called achromats, fluorites (or semi-apochromats) and apochromats (Abramowitz, 2003).

Another optical aberration that is corrected in some objectives is field curvature which means that the image of a flat object produced by a simple curved lens is not flat but curved. Flatness corrected objectives correct for this aberration bringing a wider area of the image into focus simultaneously. A high-quality plan apochromat objective thus corrects for much of both color and field curvature aberrations (Abramowitz, 2003).

Some important attributes of a microscope objective, besides the type of optical aberration correction, are its working medium, magnification and numerical aperture. The most common working medium, or the substance in which light travels from the specimen to the objective, is air, used by ”dry” objectives. There are also objectives designed for other media, such as immersion oil, which offer greater optical resolution. Dry objectives usually offer magnifications up to 60x whereas oil immersion objectives can produce magnifications of up to 100x (Abramowitz, 2003).

Numerical aperture describes the light gathering power and the optical resolution of a microscope objective (Abramowitz, 2003). Numerical aperture is defined mathematically as

$$NA = n \sin(\mu),$$

where ”NA” is numerical aperture, ”n” is the refractive index of the imaging medium (1.0 for air, 1.51 for immersion oil) and ”μ” is ½ angular aperture (or light gathering angle) of the objective (some authors use alpha, α, or theta, θ, instead of mu, μ). A high quality dry (n = 1.0) plan apochromat 20x microscope objective for instance can have an angular aperture of 98.2 degrees (μ = 49.1) and thus a numerical aperture NA of 1.0 x sin(49.1) = 0.75. The formula for calculating numerical aperture comes from the Abbe diffraction limit for a microscope,

$$r = \lambda/[2n \sin(\theta)] = \lambda/(2NA),$$

in which ”r” is the resolvable feature size and ”λ” is the wavelength of light (about 400-700 nm for visible light). This formula, created by the German physicist Ernst Abbe, describes the maximum theoretical spatial resolution for a light microscope in which the image is composed of light passing through an object (both directly and through diffraction) and being reconstituted (through interference) (Abramowitz, 2003; Wilson, 2016). There are microscopy techniques that circumvent the diffraction limit called super-resolution microscopy (Schellermeh et al., 2019). Using an objective with a numerical aperture of 0.75 to calculate the diffraction limit for green light (wavelength ~550 nm) gives $r \approx 367$ nm as the smallest resolvable feature size for the objective in question. Thus, the higher the numerical aperture of an objective is, the greater the spatial resolution it can produce.

2.2.5 Mechanical scanner movement

The mechanical scanner movement usually consists of robotics that take a glass slide from the slide feeder, move it under the optical path at a defined distance – so as to have the image in focus – and then return it to the slide feeder or a separate output tray for scanned slides. Together with the digital image sensor the mechanical scanner movement divides whole slide scanners into two main categories, namely tile based scanners and line scanners (Sucaet & Waelput, 2014b).

Tile based scanners capture rectangular image tiles, with a pixel resolution (or pixel dimensions) of for instance 2000 x 2000 pixels, which are then stitched together to form a complete whole slide image covering the entire digitized slide. Line scanners instead capture long and narrow image strips, for instance 4000 x 60 pixels in size, that are combined to make the digital slide.

To capture a vast number of individual images (in the order of hundreds or thousands for most histopathology slides) fast enough, every image frame can not be brought into focus individually because it would require too much time for the scanning process. A common strategy to deal with this issue is to only set the focus for some of the image frames, commonly predefining these into a focus map, either manually or automatically (Montalto et al., 2011). The rest of the focus values can then be calculated based on these points.

Folds in the thinly cut tissue, or the tissue not lying completely flat on the slide, can cause some parts of the slide to be digitized out of focus and also create visible seams between the individual digitized image frames. These of course can in turn hamper the interpretation of the image. Overall the focusing of the image requires a substantial amount of time and there is a tradeoff between image quality (i.e. sharply focused image tiles) and scanning time.

Z-stacking is a special case of focusing the image, where several different focal planes are captured individually to create a Z-stack (Khalbuss et al., 2011). This allows the digital slide to be examined in the Z plane, bringing the whole thickness of the tissue into focus. This can be advantageous in interpreting cytopathological samples (such as urine and pleural fluid) with the expense of longer scanning time and greater whole slide image file size.

2.2.6 Digital image sensor

The digital imaging sensor and its controlling electronics make up the part of a whole slide scanner that performs the analog-to-digital (A/D) conversion of image information. In other words, a sensor can take analog information in the form of visible light (photons) coming from the microscope objective, and turn it into digital information (electrical signals) that is ultimately stored as a digital whole slide image (U.S. Food and Drug Administration, 2016).

The imaging sensor serves the same purpose as in an ordinary digital camera e.g. on a mobile phone. A CCD (charge-coupled device) or a CMOS (complementary metal oxide semiconductor) sensor detects light by catching

photons on pixels (picture elements) made up of silicon that turn the photons into electrons and ultimately into an electrical signal (Fiete, 2010b).

The sensor is usually covered with a color filter array, such as a Bayer color filter, that defines which wavelength of light is registered by a certain pixel. Thus, initially every pixel only registers color of a certain wavelength; red, green, or blue. In the final image, data from surrounding pixels is used to estimate the full color (using all three additive primary colors red, green, and blue) of the pixel (Fiete, 2010b). The image produced by the sensor is a raster image meaning a rectangular grid made up of individual pixels (Sinard, 2017).

An important attribute of an imaging sensor is the size and number of its pixels because together with the optics, the sensor defines the optical resolution of the digital slide (Sellaro et al., 2013). A 20x objective and an imaging sensor with a pixel size of 5 micrometers (having rectangular pixels measuring 5 x 5 μm) thus produce an optical resolution of 0.25 micrometers per pixel. The smaller the individual pixels in the camera sensor are, and the greater the magnification is, the better is the resulting optical resolution (meaning that a greater number of pixels is used to cover a feature of a certain size on the slide).

2.2.7 Image processing software

Image processing software refers to computer software that handles individual images produced by the imaging sensor before they are stitched together to form the final whole slide image. The common functions of image processing software are shading correction and the setting of default image brightness, contrast, and desired color balance.

Shading correction refers to the act of producing an image with uniform illumination. Even if the whole slide scanner has been set up using Köhler illumination, for example vignetting (deficient illumination at the image periphery) can produce uneven image brightness (Leong et al., 2003). Shading correction can be achieved for example by using a blank image field to calibrate the illumination.

Color balance comprises the intensities of the colors in an image. White balance refers to the color balance stressing neutral colors – that is, colors that contain equal amounts of red, green and blue. Properties of the glass slide and light source, for instance, can cause color casts or tints of particular color in the image. Setting the color balance is a process whereby these are eliminated (Yagi & Gilbertson, 2008).

2.2.8 Image composition software

In order to create a single whole slide image the individual image tiles need to be stitched together seamlessly. Individual images can be captured with some overlap while specialized computer software is used to stitch them together (Steckhan et al., 2008). In the first generations of whole slide images the stitching was often suboptimal creating visible seams between the individual image tiles. Since then the image stitching algorithms have gotten more sophisticated creating essentially seamless boundaries between individual image tiles.

Some of the early file formats used for whole slide images were JPEG (Joint Photography Experts Group; JPEG, n.d.) and TIFF (Tagged Image File Format, Adobe Developers Association, 1992) (Sinard, 2017) which lacked support for larger amounts of image information: JPEG only supports files up to 65,535 × 65,535 pixels (Mozilla and individual contributors, 2021) and TIFF has a maximum file size limit of four gigabytes (Adobe Developers Association, 1992).

While there are different file formats for whole slide images, a standard is to use a pyramidal organization of data. This is done to counteract the problem of not being able to load the whole digital slide into the computer's working memory (DICOM, 2010).

In the pyramidal organization the scanned image information is represented with varying degrees of magnification with the image in its entirety (100% magnification) constituting the base of the pyramid and the magnification of the image decreasing going up the pyramid. The image information is thus represented several times with varying magnifications so that when it is viewed with a magnification lower than the maximum, a lower resolution image (instead of the maximum resolution image) is retrieved, thereby saving memory and access time (DICOM, 2010). The pyrimadial organization of data is supported in the DICOM (Digital Imaging and Communications in Medicine) medical imaging standard (DICOM, 2010). There are also ways of arranging the large amount of image data needed for whole slide images that do not employ the pyramidal organization, such as the JPEG 2000 imaging standard (Tuominen, 2012), that is discussed more later on.

2.2.9 Digital slide

Digital slides are raster images, meaning images composed of individual pixels, that each have their own color. The color depth (bit depth), defined as the number of

bits used to represent the color of a single pixel in a whole slide image, is usually 24. Since each bit is a binary digit (either 1 or 0) the number of possible colors in a 24 bit image is 2^24 = 16 777 216 (Sinard, 2017).

The dimensions of a digital slide are defined by the size of the digitized specimen and the optical resolution of the scanner. A histopathological tissue section can be for instance 10 x 20 mm in size. Covering this area with a tile based scanner using an image sensor with 2000 x 2000 pixel resolution (or number of pixels on the sensor), 5 x 5 micrometer pixel size and a 20x objective lens, creates 800 individual image tiles and a digital slide with a total pixel resolution of 40 000 x 80 000 pixels. The optical resolution (or resolving power) of the image is 0.25 micrometers per pixel.

The amount of storage space a digital slide takes up is determined by its pixel dimensions and color depth. Thus, a 24-bit image measuring 40 000 x 80 000 pixels takes up 24 x 40 000 x 80 000 bits = 76 800 000 000 bits = 9 600 000 000 bytes (1 byte = 8 bits) = 9600 megabytes = 9.6 gigabytes. Table 2 summarizes the effects of the specimen size and the whole slide scanner optical resolution (scanning resolution) on the resulting digital slide size on disc (the amount of storage space the digital slide takes up).

Table 2. The determinants of the size of a histopathological digital slide

glass slide specimen size	whole slide scanner optical resolution	digital slide pixel dimensions	digital slide size on disc*
10 x 20 mm	0.5 µm/pixel	20 000 x 40 000	2.4 gigabytes
	0.25 µm/pixel	40 000 x 80 000	9.6 gigabytes
20 x 30 mm	0.5 µm/pixel	40 000 x 60 000	7.2 gigabytes
	0.25 µm/pixel	80 000 x 120 000	28.8 gigabytes

* assuming a standard 24-bit color image

Whole slide scanners produce raster images, meaning images composed of a rectangular grid of individual pixels, each with an individual color. The available colors are defined by the color depth or bit depth and the used color space. A color space defines how the bits used to represent colors translate to different individual colors (Sinard, 2017; Pantanowitz, et. al., 2017; Gurcan, 2017).

The most common color space used in computers overall is RGB, with different variations, such as sRGB, Adobe RGB and ProPhoto RGB, having different numbers of individual colors. RGB is a trichromatic additive color space comprising the three primary colors red, green and blue in which all colors are produced by mixing the three individual colors in different intensities. Other color spaces are for instance CIE L*a*b* and CMYK. An image represented in a certain

color space can be converted into another color space (Pantanowitz, et. al., 2017; Gurcan, 2017; Cambride in Colour, n.d.).

The final step in the digitization process is saving of the image data as a computer file. A file format is a standardized way of representing this data. The simplest image file format is a raw file containing unprocessed data from the digital slide scanner image sensor. A raw file contains all of the information a digital imaging sensor can capture and has to be processed into a color space in order to be viewed and edited in a reproducible manner. Other common digital imaging file formats are TIFF, JPEG, JPEG 2000, and in the case of many commercial digital slide scanners, proprietary file formats based usually on the three aforementioned ones (Sinard, 2017; Gurcan, 2017).

2.3 Whole slide viewing

Whole slide images are viewed using a digital pathology workstation (Figure 3). The simplest possible workstation is a standard personal computer (or a mobile device such as a tablet or a smartphone) with no special components. A specialized digital pathology workstation can, in turn, include several medical grade displays, input devices (such as mouses and trackballs) and software applications for image viewing and manipulation as well as a software interface to the digital patient record and pathology laboratory information system (Pantanowitz et al., 2017).

When talking about the digital pathology workstation a distinction should be made between hardware and software, where the former is physical components (such as device motherboard, graphics card, keyboard and display) and the latter computer programs (or applications that can be stored and run by the hardware such as an image viewing program). A digital pathology workstation comprises in essence the following parts: computer environment – image review software – input devices – display.

Figure 3. Two different designs of a whole slide scanner: left Aperio CS2 and right Aperio VERSA, both by Leica Biosystems (screenshots from https://www.leicabiosystems.com/digital-pathology/scan/aperio-cs2/ and https://www.leicabiosystems.com/digital-pathology/scan/aperio-versa/, retrieved May 6, 2020)

2.3.1 Computer environment

The computer environment constitutes as its foundation a computer with its incorporated operating system, such as Microsoft Windows, Apple Mac or Linux. Also portable devices such as smartphones and tablets can be used as digital pathology workstations.

The hardware part of the computer environment affects the usability of viewing digital slides by way of input devices (keyboard, computer mouse, trackball), computing power (motherboard with processor and working memory, graphics processor) and output device (that is, type of display). Generally speaking, a digital pathology workstation does not require special hardware compared to standard up-to-date components.

The main software component of the computer environment is the operating system, which to some degree dictates what hardware as well as other software can be used with it. The most common platform for digital pathology workstations is Microsoft Windows, which is also the overall most common operating system used in desktop and laptop computers worldwide (Netapplications.com, n.d.).

2.3.2 Whole slide viewing software

Slide viewing software is the central part of a digital pathology workstation. Its output occurs through a computer display and its input is executed with diverse input devices. The software can be multipurpose, such as a web browser, or operated through one, or a dedicated software application operated from a hard disk drive.

Viewing digital slides is problematic because of their considerable size. This is because they are usually too big to be loaded onto a computer's working memory or random access memory. However when viewing a part of the image the whole image does not need to be retrieved from the hard disc or the internet. By organizing the image information in a pyramid made up of tiles (or in another format allowing for random image data access) makes it possible to retrieve only the tiles that are viewed at a time. This way of organizing image information makes it possible to use streaming, i.e. sending parts of the image data over the network while they are being viewed without having to download the entire image first. This on-demand sending of image data over the network to the client requires special data server hardware and software.

2.3.3 Input devices

Input devices are hardware that the digital pathology workstation uses to operate the slide viewing software. Most common ones are a standard keyboard and a computer mouse. There is, however, a multitude of options to choose from so as to make slide viewing as effortless as possible (Molin et al., 2015).

2.3.4 Computer display

The computer display is the final component of the imaging chain in whole slide imaging. The computer display is also a major factor differentiating whole slide imaging from conventional light microscopy where glass slides are viewed through ocular lenses or eyepieces. Modern computer displays usually employ LCD (liquid crystal display) or OLED (organic light-emitting diode) technology with the former being the most common one in both all-purpose consumer computers as well as dedicated digital pathology workstations (Tannas, 2020).

Most common LCD monitors can be more precisely characterized as LED-backlit TFT-LCDs or thin-film-transistor liquid crystal displays. These are flat screen displays lit by LED lamps and consist of individual pixels which in turn are made up of subpixels of usually red, green and blue colors produced using color filters. Each pixel is switched on and off using a liquid crystal, polarization filters and an electrical current controlled by a transistor (Tannas, 2020).

Further important attributes of a computer display are its physical size and pixel resolution. The size of a display is traditionally given as the length of the display diagonal in inches, for instance 24 inches (≈ 61 cm). The pixel resolution, in turn, is usually described as width by height, for instance 1920 x 1200 pixels. It can also be given in megapixels, where 1920 x 1200 pixels for instance would be 2.3 megapixels (1920 x 1200 = 2 304 000). A third way of describing the pixel resolution of a display is by the approximate number of pixels lengthwise, by which terminology 4K pixel resolution translates to 4000 lengthwise pixels. In routine diagnostics, monitors of 32-43 inches have so far turned out most practical based on pathologist feedback.

Modern digital pathology workstations usually employ a two-screen setup, where one monitor is used for viewing the digital slide and the other for displaying auxiliary information, such as the LIS (laboratory information system) interface with additional data on the patient case (Pantanowitz et al., 2017).

2.4 Whole slide image compression and storage

It is not uncommon for histopathological whole slide images to be over 20 gigabytes (GB) in size, and as such, they take up a lot of hard disk storage space (Pantanowitz et al., 2014). Large-scale use of whole slide imaging in a pathology department can generate tens (Stathonikos et al., 2013) or hundreds (Clunie et al., 2016) of terabytes (1 terabyte = 1000 GB) of image data each year, not including storage redundancy or backup, which further increase the hard disk footprint (García-Rojo, 2016; Chlipala et al., n.d.).

Image compression is a form of data compression which in turn is the act of reducing redundancy in data representation so as to facilitate data storage and communication. Some of the redundancy in image data is purely mathematical whereas some has to do with human vision (perceptual redundancy) that cannot perceive all visual data it is presented with. The beginning of data compression

research is usually credited to mathematician and engineer Claude Shannon's pioneering of information theory (Acharya & Tsai, 2005; Wolfram, 2002).

Image compression is carried out using compression algorithms (or codecs, from coding-decoding) that compress or code the original image into a compressed one. When viewing the compressed image decompression or decoding has to be carried out first. Data compression is achieved by minimizing redundancy in image data and by creating an approximation of the original data, thereby reducing the image entropy or the average information content per symbol (Acharya & Tsai, 2005).

2.4.1 Storage media

Digital slides are commonly stored on either magnetic hard disk drives (HDD) or flash-memory based solid state drives (SSD). Magnetic hard disk drives record data in the form of binary data bits by magnetizing a layer of ferromagnetic material (such as a cobalt-chrome-platinum-based alloy) on a disk called platter, made up of for instance aluminium. Magnetic hard disk drives thus employ a spinning disk and a swinging read-write arm. Solid state drives, in turn, use integrated circuit assemblies, and no moving mechanical parts, to store data. They instead make use of floating gate metal-oxide-semiconductor field-effect-transistors (MOSFETs), made up of silicon (Lacaze & Lacroix, 2014). The storage media can be local, such as hard drives in the desktop computer, or accessed through the internet, as is the case in cloud storage (Mahmoud & Xia, 2019).

A common consumer grade personal computer of the 2020s might have an SSD of 512 gigabytes, which would only hold 25 uncompressed whole slide images of 20 gigabytes each. Owing to the high costs of storing WSIs (Häger, 2016), digital archiving in a clinical setting may necessitate some form of image lifecycle management, such as deleting older WSIs from hard disks, or moving them to cheaper storage media, for example, magnetic tape (Stathanikos et al., 2013; Chlipala et al., n.d.). This, in turn, counteracts one of the main advantages of WSIs over glass slides, namely, ease of access.

2.4.2 Lossy and lossless compression

To save storage space WSIs are compressed using so-called lossy compression algorithms (García-Rojo, 2016; DICOM, 2010). Lossy image compression is

mathematically irreversible, meaning some image information is lost during the compression. Lossless image compression, on the other hand, is reversible, and no image information is lost in the process. To put it in another way, lossless data compression retains identical data whereas lossy compression results in an approximation of the original data (Acharya & Tsai, 2005).

The degree of data compression is generally expressed as a compression ratio defined as the uncompressed file size divided by the compressed file size. In histopathological whole slide imaging lossless compression generally yields compression ratios of 3:1–5:1 which is usually not thought to be enough and consequently lossy compression is used more often (DICOM, 2010).

2.4.3 Compression algorithms

The most common compression algorithms for whole slide images are JPEG and its successor JPEG 2000 (DICOM, 2010). JPEG is the acronym for ”Joint Photographic Experts Group” which includes the International Telecommunication Union (ITU), International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) (JPEG, n.d.). The group released the JPEG standard in 1992 and the first part of the JPEG 2000 standard in 2000 (JPEG, n.d; ISO, 2000).

The three central steps in lossy JPEG and JPEG 2000 compression algorithms are data transform, quantization and entropy encoding (Figure 4). The first and third steps can be thought of as providing lossless compression whereas the second step, quantization, is what achieves lossy compression (Acharya & Tsai, 2005).

Data transform means that the raster image data is transformed, mainly losslessly, from its representation in the spatial domain (as individual pixels in order) into the frequency domain. This means that the data is represented for instance by cosine waves (in the case of JPEG using discrete cosine transformation, DCT) or wavelets (in the case of JPEG 2000 using discrete wavelet transformation, DWT). By transforming the image data into the frequency domain it is possible to do further processing on the image and achieve satisfactory image compression. The transform in itself also achieves some image compression by reducing the data redundancy which exists in many images because of the tendency of neighbouring pixels to be similar (in ”smooth” image areas without edges) (Acharya & Tsai, 2005).

Quantization is the step where lossy image compression takes place. The transformed data is expressed by wave and wavelet functions whose coefficients determine the accuracy of the data representation. The less detailed these coefficients are, the more the image data is compressed. By discarding some image information the entropy (average information content per symbol) is reduced and the image compressed (Acharya & Tsai, 2005).

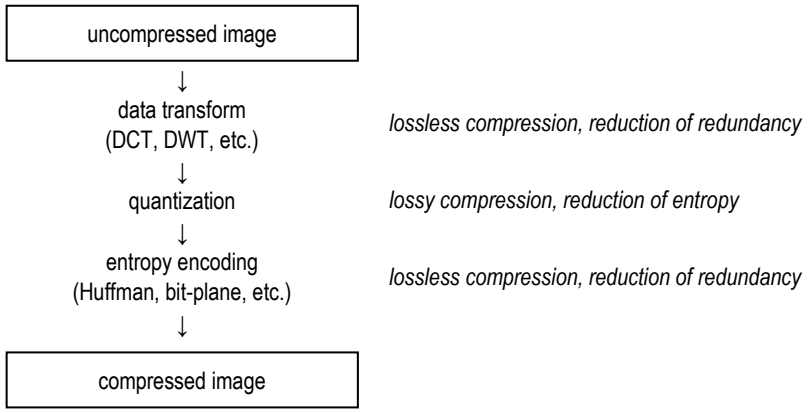
Entropy encoding, finally, achieves lossless image compression by finding a more compact representation of the compressed data. In its simplest form this can mean that a string of symbols, such as ”AAAAAA”, is shortened by not expressing every symbol individually but instead coding them in some way, for instance ”6A”. For entropy encoding JPEG uses classical Huffman coding and run-length coding whereas JPEG 2000 utilizes fractional bit-plane coding and binary arithmetic coding (Acharya & Tsai, 2005). A flowchart describing the process of an image compression algorithm is shown in Figure 4.

JPEG compression is often defined by a non-standard compression quality level, usually expressed as a value between 0 and 100, where the bigger the value, the better the resulting image quality is, and the less compression is applied. The compression ratio achieved with a given compression quality level depends on the image content and therefore the compression quality level is not directly proportional to the compression ratio (Acharya & Tsai, 2005).

Although there is no consensus regarding acceptable degrees of image compression for pathology WSIs, JPEG is thought to allow 10:1–20:1 and JPEG 2000 30:1–50:1 (mathematically lossy) data compression without loss of diagnostic information, sometimes referred to as perceptually or visually lossless compression (DICOM, 2010).

JPEG 2000 is especially suited for whole slide imaging because of its ability to handle very large images, such as whole slide images, and its use of a multi-resolution approach to image compression (Tuominen, 2012).

Figure 4. Flowchart describing the major steps in an image compression algorithm. DCT: discrete cosine transformation, DWT: discrete wavelet transformation



2.5 Digital image analysis

Computers were first used to analyze cells (Mendelsohn et al., 1965) and X-ray images (Meyers et al., 1964) more than 50 years ago. Since then substantial progress has been made in both computer performance (starting with the appearance of personal computers in the 1970s), and digital imaging, leading to considerable interest and technological advancements in computerized image analysis (Haidekker, 2010).

The National Cancer Institute at the National Institutes of Health defines digital image analysis as ”a method in which an image or other type of data is changed into a series of dots or numbers so that it can be viewed and studied on a computer. In medicine, this type of image analysis is being used to study organs or tissues, and in the diagnosis and treatment of disease” (National Cancer Institute, n.d.). This definition captures the important distinction between visual inspection and digital analysis of histopathological tissue specimens. The aforementioned comprises human visual analysis of forms, patterns and colors (represented in the two-dimensional spatial domain) while the latter involves mathematical analysis of the image data (represented as numbers), although in many cases mimicking or at least drawing inspiration from the principles underlying human visual analysis.

The Digital Pathology Association, in turn, citing The College of American Pathologists defines image analysis as ”computer-assisted detection or quantification of specific features in an image following enhancement and

processing of the image, including DNA analysis, morphometric analysis and FISH" (Digital Pathology association, n.d.). This definition, on the other hand, puts focus on the fact that image analysis can be thought of as a process of detecting and measuring structures in an image, thus representing morphometry, that is, quantitative analysis of form. Applications of digital image analysis in the field of pathology are summarized by Conway & Hewitt (2017).

Digital image analysis can be basically conceptualized as a three-step process (Bankman, 2009; Haidekker, 2010) with each step comprising a set of computer algorithms to carry out a specific task. This process starts with a digital image and ends with the output or result of the analysis algorithm. The three steps can be designated enhancement, segmentation, and quantification.

2.5.1 Preprocessing

The first step of digital image analysis, preprocessing, can also be called image processing, enhancement or image operations. This step involves processing the image in order to facilitate further steps in the analysis procedure. Image preprocessing can entail data transform from the spatial domain into the frequency domain and back, for example using wavelet transform (Bankman, 2009; Haidekker, 2010).

Examples of the uses of preprocessing algorithms are histogram manipulation, filtering and stain separation. Histogram manipulation can be used for image contrast expansion while filters are used to amplify or attenuate specific features in the image. Stain separation, finally, is used to separate different staining patterns in a histopathological slide and can be achieved for instance by color deconvolution (Haidekker, 2010; Tuominen, 2012).

2.5.2 Segmentation

Segmentation is the process of dividing the image into an object of interest and the background (which involves everything else). The object of interest is sometimes designated ROI (region of interest). Segmentation is done so as to distinguish biological entities of interest, such as cells, cell membranes, nuclei, and glands in the image (Gurcan, 2017).

Some common methods for segmentation are thresholding, region growing/splitting and edge based techniques. Thresholding involves segmentation

by pixel value (intensity-based segmentation) whereas region growing/splitting delineates similar areas in the image (region-based segmentation). Edge based segmentation acts by detecting boundaries in the image (Haidekker, 2010; Tuominen, 2012; Gurcan, 2017).

2.5.3 Quantification

Quantification is the final step in the digital image analysis process. It produces the output either in the form of measurement and/or classification. Measurements can include a wide variety of variables, such as size, form, area, number or proportion of cells, and cytoplasm/cell membrane/nucleus staining intensity as a continuous variable. Classification is done based on the measurements and can involve staining properties (positive – negative), and cell types (such as epithelial cell – fibroblast in the breast) (Bankman, 2009; Haidekker, 2010).

2.5.4 Artificial intelligence

Artificial intelligence in medicine refers to computational methods that perform tasks generally thought to require human intelligence (Tecuci, 2012; Bera et al., 2019). The term was coined in the 1950s (Tecuci, 2012) and encompasses different intelligence requiring processes such as perception, learning, and problem solving.

Machine learning is one area of artificial intelligence that has gained much interest lately. The term dates back to 1959 (Bera et al., 2019) and includes deep learning which makes use of artificial neural networks. Artificial neural networks can be thought of as computerized representations of human neural networks in which a form of learning can take place by the network adjusting its connections (Mareschal & Althaus, 2009). In digital image analysis artificial neural networks can be used either by themselves or together with existing conventional algorithms. Automatic Gleason grading of prostate cancer (Bulten et al., 2020) exemplifies the former approach whereas automatic detection of epithelial cells and their subsequent analysis using an existing non-artificial intelligence based image analysis algorithm (Valkonen et al., 2020), the latter.

Applications of artificial intelligence in digital pathology are reviewed by Bera and colleagues (2019) whereas Tizhoosh and Pantanowitz (2018) examine general challenges and opportunities facing such applications.

2.6 Histopathology of prostate and breast cancer

Prostate and breast cancer can be defined as malignant clonal proliferations of cells in the respective organ. In an overwhelming majority of cases the tumor is made up of epithelial cells and so is designated a carcinoma. Because the carcinoma usually grows in a glandular pattern it can be further characterized as an adenocarcinoma (Kumar et al., 2015; Lester, 2015; Epstein & Lotan, 2015).

The prostate and the breast can develop both invasive (or infiltrating) carcinoma as well as carcinoma in situ/intraductal carcinoma. The latter is a cancer that has not spread beyond the basement membrane underlying the epithelium and so does not grow invasively or metastasize and can be seen as a precursor to invasive carcinoma (Kumar et al., 2015).

The histopathological diagnosis of prostate and breast cancer includes the specific type of tumor (such as invasive acinar adenocarcinoma of the prostate or invasive ductal carcinoma of the breast) as well as the tumor grade (for instance Gleason score 3+4=7 in the case of prostate cancer or Nottingham grade 3 in the case of breast cancer). The tumor grade is a classification of the cancer's aggressiveness and is based on the notion of tumor differentiation: the less the cancer resembles its tissue of origin (the less differentiated it is), the more aggressively it grows (Kumar et al., 2015). The pathology report should also include other data on the tumor such as the molecular subtype of the tumor in a breast biopsy.

In order to standardize the histopathological diagnostics of cancer the World Health Organization publishes The WHO Classification of Tumors in the form of WHO Blue Books. The latest edition of the classification of tumors of the urinary system and male genital organs (including the prostate) was published in 2016 (Moch et al.) while the most recent edition of the classification of breast tumors was published in 2019 (WHO Classification of Tumors Editorial Board).

Besides the WHO there are also other organizations that publish guidelines so as to standardize the pathological diagnostics of prostate and breast cancer. The International Society of Urological Pathology (ISUP) as well as The Genitourinary Pathology Society (GUPS) publish guidelines on the grading of prostate cancer (van Leenders et al., 2020; Epstein et al., 2019) while The American Society of Clinical Oncology and The College of American Pathologists publish joint guidelines ("ASCO/CAP guidelines") on estrogen and progesterone receptor testing (Allison et al., 2020) as well as HER2 (human epidermal growth factor receptor 2) testing in breast cancer (Wolff et al., 2018).

2.6.1 Morphological diagnosis and grading

The basis of the histopathological diagnosis of prostate and breast cancer is in light microscopic examination of hematoxylin and eosin stained tissue removed most commonly through needle biopsy (Humphrey et al., 2016; Rakha et al., 2019). The main features that are analysed are cytological and architectural. Cytological (or cellular) features, that is the appearance of individual cells (their size, form and appearance of the cell nucleus) are examined at high magnification. Architectural features, such as growth patterns of groups of cells (for instance glandular formations), on the other hand, are examined at lower magnification. Finally, both immunohistochemical staining and in situ hybridization can be used as adjuncts to the morphological diagnosis of prostate and breast cancer.

Prostate cancer

Prostate cancer comes in different histopathological subtypes with usual acinar adenocarcinoma being by far the most common one. There are several variants of acinar adenocarcinoma, such as atrophic, pseudohyperplastic, and microcystic adenocarcinoma (Humphrey et al., 2016). Other types of prostate cancer including ductal adenocarcinoma, urothelial carcinoma, squamous cell carcinoma, neuroendocrine carcinoma, and sarcomas are only infrequently diagnosed.

Prostatic adenocarcinoma is graded using a system originally developed by Dr Donald Gleason in 1966. Gleason made various revisions to the grading system up till 1977 after which time it remained mainly unaltered until the 2000s. During the 1990s Gleason grading was universally adopted leaving behind other grading systems such as the Broders, the Mostofi and the MD Anderson systems (Kweldam et al., 2019). In 2004 Gleason grading was featured in the WHO blue book (Eble et al., 2004).

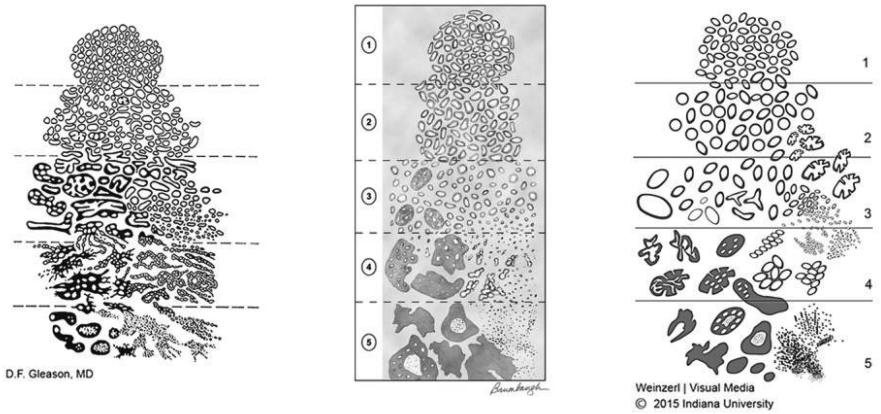
The significance of Gleason grading is evidenced by the fact that it is the single most important histopathological feature of prostate cancer. It provides both prognostic and predictive information about the cancer; prognostic information relates to the patient's prognosis of survival whereas predictive information concerns the likelihood of benefiting from a certain treatment. Gleason grading is also a central criterion in deciding whether a prostate cancer patient is treated actively (for instance with surgery or radiation therapy) or monitored (with so called active surveillance or watchful waiting) (Humphrey et al., 2016).

The Gleason grading system is based solely on the architectural growth pattern of the tumor (Gleason, 1966). It originally consisted of a composite score made up of the two most prevalent growth patterns, or grades 1-5, of the cancer, representing the differentiation of the tumor. Grade one designates a well differentiated cancer whereas grade five specifies a poorly differentiated one. If the cancer is only made up of a single growth pattern then the primary and secondary grades are the same.

The Gleason grade of a tumor is expressed as an equation: primary (most prevalent) grade + secondary (second most prevalent) grade = score, for instance 3 + 4 = 7. Originally the Gleason score ranged from two (1 + 1) to ten (5 + 5) but there were actually 25 possible scores (1 + 1...5, 2 + 1...5, and so forth) considering that 1 + 2 = 3 and 2 + 1 = 3 describe different types of tumors. When applying Gleason grading for prognostic or predictive information, however, the overall Gleason score has been mostly used (Kweldam et al., 2019).

While representing a continuous spectrum of a tumor's differentiation the cutoffs between different Gleason grades are arbitrary. Even though the different grades have been described both in writing and using drawings (see Figure 5 for an example of Gleason grading diagrams) it isn't possible to define exactly where one grade ends and another begins. This makes the grading prone to interpretation. The pathologist compares the morphology of the tumor to the textual and pictorial representations of the different grades in literature and assigns the grades that best suit the case at hand. Through the years it has become evident that there is considerable variation in the grading, both between individual pathologists (interobserver variation) and also in the same pathologists grading over time (intraobserver variation) (Bostwick, 1994; Gleason, 1992). The grading variation has been demonstrated for both general pathologists (Allsbrook et al., 2001a) as well as specialists in urological pathology (Allsbrook et al., 2001b).

Figure 5. Drawings of the Gleason grading system representing the different growth patterns of prostatic adenocarcinoma, left: the original Gleason grading system (1966-1977), center: the 2005 International Society of Urological Pathology (ISUP) modification, right: the 2015 ISUP modification (modified from Epstein et al; 2005, 2016)



With time and mounting evidence it became evident that the original findings underlying Gleason grading were not in keeping with modern prostate cancer diagnostics. When Gleason developed his grading system there was for instance no PSA testing (prostate-specific antigen, a blood test used in screening and diagnosing prostate cancer) nor immunohistochemical staining and the prostate biopsy procedure was different. This, in addition to the problems in the reproducibility of the grading, led to the need to revise the grading.

To bring the Gleason grading system up to date the International Society of Urological Pathology (ISUP) has so far held three consensus conferences on how to perform the grading. In the 2005 ISUP Modified Gleason system the five Gleason patterns were defined more comprehensively (Epstein et al., 2005). Prior to this, in 2000, the use of Gleason grades 1-2 started diminishing (Epstein, 2000). In the 2014 ISUP consensus conference a new grading system ("ISUP grading") with grade groups, based on Gleason grading, was proposed. Grade groups 1-5 corresponded roughly to Gleason scores 6-10 as Gleason scores 2-5 were not assigned any more on biopsies (Epstein et al., 2016). The latest ISUP consensus conference so far was held in 2019 and addressed issues such as reporting quantities of Gleason grades and reporting minor/tertiary patterns in addition to the primary and secondary grades (van Leenders et al., 2020).

Despite the consensus conferences, there are still controversies and uncertainty regarding the use of the Gleason grading system (Kweldam et al., 2019; Epstein et

al., 2020). These range from the correct assigning of grades to more basic phenomena of tumor biology such as the nature of intraductal carcinoma of the prostate and whether it should be graded at all (Delahunt et al., 2020). There are, also, some differences between different contemporary grading recommendations, for instance the ISUP and The Genitourinary Pathology Society (GUPS) recommendations (Smith et al., 2020).

Breast cancer

Breast adenocarcinoma is distinguished architecturally by a disturbance of the normal growth pattern of breast ducts and lobules (Lester, 2015; Rakha et al, 2019). The appearance can take many forms. Invasive cancer involves by definition invasive growth into the breast stroma. Cytologically breast adenocarcinoma is characterized by nuclear pleomorphism, that is, variability in the size, shape and staining of cell nuclei (and nucleoli).

Breast cancer can present in different morphological subtypes, such as lobular, tubular, cribriform, and mucinous carcinoma. The most common variant, however, is invasive breast carcinoma of no special type, also called invasive ductal carcinoma (Rakha et al, 2019).

According to the WHO, breast cancer should be graded using the Nottingham histologic grade, also called the Elston-Ellis modification of the Scarff-Bloom-Richardson grading system (Elston & Ellis, 1991; Rakha et al., 2019). The grading is based on one architectural feature, tubule formation, and two cytological features, nuclear pleomorphism and mitotic count. Each of the three features is assigned a score 1-3 and the three scores are summed up to make the final grade: scores 3-5 correspond to grade 1 (well differentiated tumor), scores 6-7 to grade 2 (moderately differentiated) and scores 8-9 to grade 3 (poorly differentiated tumor). The greater the grade and the score is, the more aggressive is the tumor.

Tubule formation is assessed semi-quantitatively: in score 1 cases the majority of the tumor (> 75%) exhibits tubules, in score 2 cases tubule formation is moderate (10-75% of the tumor) and in score 3 cases little or none of the tumor (< 10%) forms tubules. Nuclear pleomorphism divides the cancer cells into ones with small, regular, and uniform nuclei (score 1); nuclei with moderate increase in size and shape (score 2); and nuclei with marked variation in size and shape (score 3). Mitotic count is obtained by directly adding up mitoses from ten high-power microscopic fields of view and the mitotic frequency produces scores 1-3 depending on the field diameter (for instance for a field diameter of 0.5 mm ≤ 7

mitoses gives score 1, 8-14 mitoses score 2 and ≥ 15 mitoses score 3) (Rakha et al., 2019).

The tumor grade is one of several variables that offer prognostic and predictive information about breast cancer. Other variables include expression of hormone receptors and the HER2 receptor (human epidermal growth factor, a member of the HER family, HER1-4, of tyrosine kinase growth factor receptors) (Rakha et al., 2010; Rakha et al., 2019; Kreutzfeldt et al., 2020).

2.6.2 Immunohistochemistry and in situ hybridization

Prostate cancer

While hematoxylin and eosin (H&E) staining allows for careful examination of the tissue morphology, it is sometimes difficult to differentiate especially small foci of carcinoma from benign lesions mimicking cancer. Examples of these pseudoneoplastic conditions are atrophy and adenosis. By allowing the identification of specific cell types through staining of protein antigens immunohistochemistry (IHC) presents an important adjunct to the routine H&E staining in prostate histopathology (Humphrey et al., 2016).

Basal cells are usually absent in cancerous glands but they are not always easy to distinguish in H&E stained tissue sections. Fibroblasts for instance can be mistaken for basal cells. Also luminal epithelial cells can be difficult to tell apart from basal cells, whether because of pyknotic (shrunken) nuclei or because of their stratified formation resulting from tangential sectioning of the tissue. Staining proteins expressed by basal cells, such as high molecular weight cytokeratins (for instance 34betaE12) or the transcription factor p63, helps in differentiating the basal cells from benign luminal and carcinoma cells (Humphrey et al., 2016).

Besides the staining of basal cells, which are lacking in cancerous glands, immunohistochemistry can also be used to stain cells suspicious for cancer. AMACR (alpha-methylacyl-coenzyme-A racemase) is a protein overexpressed in prostatic adenocarcinoma cells and can be used by itself, or together with basal cell markers, to differentiate malignant glands from benign ones (Humphrey et al., 2016).

When intrepreting prostate immunohistochemistry the pathologist should always take into account the tissue morphology as examined by H&E staining (Moinfar, 2007; Pernick, 2020; Shah & Zhou, 2012).

Contemporary diagnostics of breast cancer includes molecular subtyping of the tumor irrespective of the morphological classification. The subtyping is generally done by immunohistochemistry and in situ hybridization (Szymiczek et al., 2020). Immunohistochemical staining is used to categorise breast cancers with regard to protein expression of hormone receptors (estrogen receptor, ER, and progesterone receptor, PR), the human epidermal growth factor receptor 2 (HER2, ERBB2) and the Ki-67 cell proliferation marker. In situ hybridization (ISH) in turn, usually either fluorescence in situ hybridization (FISH) or chromogenic in situ hybridization (CISH), is used to study the *HER2* status for potential gene amplification (Wolff et al., 2018). The molecular subtypes of breast cancer commonly used in breast cancer diagnostics are ”luminal A-like” (ER positive, PR positive, HER2 negative and low Ki-67 proliferation index), ”luminal B-like HER2 negative” (ER+, HER2-, high Ki-67 index and/or negative-low PR), ”luminal B-like HER2 positive” (ER+, HER2+), ”HER2 positive non-luminal” (ER-, PR-, HER2+) and ”triple-negative” (ER-, PR-, HER-) (Rakha et al., 2019).

Estrogen receptor, progesterone receptor and HER2 play a part in the growth of the cancer. Ki-67 is, in turn, a protein expressed in the active phases of the cell cycle and thereby a marker for proliferating cells (Lester, 2015). ER and PR are expressed in normal luminal breast epithelium but their expression can be lost during carcinogenesis (Hicks, 2011; Rakha et al., 2019). HER2 is also expressed in normal breast epithelium at low levels and can conversely be upregulated in cancer making it a proto-oncogene (Kumar et al., 2015; Rakha et al., 2019). HER2 overexpression is almost exclusively due to gene amplification on chromosome 17q12 (Lester, 2015).

The HER2 status of breast cancer offers both prognostic and predictive information about the tumor. HER2 positive cancers generally carry a worse prognosis than tumors without it. Most importantly HER2 positive cancers can be treated effectively with anti-HER2 therapies such as trastuzumab (a humanized monoclonal antibody targeting the extracellular domain of HER2, Herceptin®), lapatinib (a dual HER1/HER2 kinase inhibitor), pertuzumab (a dual HER2/HER3 monoclonal antibody), and trastuzumab emtansine (a conjugate of trastuzumab and the chemotherapeutic agent emtansine, T-DM1). Anti-HER2 treatments provide significant increases in survival rates (Rakha et al., 2019) and are generally thought to have brought about a paradigm shift in the treatment of breast cancer (Kreutzfeldt et al., 2020). However at least trastuzumab exhibits cardiotoxicity

which stresses the importance of correct qualifying of patients for anti-HER2 treatment (Dias et al., 2016).

Defining the HER2 status of a tumor is not always easy. For one thing, there is no consensus as to whether in situ hybridization or immunohistochemistry should be regarded as the ”gold standard” method of testing. For another, in situ hybridization is carried out both with a single-probe method (labeling the *HER2* gene) and with a dual-probe method (labeling *HER2* and the chromosome 17 centromere region, for instance with the *CEP17* probe) producing in total three sets of criteria for *HER2* amplification/overexpression (IHC, single-probe ISH, and dual-probe ISH). A third issue in HER2 diagnostics is the fact that there exists substantial interobserver variation in interpreting the immunohistochemical staining. In addition to the aforementioned issues the quality of the HER2 staining methods varies between laboratories which is why guidelines recommend the use quality control measures. The ASCO/CAP guidelines on HER2 testing in breast cancer were first published in 2007 (Wolff et al., 2007) in order to improve the accuracy of HER2 testing in breast cancer. They have since been updated twice, first in 2013 (Wolff et al., 2013), and subsequently in 2018 (Wolff et al., 2018).

2.6.3 Digital image analysis

Through the advancement of digital pathology it has become possible to use conventional digital image analysis as well as applications of artificial intelligence in prostate and breast cancer histopathology.

Breast cancer

The most widely used applications of digital image analysis in breast histopathology are algorithms for analysing immunohistochemical stains, especially for cancer diagnostics including the analysis of hormone receptor and HER2 status (Conway & Hewitt, 2017). In a supplement to the 2013 ASCO/CAP guidelines on HER2 testing (Data Supplement 7: IHC Interpretation Criteria) the use of quantitative image analysis is encouraged for cases with weak immunohistochemical staining (1-2+) to improve the consistency of the interpretation (American Society of Clinical Oncology & College of American Pathologists, 2014). The College of American Pathologists has since published a guideline on the use of quantitative image analysis in HER2 immunohistochemistry (Bui et al., 2019). Some of the image analysis algorithms used are free and open source meaning the source code is

publicly available whereas others are proprietary/commercial (Tuominen et al., 2010).

Prostate cancer

Both the ISUP and GUPS recommendations on prostate cancer grading discuss the application of artificial intelligence and other digital pathology methods to not only prostate cancer grading but also detection, quantification and prognostication (Epstein et al., 2019; van Leenders et al., 2019). In Gleason grading, for instance, deep neural networks have been shown to outperform general pathologists and match specialists in uropathology in grading accuracy (Nagpal et al., 2020; Ström et al., 2020). Deep learning has also been used to differentiate cancer from benign prostate tissue reliably (Campanella et al., 2019; Ström et al., 2020). Both the ISUP and GUPS recommendations conclude that although the applications of artificial intelligence are promising there is still work to be done before they can be applied to routine clinical practice.

3 AIMS OF THE STUDY

The aim of the study was to develop applications of digital pathology and test whether these can improve the histopathological diagnosis of prostate and breast cancer by:

1. advancing the teaching and standardization of Gleason grading of prostate cancer. (Study I: Whole slide imaging in Gleason grading)

2. aiding in the interpretation of immunohistochemical staining of prostate cancer. (Study II: Simultaneous viewing of whole slide images)

3. aiding in the interpretation of immunohistochemical staining of breast cancer. (Study III: Image analysis of HER2 immunohistochemistry)

4. facilitating the implementation of digital pathology in larger scale by

    a. introducing a novel whole slide image optimized image compression algorithm. (Study IV: WSI-optimized JPEG 2000 image compression)
    b. mapping the determinants of an optimal imaging resolution for whole slide scanners (Study V: Determining the optimal imaging resolution in a whole slide scanner)

4 MATERIALS AND METHODS

4.1 Tissue samples

Studies I-IV utilized histopathological tissue samples. The studies were carried out in 2004-2005 (Study I: Whole slide imaging in Gleason grading, and Study II: Simultaneous viewing of whole slide images), 2015 (Study III: Image analysis of HER2 immunohistochemistry), and 2017 (Study IV: WSI-optimized JPEG 2000 image compression). Studies I, III and IV made use of archived pathological tissue samples originally acquired in routine clinical practice. The samples were retrieved after they had been used for histopathological diagnosis and there was no intervention in the diagnostic or therapeutic process, or in the integrity of the persons from which the tissue samples came. In Study II new tissue sections were cut from archived tissue blocks and the blocks were selected so as to include sufficient amount of tissue for the extra sections and still have tissue left for potential additional sections needed in the future.

Study IV employed a set of histopathology slides without any accompanying patient information whereas the clinical data in studies I-III was restricted to pathological-anatomical diagnoses and specimen ID codes. The ID codes were handled for a brief amount of time until slide digitization after which purely anonymised data, that is, new unique specimen identifiers and their corresponding pathological-anatomical diagnoses, were used.

The tissue samples were acquired from Tampere University Hospital (courtesy of Heikki Helin), Seinäjoki Central Hospital (courtesy of Mervi Jumppanen), and Helsinki University Central Hospital (courtesy of Heikki Helin), with applicable permits from the institutions in question. The studies were carried out before the passing of the Data Protection Act (1050/2018) and the Act on the Secondary Use of Health and Social Data (552/2019). The studies did not employ any data pertaining to biobanks.

4.1.1 Whole slide imaging in Gleason grading (Study I)

Sixty-two consecutive patients who underwent radical prostatectomy preceded by a needle biopsy were identified in the patient records of Tampere University Hospital, Finland. For each patient a representative, archived, hematoxylin and eosin stained glass slide of a needle biopsy was selected to be digitized.

4.1.2 Simultaneous viewing of whole slide images (Study II)

Twenty-three cases with formalin fixed, paraffin embedded prostatic needle biopsies were identified in the archive of the Pathology Department at Central Hospital Seinäjoki, Finland. Specimens were selected so as to favor challenging diagnoses such as small focus carcinoma, PIN (prostatic intraepithelial neoplasia), ASAP (atypical small acinar proliferation), proliferative inflammatory atrophy and suspicion for carcinoma without a more detailed description. All diagnoses were based on the original hematoxylin and eosin staining only. In addition to challenging diagnoses, also common diagnostic entities were included. Ordinary 3 to 4 μm tissue sections were cut from paraffin blocks and mounted on charged SuperFrost™ Plus slides to avoid detachment of the tissue sections from the slides. The slides were stained routinely with hematoxylin and eosin, and digitized. Afterwards the tissue sections were destained, restained and digitized again, as described later on.

4.1.3 Image analysis of HER2 immunohistochemistry (Study III)

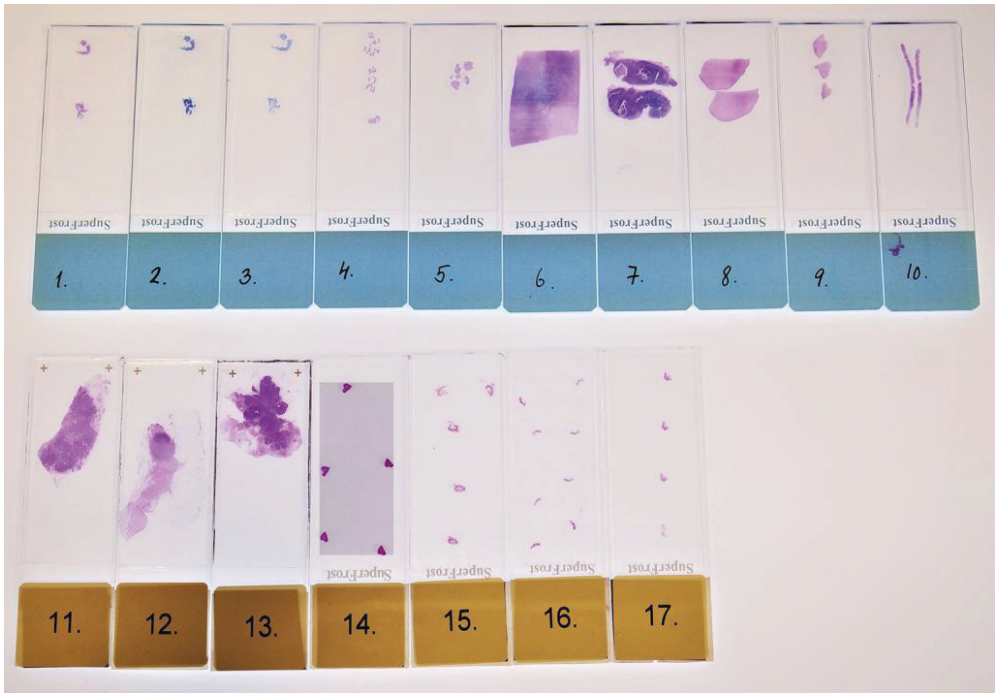
A database search was conducted at the Department of Pathology, Helsinki University Central Hospital, Finland, in order to identify invasive breast cancer cases tested for HER2 from the period of 1 January 2010 to 1 July 2011. The search yielded 1249 cases. The slides for the corresponding surgical resection specimens were retrieved from the archives, and starting from the earliest case, 750 consecutive cases (one slide per case) were included in the study. The 750 cases represented the period of 1 January 2010 through 23 May 2011, in which period 1186 breast cancer cases were histologically diagnosed. The missing 436 cases (1186-750) were not found in the archives at the time of retrieval and were distributed along the whole period of time taken for the study. The largest number of consecutive cases missing was 28, representing a period of 14 days. The

specimens had been routinely fixed for a period of 24–48 h in neutral formalin and embedded in paraffin.

4.1.4 WSI-optimized JPEG 2000 image compression (Study IV)

A set of seventeen histopathological slides were selected to be digitized from the Tampere university hospital pathology archive (see Figure 6). The slides were chosen to reflect the routine workload of a general pathologist and included biopsies (n = 10) as well as surgical sections (n = 7). The tissue sections were stained with hematoxylin and eosin (n = 15) and modified Giemsa (n = 2). In addition a gastric biopsy slide was digitized so as to test the visualization of Helicobacteria.

Figure 6. The seventeen routine histopathological glass slides used in Study IV: the shaded rectangle in slide 14 demonstrates the area that makes up a whole slide image consisting of both tissue and empty slide area



4.2 Immunohistochemistry and in situ hybridization

4.2.1 Simultaneous viewing of whole slide images (Study II)

After slide scanning the slide coverslips were removed by soaking the slides in xylene until they detached. The slides were then washed with absolute ethanol, air dried, and immersed in antigen retrieval buffer, composed of 0.5 M (molar) tris (tris(hydroxymethyl)aminomethane) and 1 mM (millimolar) ethylenediaminetetraacetic acid, pH 9. Antigen retrieval was done in an autoclave at 105°C for 5 minutes, followed by a 20-minute cooling period. This procedure also removed the hematoxylin and eosin staining completely. After rinsing and endogenous peroxidase quenching primary antibody incubation was done using a cocktail of antibodies to p63 (clone 4A4+Y4A3, Novocastra, Newcastle, United Kingdom), dilution 1:200, and to AMACR (clone P504S), dilution 1:200, for 30 minutes at room temperature. Antibodies were detected using a PowerVision+™ reagent kit according to the manufacturer's instructions. The slides were immersed for 10 minutes in DAB (diaminobenzidine) and enhanced with 0.5% copper sulfate for 5 minutes. The slides were counterstained with hematoxylin, dehydrated in graded series of ethanol, cleared in xylene, and coverslipped. The laboratory work was carried out at the Institute of Medical Technology, University of Tampere (cancer biology research group, professor Jorma Isola).

4.2.2 Image analysis of HER2 immunohistochemistry (Study III)

Immunohistochemistry had originally been performed according to the manufacturer's instructions using the BenchMark XT automated staining system (Ventana Medical Systems, Tucson, Arizona, The United States) with PATHWAY anti-HER2/neu (4B5) rabbit monoclonal antibody at a dilution of 6 µg/ml and Ventana ultraView Universal DAB Detection Kit (both Ventana Medical Systems). The slides were counterstained using Ventana Hematoxylin II (Ventana Medical Systems) as part of the automated staining procedure. The HER2 status of the specimens was established by immunohistochemical staining, and positive (3+) and equivocal (2+) cases were further subjected to in situ hybridization to classify them into positive and negative with regard to HER2 gene amplification.

In situ hybridization had been carried out in the BenchMark XT automated staining system using the INFORM HER2 DNA and the INFORM Chromosome 17 probes and the ultraVIEW SISH Detection Kit (all Ventana Medical Systems) according to the manufacturer's instructions. Consecutive sections of the specimens were hybridized with the probes for HER2 and chromosome 17, respectively, and the specimens' HER2 gene status was classified as amplified if the ratio of HER2 to chromosome 17 was over 2.2. Both the ISH and IHC assays had been subjected to external quality assessment by the Nordic Immunohistochemical Quality Control organization.

Both the immunohistochemistry and the in situ hybridization were carried out as part of the routine clinical practice at HUSLAB, Division of Pathology and Genetics, Helsinki University Central Hospital.

4.3 Whole slide imaging

Apart from Study IV, all of the whole slide scanning was done at the Institute of Medical Technology cancer biology research group (professor Jorma Isola). The web server in Study II was used courtesy of Johan Lundin and Mikael Lundin, Biomedical Informatics Group, Department of Oncology, University of Helsinki.

4.3.1 Whole slide imaging in Gleason grading (Study I)

A Zeiss Axioskop2 MOT microscope (Zeiss Gmbh, Göttingen, Germany) was equipped with a 40x NeoFluar® immersion oil objective and Märzhauser motorized specimen stage (Märzhauser, Wetzlar, Germany), which holds 8 standard microscope slides at a time. Images were captured at a 0.26-μm resolution with a CCD camera (Zeiss Axiocam HR; sensor size 6.8 x 7.6 mm; capture resolution 1300 x 1030 pixels). The camera was attached to the microscope with a 0.63x magnifying phototube. Image acquisition was controlled by the KS400 software (version 3.0, Zeiss) running on a standard MS Windows workstation. A custom macro command script running within the KS400 controlled for sequential autofocus, image acquisition, shading correction, and stage movement consecutively for 8 slides on the specimen stage. The acquired image files were digitally sharpened and stitched into a single montage file, which was then compressed. This robotic microscope was to the best of my knowledge the first

fully functioning whole slide scanner in Finland and probably in the Nordic countries as well.

4.3.2 Simultaneous viewing of whole slide images (Study II)

An Axioskop2 MOT microscope (Zeiss GmBH, Göttingen, Germany) was equipped with a NeoFluar® immersion oil 40 objective (numerical aperture 1.3) and a motorized specimen stage. A contiguous array of digital images covering the entire biopsy was captured at 0.3 μm per pixel resolution with a color sensor camera (capture resolution 1,300 x 1,030 pixels in 3 color scanning mode). The image capture process, i.e. stage movement, autofocus, shading correction and image capture, was automated using KS400 software, version 3.0 (Zeiss GmBH). The acquired image files were digitally sharpened and stitched into a single montage file, which was compressed. The compressed virtual slides were uploaded onto a web server running the Image Web Server software (Earth Resource Mapping Pty, West Perth, Australia). The virtual slide pairs (hematoxylin and eosin, and IHC) were exactly aligned for synchronized viewing by keeping the hematoxylin and eosin virtual slide as a reference and adjusting the position of the corresponding IHC virtual slide.

4.3.3 Image analysis of HER2 immunohistochemistry (Study III)

The 30 cases, one glass slide per case, were scanned as whole slide images using the Objective Imaging Surveyor with Turboscan digital slide scanner (Objective Imaging Ltd., Cambridge, UK) employing a ×20 Plan Apo microscope objective (scanning resolution 0.23 μm per pixel).

4.3.4 WSI-optimized JPEG 2000 image compression (Study IV)

The slides were digitized with whole slide scanners from four different vendors, including two line scanners (Aperio and Hamamatsu) and two tile-based scanners (Pannoramic and SlideStrider). The scanner setups were as follows:

1. Aperio ScanScope AT2 (Leica Biosystems, Nussloch, Germany; courtesy of Juha Näpänkangas, Department of Pathology, Oulu University

Hospital) brightfield line scanner, Piranha Color 2k PC-30-02K80 camera (Teledyne DALSA, Ontario, Canada) with 2048 × 3 pixel resolution, pixel size 14 × 14 μm, ×20 Olympus Plan-Apo objective lens with a numerical aperture (NA) of 0.75, and scanning resolution of 0.5 μm/pixel

2. Hamamatsu NanoZoomer XR (Hamamatsu Photonics, Hamamatsu, Japan; courtesy of Teemu Tolonen, Department of Pathology, Fimlab Laboratories, Tampere University Hospital) brightfield line scanner, charge-coupled device (CCD) camera with 4096 × 64 pixel resolution, pixel size 8 × 8 μm, ×20 Olympus Plan-Apo objective lens (NA 0.75) and ×1.75 relay lens, scanning resolution of 0.46 μm/pixel

3. Pannoramic SCAN (3DHISTECH Ltd, Budapest, Hungary; courtesy of Fang Zhao, HUSLAB, Division of Pathology and Genetics, Helsinki University Central Hospital) brightfield tile-based scanner, CIS 3CCD camera with 2048 × 2048 pixel resolution, pixel size 5.5 × 5.5 μm, ×20 Carl Zeiss Plan-Apochromat objective lens (NA 0.8) and ×1 phototube, scanning resolution of 0.24 μm/pixel

4. SlideStrider (Jilab Inc, Tampere, Finland; courtesy of Jorma Isola, Jilab Inc.) brightfield tile-based scanner, Lumenera Lt1265R CCD camera (Lumenera Corporation, Ottawa, Ontario, Canada) with 4240 × 2832 pixel resolution, pixel size 3.1 μm × 3.1 μm, ×10 (NA 0.4) and ×20 (NA 0.75) Olympus UPLSAPO objective lenses, scanning resolution 0.16–0.31 μm/pixel.

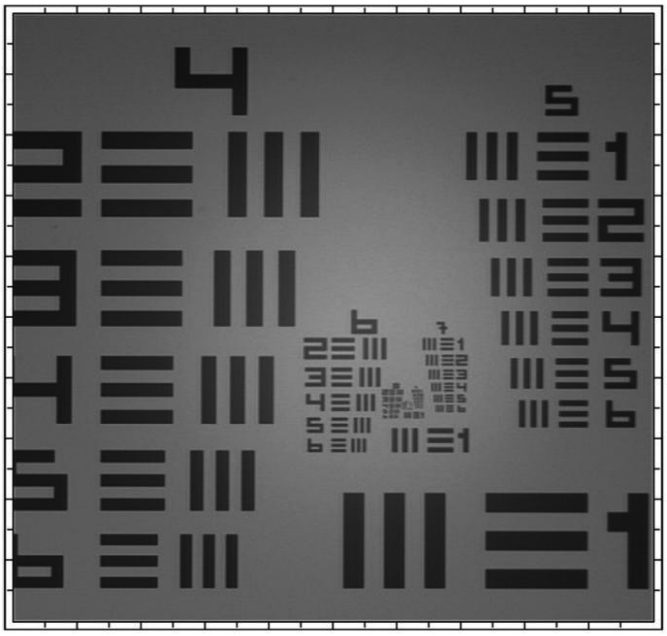
The area included in the WSI was the smallest rectangle covering all individual tissue fragments on the slide. For the SlideStrider scans, the non-scanned empty slide areas required to fill in the WSI rectangle were copied automatically from a standard empty slide image tile. All WSIs were scanned as 24-bit RGB color images.

4.3.5 Determining the optimal imaging resolution in a whole slide scanner (Study V)

The Extreme Resolution 1951 USAF Target slide (model 2012A; Ready Optics, California, USA; Figure 7) was digitized with the tile-based SlideStrider scanner

(Jilab Inc, Tampere, Finland) using Koehler-adjusted brightfield illumination, three objective lenses and two cameras, producing in total six whole slide images (pairing each camera with each objective lens). The cameras used were Lumenera Lt1265R CCD (pixel size 3.1 x 3.1 μm, 4240 x 2832 pixels in 1-inch sensor format) and Lumenera Lt425 (pixel size 5.5 x 5.5 μm, 2048 x 2048 pixels in 1-inch sensor format). The objective lenses were plan apochromats, Olympus UPLSAPO 10x NA 0.4, 20x NA 0.75, and 40x NA 0.95. All images were acquired with 8-bit pixel depth resulting in 24-bit color images. The images were saved as uncompressed bitmap images.

Figure 7. Photograph of the Extreme Resolution 1951 USAF Target slide. Element 1 from group 4 is placed in the lower right hand corner with the rest of the elements in group 4 (2-6) in the left side. Group 5 (elements 1-6) is to the right with the rest of the groups (6-11) being placed in a spiraling manner in the center. The line width of group 4 element 1 is 31 μm or 0.031 mm.



The target slide consists of 48 elements divided into 8 groups (groups 4-11) with each group being further divided into 6 individual elements (elements 1-6). Each element is made up of three horizontal and three vertical lines. The elements are

arranged according to their size with the biggest element being number 1 in group 4 and the smallest element 6 in group 11. The line width of the biggest element (group 4, element 1) is 31 μm and the period (the space between two lines) 62 μm; the corresponding width and period of the smallest element (group 11, element 6) are 137 nm and 274 nm. The groups and elements are arranged on the slide in a spiral fashion with the biggest elements (groups 4 and 5) in the periphery and the smaller elements getting successively nearer the center.

The sampling resolution of a slide scanner-objective combination was experimentally defined by looking up the dimensions of the smallest element in the resolution target whose individual lines could still be distinguished in the digitized image without them blending into each other.

4.4 Digital image compression

The image compression was carried out at the Institute of Medical Technology (studies I-III), University of Tampere (cancer biology research group, professor Jorma Isola) and Jilab Inc (Study IV), Tampere.

4.4.1 Whole slide imaging in prostate cancer Gleason grading & Simultaneous viewing of prostate cancer morphology and immunohistochemistry (Studies I-II)

After being scanned, the whole slide images were compressed using the ECW algorithm (enhanced compression wavelet) which was done by the ERMapper software (Earth Resource Mapping Pty, West Perth, Australia).

4.4.2 Image analysis of HER2 immunohistochemistry (Study III)

The whole slide images were compressed using standard JPEG 2000 compression.

4.4.3 WSI-optimized JPEG 2000 image compression (Study IV)

WSIs from the Aperio, Hamamatsu, and Pannoramic scanners were saved without compression and then compressed with JP2-WSI. The same scanned WSIs were also saved using the manufacturers' proprietary file formats and their default compression schemes. Aperio SVS format used JPEG tile compression with compression level set at 70/100. Both Hamamatsu NDPI format and Pannoramic MRXS format employed JPEG tile compression with quality level 80/100. The WSIs scanned with SlideStrider were first saved losslessly and then converted to either fixed 35:1 ratio JPEG 2000 or the developed JP2-WSI compression. The compression method of the SlideStrider software is based on the Kakadu software development kit library implementation of JPEG 2000 (version 7.5, Kakadu Software Inc., NewSouth Innovations Pty Limited, Sydney, Australia). The four scanners all had different sampling resolutions, Aperio 0.5 μm/pixel, Hamamatsu 0.46 μm/pixel, Pannoramic 0.24 μm/pixel, and SlideStrider 0.16–0.31 μm/pixel.

5 RESULTS

5.1 Whole slide imaging in teaching and standardizing prostate cancer Gleason grading (Study I)

A public website was constructed (http://www.webmicroscope.net/gleason) for the purpose of conducting the study and presenting its results. The website is no longer maintained and is only available for viewing through The Internet Archive (http://archive.org/). The website served as a platform for self-testing and learning of Gleason grading. The user could choose whether to randomly view slides from the pool of sixty-two slides included in the study or from a selected smaller set of biopsies representing an educationally more meaningful distribution of Gleason grades. The user could assign each biopsy a Gleason score and compare it with the scores given by the expert urological pathologists in the study. Grading could be conducted in an interactive learning mode with instant feedback, or the user could choose to assess their agreement with the expert urological pathologists after finishing a grading session. Additional features of the website included an option to save a scoring set to later return to and score the same slides again in random order and thus assess one's own intraobserver agreement. Figure 8 shows a screenshot of the summary statistics page of the website.

One of the expert urological pathologists viewed the sixty-two prostate biopsies from the original glass slides with a conventional light microscope and again six weeks later. The latter slide review was done with the pathologist's ordinary office computer, screen, and internet connection using the virtual slides on the website in a random order and without specimen codes. The intraobserver agreement (glass slides versus virtual slides) was excellent (weighted kappa = 0.73; 95% confidence interval [CI] 0.61-0.85) in that 48 (77%) of the 62 scores were identical and the 14 remaining scores differed only by 1 point on the Gleason scale. When the data were categorized into Gleason score groups (< 7, 7, > 7) as in Table 3, the weighted kappa was 0.75 (95% CI 0.63-0.88). Recorded from the web server log, the average time needed to grade a biopsy using a digital slide was 3 minutes.

Figure 8. Screenshot of the summary statistics page of the interactive Gleason grading website

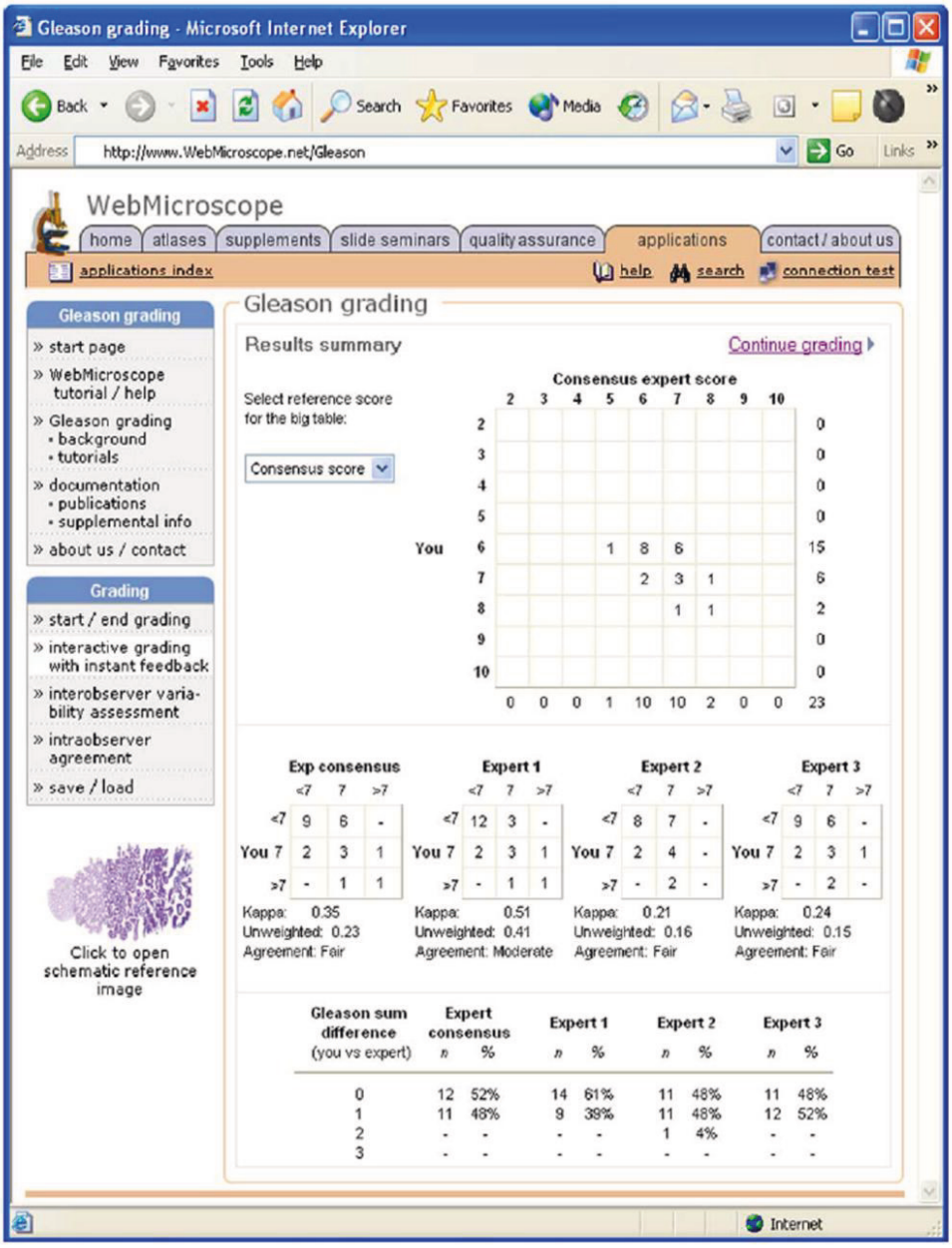


Table 3. Intraobserver variation in Gleason grading using digital slides over the internet vs. glass slides

Digital slide Gleason score	Glass slide Gleason score			Total
	< 7	7	> 7	
< 7	32	4	0	36
7	1	14	5	20
> 7	0	1	5	6
	33	19	10	62

The sixty-two digital slides were independently assessed from the website by three expert urological pathologists with their own ordinary office computers, screens, and internet connections. No efforts were made to standardize the Gleason scoring for the purpose of the study. The Gleason scores of the three pathologists were compared pairwise, that is, pathologist 1 versus pathologist 2, pathologist 2 versus pathologist 3, and pathologist 1 versus pathologist 3 (see Table 4). The corresponding weighted kappa coefficients for categorized Gleason scores (< 7, 7, > 7) were 0.66 (95% CI 0.50-0.82), 0.59 (95% CI 0.42-0.77), and 0.61 (95% CI 0.43-0.78), whereas the coefficients calculated for uncategorized Gleason scores were 0.62 (95% CI 0.47-0.78), 0.58 (95% CI 0.42-0.75), and 0.55 (95% CI 0.39-0.70), reflecting moderate to substantial interobserver agreement.

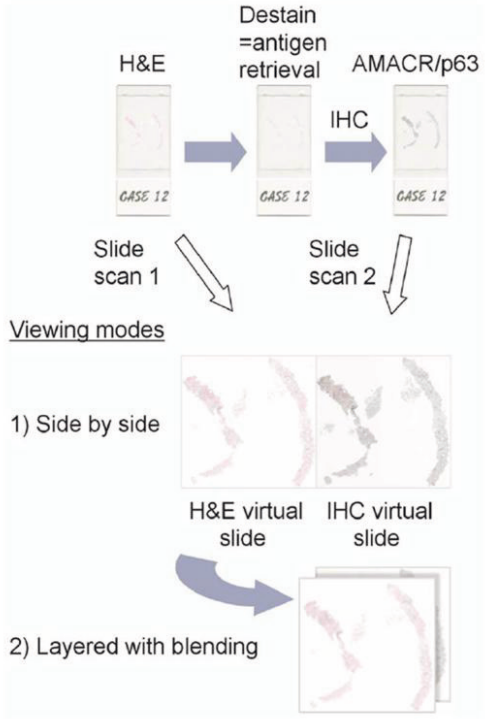
Table 4. Interobserver variation of Gleason grading using web based digital slides

	Gleason score	< 7	7	> 7	Interobserver agreement
Pathologist 1		Pathologist 2			
	< 7	23	5	0	
	7	5	21	1	0.66
	> 7	0	3	4	
Pathologist 3		Pathologist 2			
	< 7	25	11	0	
	7	3	16	1	0.59
	> 7	0	2	4	
Pathologist 3		Pathologist 1			
	< 7	25	10	1	
	7	3	16	1	0.61
	> 7	0	1	5	

5.2 Simultaneous whole slide imaging of prostate cancer morphology and immunohistochemistry (Study II)

The routinely hematoxylin and eosin stained slides were first digitized using an established whole slide scanning technique. The slide coverslips were then removed in xylene and the tissue sections destained (with the destaining doubling as antigen retrieval) and restained immunohistochemically with AMACR and p63, using a standard protocol. The quality of the following immunostaining was deemed similar to staining carried out on ordinary unstained tissue. After the immunohistochemical staining the slides were digitized again. Removing the coverslip and performing the AMACR/P63 immunostaining required approximately 3 hours, followed by slide scanning, which required 0.5 to 1 hour. The staining and scanning protocol is shown diagrammatically in Figure 9.

Figure 9. Diagram showing the process of creating whole slide images with simultaneous viewing of hematoxylin and eosin (H&E) and immunohistochemically stained (AMACR/p63) glass slides

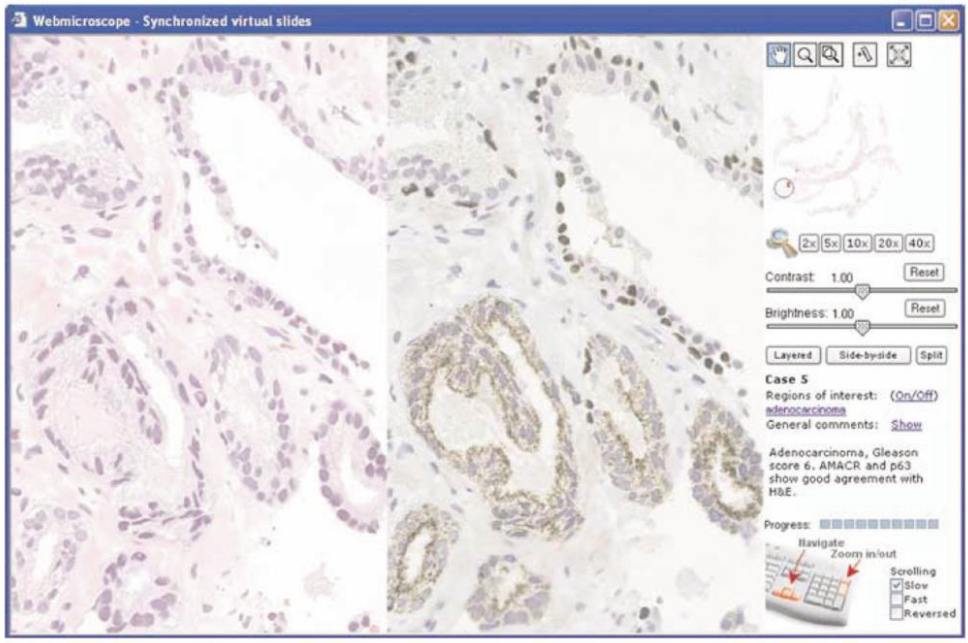


A publicly open website was created (http://www.webmicroscope.net/AMACRp63) for the viewing of the thirty digital slides. The website is no longer maintained and is only available for viewing through The Internet Archive (http://archive.org/). The paired digital slides consisting of hematoxylin and eosin, and IHC stains, could be viewed using a standard web browser (Internet Explorer or Mozilla® Firefox™). Figure 10 is a screenshot of the digital slide viewing interface.

The website user was able to view the stainings either side-by-side or blended on top of each other. In the blending mode, a slider was used to control the transparency of the immunohistochemical staining on top of the H&E staining. This yielded a result of seeing only the H&E staining, only the IHC staining, or a blend of the two. Almost perfect alignment of the hematoxylin and eosin, and the IHC stainings was achieved, allowing the comparison of the H&E staining morphology and immunophenotype even at a single cell level. When viewing the stainings side by side, zooming (magnification change) and navigation within the sample were synchronized, i.e. every movement took place in the hematoxylin and eosin, and IHC viewing windows simultaneously and similarly.

The website not only demonstrated the simultaneous viewing of sequentially stained digital slides, but also served as a self-educational tool for learning the correlation between prostate histopathology and the interpretation of AMACR/p63 immunostaining. To improve the pedagogic value of the website, an experienced uropathologist (HJH) defined and annotated regions of interest (ROIs) in the slides with foci of particular diagnostic interest. ROIs were indicated on the screen by circles and they represented carcinomatous and suspicious glands, PIN, ASAP and post-atrophic hyperplasia. ROI indicators and the annotations were initially hidden and could be turned on by the user.

Figure 10. Simultaneous viewing of two whole slide images with different stainings. Screenshot of the side-by-side viewing interface



5.3 Image analysis of HER2 immunohistochemistry (Study III)

Of the original database search result of 1249 cases, 157 cases (12.6%) were positive by ISH, and of the 750 case subset in the study, 74 (9.9%). In situ hybridization had been performed in 314 (41.9%) cases (originally 2+ and 3+ by IHC). The frequencies for a positive (3+) IHC staining for the original pathologist, the researcher and ImmunoMembrane image analysis were 8.4% (63 cases), 9.5% (71 cases) and 9.1% (68 cases), respectively. Equivocal (2+) staining was reported by the pathologist, researcher and ImmunoMembrane image analysis in 34% (255 cases), 43.7% (328 cases) and 10.1% (76 cases), respectively. Negative staining (0/1+) showed, for the pathologist, the researcher and ImmunoMembrane image analysis, frequencies of 57.6% (432 cases), 46.8% (351 cases) and 80.8% (606 cases), respectively. These data are summarised in Table 5.

Table 5. Frequencies of the HER2 IHC scores in percent, number of cases in parentheses

	Pathologist visual score	Researcher visual score	ImmunoMembrane image analysis score
Negative (0/1+)	57.6 (432)	46.8 (351)	80.8 (606)
Equivocal (2+)	34 (255)	43.7 (328)	10.1 (76)
Positive (3+)	8.4 (63)	9.5 (71)	9.1 (68)
Total	100 (750)	100 (750)	100 (750)

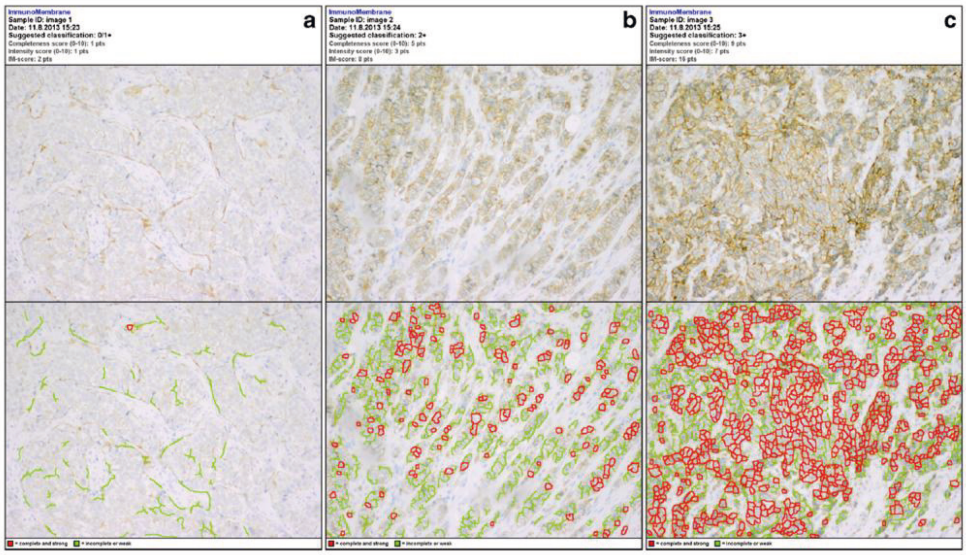
Of the 750 cases, 6 (0.8%) were false positives in ImmunoMembrane image analysis, defined as 3+ positive by IHC but negative by ISH. The corresponding number of false positives was for the researcher 8 (1.1%) and for the pathologist 0. Similarly, there were 6 (0.8%) false negative cases by ImmunoMembrane image analysis (0/1+ negative by IHC while positive by ISH). The number of false negatives was 0 for the researcher whereas it could not be defined for the pathologist because the laboratory performed ISH only on cases originally scored as 2+ or 3+ by IHC. A cross tabulation comparing the IHC scores of the pathologist, the researcher and ImmunoMembrane image analysis with in situ hybridization is presented in Table 6. Examples of images from the ImmunoMembrane analyses are shown in Figure 11.

Table 6. Cross tabulation comparing HER2 in situ hybridization to immunohistochemistry by visual scoring and by image analysis in numbers of cases

	Pathologist IHC			Researcher IHC			ImmunoMembrane IHC			Total
	0/1+	2+	3+	0/1+	2+	3+	0/1+	2+	3+	
No ISH	427	5	4	322	110	4	424	7	5	436
ISH -	5	235	0	29	203	8	176	58	6	240
ISH +	0	15	59	0	15	59	6	11	57	74
Total	432	255	63	351	328	71	606	76	68	750

The six false positive and six false negative cases (each representing 0.8% of the total 750 cases) in ImmunoMembrane image analysis were divided into four categories: truly discrepant IHC and ISH, borderline ImmunoMembrane IM-score, heterogeneous staining, and low-contrast staining. Of the six false positive cases, only two exhibited true discrepancy between IHC and ISH after review by an experienced breast pathologist, who scored the cases as 3+ positive in line with ImmunoMembrane while ISH was negative. In both cases, the original pathologist had scored the cases 2+ and the researcher 3+.

Figure 11. Three result images from ImmunoMembrane image analysis using digital photomicrographs. All three cases were originally scored as equivocal (2+) by a pathologist. The first image (a) receives an IM score of 2 points and is thus classified as negative (0/1+), the second image (b) 8 points (classified as equivocal, 2+), and the third image (c) 16 points (classified as positive, 3+). The first two cases were classified visually as 2+ by the researcher and the last one 3+. The first two cases are negative by ISH and the last one positive.



Two false positive cases showed a borderline IM score of 10 (in the scale of 0–20) which, according to our cutoff, value qualifies them as positive, although they were negative by ISH. In both cases, both the original pathologist and the researcher had scored the cases 2+.

Two false positive cases were found to exhibit heterogeneous staining when reviewed by an experienced breast pathologist. In these cases, some of the photomicrographs were classified by ImmunoMembrane in agreement with ISH whereas a greater number of images showed a negative score, which rendered the overall score in disagreement with ISH. One of the cases was scored 2+ by the original pathologist and 3+ by the researcher whereas the other case was scored 2+ by both.

Of the six false negative cases, one was found to be discrepant between IHC and ISH after review by an experienced breast pathologist (0/1+ IHC-negative in both the pathologist and ImmunoMembrane image analysis but positive by ISH). The case was scored 2+ by both the original pathologist and the researcher. Heterogeneous staining was observed in three false negative cases. All three cases

were scored 2+ by both the original pathologist and the researcher. Two false negative cases exhibited low-contrast staining when reviewed by an experienced breast pathologist. Both cases were scored 2+ by the original pathologist and the researcher.

Table 7 presents a cross tabulation of HER2 IHC scores rendered by a pathologist with the help of ImmunoMembrane using digital photomicrographs and by a researcher using ImmunoMembrane directly on whole slide images. Of the thirty cases two were discordant. Of these the first was scored 2+ using ImmunoMembrane with photomicrographs and 3+ using WSI and the second one vice versa (3+ using WSI and 2+ using photomicrographs). Interobserver variability, calculated using linearly weighted kappa statistic, was almost perfect at 0.92 (95% confidence interval 0.82–1.0).

Table 7. Cross tabulation of cases analysed with ImmunoMembrane using digital photomicrographs and whole slide images

ImmunoMembrane, photomicrographs	ImmunoMembrane, whole slide images			
	0/1+	2+	3+	Total
0/1+	10	0	0	10
2+	0	9	1	10
3+	0	1	9	10
Total	10	10	10	30

5.4 WSI-optimized JPEG 2000 image compression (Study IV)

Table 8 presents the pixel dimensions, the ratios of empty slide to tissue area, and the file sizes of the 17 slides digitized with the SlideStrider scanner at 0.31 μm/pixel. The uncompressed file sizes ranged from 2.6 to 30 GB. Lossless JPEG 2000 compression yielded compression ratios ranging from 3:1 to 56:1 and file sizes from 341 megabytes (MB) to 5.9 GB. The fixed ratio JPEG 2000 algorithm compressed all images to the 35:1 extent, except for two cases (slides 14 and 16), for which higher compression ratios were achieved with the lossless algorithm. The file sizes ranged from 74 MB to 686 MB. The developed JP2-WSI compression produced overall compression ratios varying from 41:1 to 1487:1, and file sizes of 8 MB to 442 MB. As an average, using JP2-WSI, we obtained file sizes that were 33% of fixed-ratio lossy compressed JPEG 2000. Of the individual scanned histopathology test slides, JP2-WSI reduced file sizes most effectively in biopsy slides containing multiple small tissue fragments and abundant empty slide area

(slides 14, 16, and 17 in our test set, see Figure 6). The ratio of empty slide area to tissue-containing slide area showed an approximately linear relationship with the overall compression ratio achieved with JP2-WSI.

Figure 12 allows comparison of the visual image qualities obtained with JP2-WSI compression, JPEG quality level 80 compression, and fixed 35:1 ratio JPEG 2000 compression. In this figure, visually detectable differences can be seen only with zoom levels well over 100%, which represent purely digital magnification. The magnified screenshots come from WSIs with file sizes of 528 MB and 21 MB (JPEG 2000 35:1 and JP2-WSI, respectively) and 611 MB and 18 MB (JPEG 2000 35:1 and JP2-WSI, respectively). The slides were scanned using the SlideStrider whole slide scanner with x10 objective lens and a charge-coupled device camera with 3.1 μm pixel size, resulting in 0.31 μm/pixel scanning resolution. Figure 13 presents zoomed screenshots of Helicobacteria in a gastric biopsy, scanned with resolutions of 0.31 μm/pixel and 0.16 μm/pixel (Plan-Apo ×10 and ×20 objective lenses, respectively). At 0.31 μm/pixel, there are subtle visible differences in the image quality. JP2-WSI eliminates random noise, resulting in a smooth or blurry image appearance, whereas JPEG produces a grainier or noisier image. With the higher optical scanning resolution we were unable to detect diagnostic differences in image quality.

Table 8. Comparison of whole slide image file sizes produced by three different parametrizations of JPEG 2000 image compression

Slide	WSI dimensions in pixels	Ratio of empty slide to tissue area	Un-compressed file size (MB)	JP2-lossless file size (MB)	JP2-35:1 file size (MB)	JP2-WSI file size (MB)
1	19,728 x 71,824	9.0	4251	387	122	13
2	22,544 x 67,600	10.1	4572	347	131	12
3	25,360 x 67,600	11.5	5143	387	147	12
4	19,728 x 80,272	6.7	4751	593	136	17
5	25,360 x 33,808	3.2	2572	448	74	12
6	64,784 x 92,944	0.3	18064	5545	517	442
7	67,600 x 67,600	1.0	13709	3431	392	289
8	56,336 x 63,376	1.0	10711	2382	306	120
9	19,728 x 59,152	1.9	3501	634	100	26
10	22,544 x 80,272	3.0	5429	928	155	38
11	64,784 x 92,944	1.2	18064	3512	517	304
12	67,600 x 118,288	0.6	23989	5896	686	284
13	67,600 x 67,600	0.2	13709	4091	392	248
14	67,600 x 147,856	49.0	29985	528	528	21
15	56,336 x 135,184	32.3	22847	669	648	25
16	61,968 x 143,632	49.0	26702	611	611	18
17	19,728 x 13,9408	24.0	8251	341	236	8

Figure 12. Comparison of image quality between JP2-WSI (a and b), JPEG quality level 80 (c and d) and fixed 35:1 ratio JPEG 2000 (e and f) compression. Digitally magnified screenshots of whole slide images 14 (a, c, e) and 16 (b, d, f).

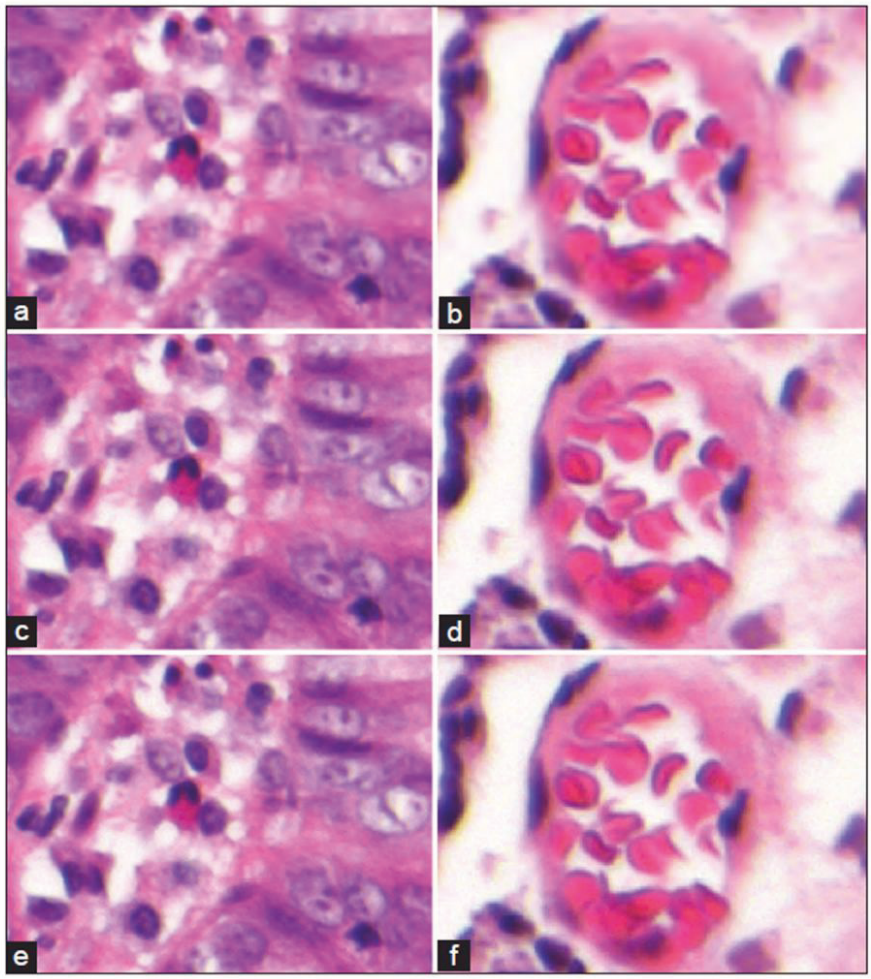


Figure 13. Effects of image compression and whole slide image scanning resolution on image quality exemplified by the detection of Helicobacteria in a gastric biopsy. JP2-WSI (a and c) and JPEG quality level 80 (b and d) compressed images scanned at 0.31 µm/pixel (a and b) and 0.16 µm/pixel (c and d) sampling resolutions.

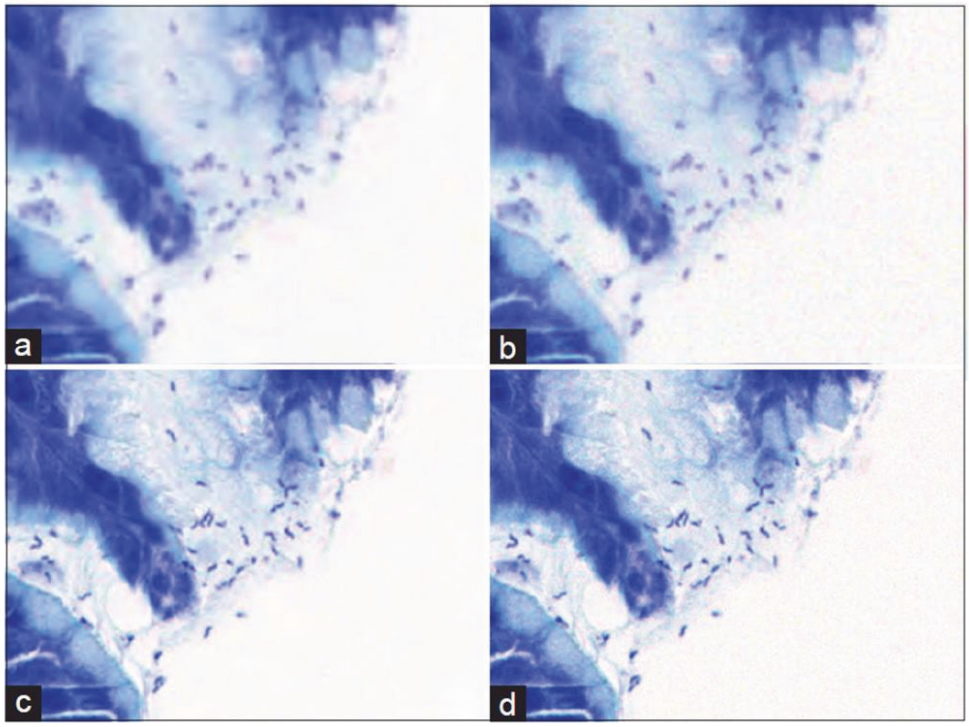


Table 9 shows the file sizes resulting from digitizing the set of 17 slides with 3DHISTECH, Aperio and Hamamatsu scanners. For each scanner, three different file sizes are shown per slide: the raw uncompressed file size, the file size using the scanner's default compression method, and the file size using JP2-WSI compression. The file sizes are not comparable between scanners because of different scan area dimensions and different scanning resolutions.

The uncompressed file sizes ranged from 4.84 GB to 63.15 GB. 3DHISTECH default compression produced compression ratios of 10–158:1 while JP2-WSI compressed the same images with ratios of 66–2250:1. For Aperio, the compression ratios were 7–32:1 for default compression and 48–1289:1 for JP2-WSI. Hamamatsu default compression produced compression ratios of 10–42:1 with JP2-WSI producing compression ratios of 49–1342:1. JP2-WSI compression had the widest range of overall compression ratios, 66–2250:1, 48–1289:1 and 49–1342:1 for 3DHISTECH, Aperio, and Hamamatsu scanned images, respectively.

JP2-WSI compressed images were smallest and had the highest overall compression ratios in every case. All of the compression methods produced the highest compression ratios for the biopsy slides 14–17. These slides had the highest ratios of empty slide to tissue area. JP2-WSI compressed file sizes were 15%, 9%, and 16% of the file sizes produced by 3DHISTECH, Aperio, and Hamamatsu default compression methods (Table 10).

Table 9. File sizes (gigabytes) of slides digitized with different scanner and compression method combinations with compression ratios in parentheses

	3DHISTECH			Aperio			Hamamatsu		
	Raw	Default	JP2-WSI	Raw	Default	JP2-WSI	Raw	Default	JP2-WSI
1	16.77	0.27 (63)	0.03 (643)	6.34	0.32 (20)	0.02 (270)	11.89	0.36 (33)	0.03 (446)
2	18.48	0.30 (62)	0.03 (678)	7.10	0.31 (23)	0.02 (333)	13.08	0.39 (34)	0.03 (498)
3	18.58	0.28 (66)	0.03 (709)	9.98	0.44 (23)	0.02 (479)	11.01	0.34 (32)	0.03 (432)
4	21.15	0.37 (57)	0.03 (649)	7.24	0.43 (17)	0.03 (243)	9.87	0.33 (30)	0.03 (301)
5	7.91	0.29 (28)	0.04 (221)	6.62	0.47 (14)	0.03 (205)	4.84	0.23 (21)	0.03 (143)
6	39.63	4.13 (10)	0.60 (66)	36.07	5.35 (7)	0.76 (48)	40.87	4.10 (10)	0.84 (49)
7	33.90	2.56 (13)	0.48 (70)	36.46	3.30 (11)	0.40 (91)	27.35	2.32 (12)	0.44 (62)
8	25.97	1.59 (16)	0.29 (89)	14.80	1.94 (8)	0.27 (55)	20.35	1.53 (13)	0.29 (70)
9	18.48	0.48 (38)	0.07 (271)	7.00	0.48 (15)	0.06 (116)	10.66	0.50 (22)	0.07 (151)
10	16.89	0.57 (29)	0.07 (238)	7.79	0.61 (13)	0.06 (124)	10.70	0.50 (21)	0.07 (151)
11	47.22	2.61 (18)	0.47 (100)	28.07	2.89 (10)	0.40 (70)	37.79	2.57 (15)	0.46 (83)
12	56.35	3.28 (17)	0.50 (112)	40.37	3.73 (11)	0.42 (96)	49.81	2.87 (17)	0.47 (105)
13	34.23	3.06 (11)	0.52 (66)	25.82	3.41 (8)	0.44 (59)	29.37	2.49 (12)	0.47 (62)
14	59.32	0.37 (158)	0.03 (1739)	43.14	1.48 (29)	0.03 (1289)	52.85	1.27 (42)	0.04 (1342)
15	62.33	0.51 (123)	0.04 (1641)	49.37	1.55 (32)	0.04 (1189)	49.68	1.23 (40)	0.05 (1033)
16	63.15	0.50 (127)	0.03 (2250)	37.20	1.56 (24)	0.04 (905)	49.68	1.24 (40)	0.05 (1078)
17	35.64	0.24 (146)	0.02 (2093)	13.22	0.58 (23)	0.02 (671)	16.21	0.41 (39)	0.02 (862)

Table 10. Mean whole slide image file sizes produced by different scanner and compression method combinations

Compression method	Mean whole slide image file size in gigabytes (mean compression ratio)		
	3DHISTECH	Aperio	Hamamatsu
No compression	33.88	22.15	26.24
Default compression	1.26 (27)	1.70 (13)	1.33 (20)
JP2-WSI compression	0.19 (176)	0.18 (122)	0.20 (131)

5.5 Determining the optimal imaging resolution for whole slide images

The scanner setup of the two alternative cameras and three different objective lenses produced in total six digital images, for which we calculated a number of parameters as well as measuring the sampling resolution from the uncompressed digital images (Table 11).

For each objective lens, the theoretical optical resolving power was calculated as the Rayleigh resolution limit (Sellaro et al., 2013; Murphy & Davidson, 2013a) using the formula

$$R = 0.61 \lambda / NA,$$

where R is the smallest resolved distance in μm, λ is the wavelength in μm, and NA is the numerical aperture of the objective lens. The theoretical optical resolving power ranged from 0.839 to 0.353 μm.

In the same manner, each objective lens and camera sensor combination has a theoretical resolving power (theoretical sensor resolving power using the microscope magnification). This is defined by the sensor pixel size, the objective lens magnification, and the Nyquist criterion for image sampling (Sellaro et al., 2013; Murphy & Davidson, 2013b) according to which a minimum of two pixels is needed to cover the distance to be resolved. Thus, a sensor with 5,5 μm pixel size and a 10x magnification produces 5,5 μm x 10 x 2 = 1,1 μm resolving power. The theoretical sensor resolving power ranged from 1.1 to 0.155 μm. Expressed in μm/pixel, these values produce the theoretical sampling resolutions of the scanner setups, which ranged from 0.55 to 0.078.

By comparing the optics resolving power to the camera sensor resolving power we attained a percent value illustrating the extent to which the resolving power of the optics is utilised by the camera sensor. Percentages below one hundred indicate that the camera sensor is the limiting factor of the total resolving power of the scanner and that the image resolving power is not limited by the optics (image formation through diffraction and interference). A value of 100% signifies that the sampling resolution and the optical resolution are matched according to the Nyquist criterion, and that the image resolving power is diffraction limited. Percentages over one hundred mean that the whole resolving power of the optics is utilised, and that there is additional ”empty” digital resolution, which is not matched by optical resolution (Abramowitz, 2003).

Table 11. Determinants of the resolving power and sampling resolution of a WSI scanner and sampling resolutions measured from a resolution target slide

Objective lens	Theoretical optical resolving power[1], µm	Sensor pixel size, µm	Theoretical sensor resolving power[2], µm	Optics resolving power utilised by sensor[3]	Theoretical sampling resolution[4], µm/pixel	Measured sampling resolution, µm/pixel (% of theoretical)
10x NA 0.4	0.839	5.5	1.1	76%	0.55	0.87 (63%)
10x NA 0.4	0.839	3.1	0.62	135%	0.31	0.49 (63%)
20x NA 0.75	0.447	5.5	0.55	81%	0.28	0.39 (72%)
20x NA 0.75	0.447	3.1	0.31	144%	0.16	0.27 (59%)
40x NA 0.95	0.353	5.5	0.275	128%	0.14	0.27 (52%)
40x NA 0.95	0.353	3.1	0.155	228%	0.078	0.22 (35%)

NA = numerical aperture. 1: According to the Rayleigh resolution limit calculated for 550 nm wavelength. 2: Smallest detectable object according to sampling theory. 3: Percentages below 100 mean that the camera sensor is the limiting factor in the sampling resolution whereas percentages over 100 mean empty digital resolution. 4: Sampling resolution calculated according to sampling theory using the measured total resolving power.

6 DISCUSSION

The present study was carried out with the purpose of developing and testing applications of digital pathology in order to improve the histopathological diagnosis of prostate and breast cancer. The first two studies, I and II, represent the very beginnings of digital pathology, and the field has advanced since then, perhaps in small part due to the findings from these studies (Weinstein, 2005; Camparo et al., 2012).

Cancer histopathology is a field under constant progress. The pathologist shortage observed in many countries as well as the need for ever more detailed characterizations of cancer present challenges for this advancement of diagnostics.

6.1 Web-based whole slide imaging in cancer histopathology

Historically, the histopathological diagnosis of cancer has been made with light microscopy. Transitioning from physical glass slides to whole slide imaging makes it possible to take advantage of all the features that a computer environment offers.

The studies I, II, and IV were carried out employing whole slide images. In Study II photomicrographs as well as whole slide images were used. Whole slide imaging can be thought of as the pinnacle of digital pathology: it translates the work done with light microscopes into the digital realm. Tissue microarrays and photomicrographs, used in some studies, for instance, only emulate the task of making a histopathological diagnosis on a complete tissue specimen.

WSIs have been used for quite some time in certain environments, such as undergraduate teaching of histology and pathology (Saco et al., 2016; Nauhria & Ramdass, 2019). However, much of the material on Gleason grading is still photomicrographs and textual descriptions. The International Society of Urological Pathology, for example, has a Pathology Imagebase with a section on prostate cancer grading aiming to standardize the grading (The International Society of Urological Pathology, n.d.). The Imagebase is made up of photomicrographs. The WHO Classification of Tumors, as another example, is published in book form

while some volumes are also published on a website (International Agency for Research on Cancer, 1965). The ones that have been published online feature almost exclusively photomicrographs instead of whole slide images.

Moving from the light microscope to a computer enables, among other things, more flexible viewing of the histopathological sample. Using a computer display instead of microscope oculars confers advantages in ergonomics as well as taking advantage of the computer graphical interface in annotating the slides and using the cursor to indicate areas of interest. The simultaneous viewing of a single slide by several people without a multi-viewer microscope is also advantageous.

A novel application of digital pathology – compared to light microscopy – is the viewing of two slides simultaneously and synchronously. Some light microscope stages permit having two slides on at the same time but moving between the two – and locating the corresponding tissue areas – is cumbersome at best. After the methodology of simultaneous viewing of digital slides (Study III) was published, it has become almost standard in commercial whole slide imaging systems. Our method involved aligning the two slides partially by hand whereas the currently used applications do this fully automatically. A novel application of the simultaneous viewing of two whole slide images in breast cancer HER2 diagnostics is the viewing of an immunohistochemically stained section simultaneously with an in situ hybridization section.

The method of destaining and restaining histopathological sections has been described before (Halmi, 1978). Even though its use in routine work is limited by labour intensiveness it can have practical implications, for instance, in cases where there is insufficient amount of tissue to cut additional sections for new stains. The method can also be applied in education and research. We showed that the destaining procedure also functions as antigen retrieval in immunohistochemistry.

The Studies I and II used web-based software which made them accessible through the internet. With high speed internet access being practically ubiquitous it is possible to operate applications such as whole slide viewers over the internet even when fast transfer of large amounts of data is needed. Compared to locally installed software web-based applications are available to a much broader audience benefiting the dissemination of information. Some hospital networks, for instance, have firewalls and other security measures that prohibit installing software onto local workstations which further increases the feasibility of internet based software applications.

Whether considering on-site use through local area computer networks or web based applications such as off-site diagnostics or second opinion use, the

introduction of whole slide imaging into routine clinical practice has been slow (Griffin & Treanor, 2017; Jahn et al., 2020). Even though the shift from glass slides to whole slide images is in many ways advantageous for routine diagnostics, the process of going digital can be strenuous (Stathonikos et al., 2021).

In Finland, the first digital pathology systems – then called virtual microscopy – were developed in the early 2000s with the first scientific publications documenting the technology published in 2004 (Lundin et al., 2004a; Lundin et al., 2004b). Since then, digital pathology has been successfully used in Finland in pre- and postgraduate medical education, research, and clinical pathology quality control (Tolonen et al, 2021). However, the introduction of whole slide imaging into routine histopathology practice has taken over fifteen years, with the first workflows having been adopted into clinical practice in the 2010s and 2020s in Oulu University Hospital, Hyvinkää Hospital, Tampere University Hospital, and Turku University Hospital (Teemu Tolonen, personal communication, 18 January, 2021).

6.2 The evolving nature of histopathological and immunohistochemical entities

Since the present study was carried out, both Gleason grading of prostate cancer and HER2 diagnostics of breast cancer have changed. This is part of the nature of histopathological grading and immunophenotyping of malignant tumors: as more knowledge is acquired, changes are made to the diagnostic/grading criteria. The histopathological grading and immunophenotyping is in essence subjective definitions agreed upon by the experts in the field: the biological entities themselves express continuous variability, whereas the histopathological and immunohistochemical definitions are made up of categories with subjectively defined cutoffs.

Study I dealt with teaching and standardizing Gleason grading in prostate cancer. Since the study was carried out, The International Society of Urological Pathology (ISUP) has held three consensus conferences on how to perform the grading (Epstein et al., 2005; Epstein et al., 2016; van Leenders et al., 2020). The Genitourinary Pathology Society has also published its white paper on the contemporary grading of prostate cancer (Epstein et al., 2020). Even though the details of how to perform the grading have changed after Study I was published,

the basic principle of assigning grades according to the architectural growth patterns of the cancer has remained intact.

Despite the consensus conferences and other recommendations, there are still issues to be resolved regarding the exact use of Gleason grading (Kweldam et al., 2019; Sopyllo et al., 2020; Epstein et al., 2020; Smith et al., 2020). The methodology presented in Study I can be of value in spreading guideline recommendations and in proficiency testing.

The three versions of the ASCO/CAP guidelines on breast cancer HER2 diagnostics published so far – in 2007 (Wolff et al., 2007), 2013 (Wolff et al., 2013) and 2018 (Wolff et al., 2018) – have gone back and forth in how the HER2 status of a cancer should be determined. Study III was published after the issuing of the 2013 guidelines and subsequently the 2018 guidelines have changed the definition of the HER2 status. This change was mostly confined to the definition of HER2 positivity using in situ hybridization. However, the definition of IHC category 2+ (equivocal staining) was changed from "circumferential membrane staining that is incomplete and/or weak/moderate and within > 10% of tumor cells or complete and circumferential membrane staining that is intense and within ≤ 10% of tumor cells" (2013) to "weak to moderate complete membrane staining observed in > 10% of tumor cells" (2018) (Wolff et al., 2018). It seems that there are still unresolved issues in how to exactly interpret the cutoffs between the IHC categories (Taylor et al., 2021). More research is warranted, including evaluating the methodology presented in Study III, in order to clarify the best way of determining the HER2 status of a breast cancer by immunohistochemistry.

6.3 Black boxes and proprietary technology

Artificial intelligence, especially deep neural networks, hold great promise for advancing the histopathological diagnosis of cancer. In the case of prostate cancer, neural networks have been shown to be able to differentiate cancer from normal tissue and assign Gleason grades with high sensitivity and specificity (Nagpal et al., 2020; Ström et al., 2020; Bulten et al., 2020).

In theory, neural networks can analyze histopathological images with more precision than human pathologists. A neural network registers every single pixel in a digital image and can process vast numbers of individual images in a matter of seconds.

A problem with diagnoses and other categorizations made using neural networks is that the process by which the network arrives at a conclusion (such as an image representing ”cancer” versus ”not cancer”) is not easy to follow for a human observer. Whereas a human pathologist registers commonly accepted anatomical entities (cells, nuclei, nucleoli) and pathological changes (dedifferentiation or anaplasia, invasion, inflammation), a neural network can, at least in theory, analyze aspects with no clear correlates in conventional pathophysiology. This is sometimes referred to as the black box nature of artificial neural networks (Sakamoto et al., 2020; Liu et al., 2019; Block, 2005). Making diagnoses such as cancer, that can impact a person's life profoundly, without being able to clearly state what the diagnosis is based on, is by no means unproblematic.

In a broader sense, the black box theory can be applied to all forms of image analysis algorithms, not merely those using neural networks. By ”looking inside” the black box one can find that algorithms can exhibit different forms of bias (Danks & London, 2017) and that they are not as objective as sometimes made out to be (Tadrous, 2010). One incentive for using black box technology, or at least not actively avoiding it, could be financial gain. This is exemplified by ”trade secrets” giving an advantage to one proprietary technology over another.

A contrasting approach to black box applications and proprietary technology is presented by free and open source software. It is defined as software distributed under terms that allow users to use, modify, and redistribute the software freely (Feller et al., 2005). Free and open source software, such as the ImmunoMembrane application used in Study III, can be seen as more conducive for the scientific endeavor than proprietary software (Kelty, 2005). The advantages include concrete aspects (such as vendor lock-in) as well as more abstract ones (for instance peer review). The notions of freedom and openness can be extended from software to hardware as well as publishing and educational resources (Marée, 2019; Pearce, 2020).

6.4 Striking a balance between whole slide image fidelity and file size

With advancements in digital cameras and, by extension, in whole slide scanning, the fidelity of whole slide images has reached a point where greater scanning resolution does not automatically equal better usability in the clinical setting. Too big scanning a resolution produces too big images with more time needed to scan,

view and upload/download the images. Too big whole slide images also make up a greater storage footprint. The more widely digital pathology is adopted in a hospital setting, for instance, the more important it is to optimize the file size of whole slide images.

The two main approaches to optimizing whole slide image file size are using image compression algorithms and avoiding excessive optical resolution in the first place. In Study IV, a novel whole slide image optimized parametrization of JPEG 2000 was introduced and found to have advantages compared to conventional compression methods used in whole slide scanners.

When it comes to optimizing the scanning resolution of a whole slide image, an important aspect is the matching of the components in the imaging chain of the whole slide scanner, particularly the camera and the objective lens. To study this, we digitized an optical resolution target using a whole slide scanner with two alternative cameras and three different objective lenses.

The sampling resolution of a WSI can either be expressed as a theoretical value calculated from the sensor pixel size and the objective lens magnification, or obtained empirically by measuring the total resolving power of the scanner (in μm) using the USAF 1951 Extreme Resolution Target slide (see Figure 7). By matching the measured total resolving power to the corresponding sampling resolution given by the Nyquist criterion, we were able to compare the theoretical and measured sampling resolutions (see Table 11).

By optimizing the sampling resolution, one can make sure that the image resolving power is diffraction limited (that is, the whole resolving power of the optics is utilised) and that there is no empty digital resolution (that is, too many pixels to cover the smallest image detail discernible by the optics) (Abramowitz, 2003).

Our study showed that by using a camera sensor with too large a pixel size (5.5 μm for 10x NA 0.4 and 20x NA 0.75 objective lenses), the resulting resolving power is limited by the camera sensor, meaning only part of the resolving power of the optics is utilised (76% for the 10x and 81% for the 20x objective lenses). By using a camera sensor with a smaller pixel size (3.1 μm for the 10x and 20x objective lenses), the optics resolving power is obtained, meaning the camera sensor has 2 to 3 pixels covering the smallest image detail discernible by the optics (translating to 100-150% optics resolving power utilised by sensor in Table 11). In the case of the 40x NA 0.95 objective lens, the camera sensor with 3.1 μm pixels produces so called oversampling, or empty digital resolution. Oversampling can be defined as the use of a sampling resolution greater than 1-1.5 times the Nyquist

criterion (that is, over 100-150% utilization of the optics resolving power by the camera sensor). The extra pixels do not convey additional image information because the resolving power of the optics has already been met, but they increase the image file size. Thus, with the 40x objective lens, the camera sensor with 5.5 µm pixel size is better.

An additional observation when it comes to sampling resolution is, that the theoretical values are not reached empirically. With very high sampling frequencies the limits posed by the laws of optics are reached. The so called Abbe limit,

$$d = \lambda / 2NA,$$

where d is the smallest periodicity in a structure that can be detected in the microscope image, λ is the wavelength of light being used, and NA is the numerical aperture, gives 0.275 µm as the resolving power limit for an objective lens with NA 1.0 used with light having a wavelength of 550 nm. This is in keeping with previous observations that the spatial resolution limit of a light microscope is roughly half of the wavelength of the light used for illumination (Schermelleh et al., 2019).

The higher the theoretical sampling resolution is, the smaller is the proportion of it being realised in practice. This is in part due to empty digital resolution, while camera sensor noise can also be partly responsible.

WSI sampling resolution, meaning the resolution in µm/pixel at which a WSI is digitized, is, together with image compression, an important factor in determining WSI file size and image quality. By matching the sampling resolution to the resolving power of the scanner optics, the file size and spatial resolution of a WSI can be optimized. When the sampling resolution is too low in relation to the optics resolving power, the camera sensor limits spatial resolution, and pixelation occurs, even when the optics would have allowed for more detail to be displayed. In the opposite scenario, too high a sampling resolution results in oversampling, creating empty digital resolution, which means redundant pixels that do not convey extra image information but that do increase the file size. In determining the exact sampling resolution of a scanner, the theoretical calculations do not seem to be reached in practise, and therefore, it is better to use an experimental procedure.

6.5 The rationality of digital pathology technology

Digital histopathology represents new technology that is meant to augment or replace existing diagnostic equipment, namely the light microscope. Despite the many advantages presented by working with computers it is worthwhile to consider the presuppositions underlying the use of this new technology and whether ”going digital” is a rational choice or not. Light microscopy is, after all, compared to many digital pathology solutions, a reliable, cheap, and easy-to-use technology.

A framework for considering the rationality of technology is presented by the Swedish philosopher Ingemar Nordin (Nordin, 1988). As the goal for rational technology he designates usefulness, defined as techniques producing their effects reliably and efficiently. The reliability and efficiency – and by extension the usefulness of a technology – are thus ultimately judged subjectively by the users of the technology in question.

According to Nordin’s model of technology, digital histopathology and conventional light microscopic histopathology can be thought of as two separate technological paradigms (or parapraxes). They are competing ways of solving the practical problem of making histopathological diagnoses. Proponents of each paradigm see their own technology as useful and use it to solve new problems. In a technologically pluralistic society, the two paradigms are able to coexist and compete with each other, thereby allowing the user to choose the one they find most suitable for the task at hand.

In line with this notion, digital pathology can be seen as a rational technology for cancer histopathology if the persons employing the technology, namely practicing pathologists, choose to use it over light microscopy. As we have already discussed, the shift from light microscopy to digital whole slide images has already begun and is anticipated to continue. The ambition of this shift is to advance the histopathological diagnostics of cancer, thereby improving the care of patients dealing with "the emperor of all maladies" (Mukherjee, 2010).

6.6 Limitations of the study

The current study dealt with different aspects of digital pathology, namely the use of whole slide scanning and viewing, digital image analysis, as well as digital image compression and storage. All of the aforementioned were studied in the context of light microscopic histopathology. Digital pathology as a whole, however,

encompasses a wider range of applications such as cytopathology, fluorescence microscopy, as well as special visualization techniques including spectral imaging (Pantanowitz & Parwani, 2017).

The study also focused on digital pathology from the perspective of high-income countries with good pathology resources, funding, and infrastructure. Yet, from the point of view of developing countries and rural regions, digital pathology holds special promise in alleviating the lack of pathology services (Holmström, 2020).

The individual studies examined applications of digital pathology in what might be considered a proof of concept manner. The sample sizes were limited and applications in routine diagnostics probably call for more validation studies.

6.7 Future prospects

The single hottest topic in digital pathology at the moment, at least from the perspective of cancer diagnostics, is most likely the use of deep learning algorithms in image analysis. At the same time, the use of magnetic resonance imaging can in some instances almost parallel histopathology (Bardis et al., 2020). An exciting prospect would be to combine histopathological and magnetic resonance imaging data as substrate for deep learning applications in otder to help discover new aspects of cancer pathophysiology (Kartasalo et al., 2021).

Another possible implementation of neural networks in cancer diagnostics would be to train them with large prospective cohort data sets in otder to directly produce information about clinical events, such as metastasis or mortality, instead of histopathological entities (Kartasalo et al., 2021).

Digital image analysis could help the pathologist in some of their less challenging diagnostic tasks and free up time for more demanding work thereby mitigating the aforementioned pathologist shortage. Examples of such tasks could be the detection of carcinoma metastasis in lymphadenectomy preparations (Sakamoto et al., 2020), and the preanalyzing of common non-malignant tissues such as ordinary cholecystectomy and appendectomy specimens (Wrenn et al., 2017).

A long standing issue in studying various aspects of histopathology is the variety of practices in the preanalytical phase (Bass et al., 2014). This refers to all of the steps in turning a piece of tissue into a stained histopathological slide ready to be analysed by the pathologist. For instance, factors concerning the sectioning,

fixation, and staining of the tissue need to be standardised in order for the resulting specimens to be genuinely comparable between different laboratories. The same goes for factors concerning whole slide image scanning. By way of harmonizing the input material for large neural networks, or through large scale international quality control systems, digital pathology might help in standardizing the preanalytical phase of histopathology.

7 SUMMARY AND CONCLUSIONS

Digital histopathology presents a new way of carrying out the central task of a pathologist in managing cancer patients, namely making the pathological-anatomical diagnosis of tissue removed through surgery or needle biopsy. Making the shift from a light microscope to a computer environment offers many benefits, some of which have been examined in this dissertation.

The aim of this thesis was to develop applications of digital pathology and test whether these can improve the histopathological diagnosis of prostate and breast cancer. We demonstrated that whole slide images can be used to assess the Gleason grade of a prostate biopsy and that the use of an internet based platform can be beneficial in assessing interobserver variation and teaching and standardizing of Gleason grading.

Besides Gleason grading, another important aspect of prostate histopathology is the interpretation of immunohistochemistry. We created a method of viewing two whole slide images simultaneously and synchronously and tested this method in visualizing the AMACR-p63 double stain along with normal hematoxylin and eosin staining of prostate biopsies. We showed that this technique can be used for histopathology education as well as in clinical diagnostics in selected cases.

A key issue in breast cancer diagnostics is defining the HER2 status of a tumor, that is, whether the tumor expresses the molecule and can then be treated with anti-HER2 drugs. We studied the use of digital image analysis, using both photomicrographs and whole slide images, in aiding the pathologist in their task to define the HER2 status of a breast cancer surgical resection specimen. We showed that using a free and publicly available image analysis software can help to resolve cases otherwise deemed equivocal by manual microscopy.

The introduction of digital histopathology into routine diagnostic work is underway. One technical challenge is managing the large amounts of image data generated by whole slide images. When there is a need to store large numbers of whole slide images, it is essential to strike a balance between image fidelity and file size. To deal with this issue, we studied the optimal imaging resolution of a whole slide scanner using a methodology that can be utilized, for instance, in comparing

whole slide scanners before acquiring one. Lastly, we introduced a novel way of image compression suited for whole slide images in order to reduce the storage footprint, and cost, of whole slide images.

The first two studies in this dissertation represent the very beginnings of whole slide imaging in pathology, and the field has advanced since then, perhaps in small part due to the findings of these studies. Taken together, the findings in this dissertation can hopefully advance the use of digital pathology in cancer diagnostics and thereby improve the care of cancer patients.

8 REFERENCES

Abramowitz, M. (2003). Microscope – Basics and Beyond (Revised ed.). Olympus America Inc.

Abramowitz, M., Keller, H. E., Spring, K. R., Flynn, B. O., Long, J. C., Parry-Hill, M. J., Tchourioukanov, K. I., Davidson, M.W. (2015, November 13) Sixteenth and Seventeenth Century Microscopes. Molecular Expressions Microscopy Primer. Retrieved April 24, 2020, from https://micro.magnet.fsu.edu/primer/anatomy/anatomy.html

Adobe Developers Association. (1992). TIFF™, Revision 6.0, Final – June 3, 1992. Retrieved July 11, 2021, from https://www.adobe.io/content/dam/udp/en/open/standards/tiff/TIFF6.pdf

American Society of Clinical Oncology & College of American Pathologists. (2014, May 23). Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology /College of American Pathologists Clinical Practice Guideline Update Data Supplement. Retrieved October 15, 2020, from https://documents.cap.org/documents/2014-asco-her2-data-supplement.pdf

Association for Pathology Informatics. (n.d.). About API. Retrieved April 27, 2020, from https://www.pathologyinformatics.org/about_api.php

Acharya, T. & Tsai, P-S. (2005). Introduction to Data Compression. In JPEG2000 Standard for Image Compression. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471653748.ch1

Allison, K. H., Hammond, M. E. H., Dowsett, M., McKernin, S. E., Carey, L. A., Fitzgibbons, P. L., Hayes, D. F., Lakhani, S. R., Chavez-MacGregor, M., Perlmutter, J., Perou, C. M., Regan, M. M., Rimm, D. L., Symmans, W. F., Torlakovic, E. E., Varella, L., Viale, G., Weisberg, T. F., McShane, L. M., & Wolff, A. C. (2020). Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 38(12), 1346–1366. https://doi.org/10.1200/JCO.19.02309

Allsbrook, W. C., Mangold, K. A., Johnson, M. H., Lane, R. B., Lane, C. G., & Epstein, J. I. (2001a). Interobserver reproducibility of Gleason grading of prostatic carcinoma: General pathologist. Human Pathology, 32(1), 81–88. https://doi.org/10.1053/hupa.2001.21135

Allsbrook, W. C., Mangold, K. A., Johnson, M. H., Lane, R. B., Lane, C. G., Amin, M. B., Bostwick, D. G., Humphrey, P. A., Jones, E. C., Reuter, V. E., Sakr, W., Sesterhenn, I. A., Troncoso, P., Wheeler, T. M., & Epstein, J. I. (2001b). Interobserver reproducibility of Gleason grading of prostatic carcinoma: Urologic pathologists. Human Pathology, 32(1), 74–80. https://doi.org/10.1053/hupa.2001.21134

Ball, C. S. (1966). The Early History of the Compound Microscope. Bios, 37(2), 51–60. JSTOR.

Bankman, I. N. (Ed.). (2009). Handbook of Medical Image Processing and Analysis (2nd ed.). Academic Press.

Barbareschi, M., Demichelis, F., Forti, S., & Dalla Palma, P. (2000). Digital Pathology: Science Fiction? International Journal of Surgical Pathology, 8(4), 261–263. https://doi.org/10.1177/106689690000800401

Bardis, M. D., Houshyar, R., Chang, P. D., Ushinsky, A., Glavis-Bloom, J., Chahine, C., Bui, T.-L., Rupasinghe, M., Filippi, C. G., & Chow, D. S. (2020). Applications of Artificial Intelligence to Prostate Multiparametric MRI (mpMRI): Current and Emerging Trends. Cancers, 12(5), E1204.

Bass, B. P., Engel, K. B., Greytak, S. R., & Moore, H. M. (2014). A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: How well do you know your FFPE specimen? Archives of Pathology & Laboratory Medicine, 138(11), 1520–1530.

Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology—New tools for diagnosis and precision oncology. Nature Reviews. Clinical Oncology, 16(11), 703–715. https://doi.org/10.1038/s41571-019-0252-y

Block, P. N. (2005). Black box. In The Oxford Companion to Philosophy. Oxford University Press. http://www.oxfordreference.com/view/10.1093/acref/9780199264797.001.0001/acref-9780199264797-e-261

Bostwick, D. G. (1994). Grading prostate cancer. American Journal of Clinical Pathology, 102(4 Suppl 1), S38-56.

Bui, M. M., Riben, M. W., Allison, K. H., Chlipala, E., Colasacco, C., Kahn, A. G., Lacchetti, C., Madabhushi, A., Pantanowitz, L., Salama, M. E., Stewart, R. L., Thomas, N. E., Tomaszewski, J. E., & Hammond, M. E. (2019). Quantitative Image Analysis of Human Epidermal Growth Factor Receptor 2 Immunohistochemistry for Breast Cancer: Guideline From the College of American Pathologists. Archives of Pathology & Laboratory Medicine, 143(10), 1180–1195. https://doi.org/10.5858/arpa.2018-0378-CP

Bulten, W., Pinckaers, H., van Boven, H., Vink, R., de Bel, T., van Ginneken, B., van der Laak, J., Hulsbergen-van de Kaa, C., & Litjens, G. (2020). Automated deep-learning system for Gleason grading of prostate cancer using biopsies: A diagnostic study. The Lancet. Oncology, 21(2), 233–241. https://doi.org/10.1016/S1470-2045(19)30739-9

Cambride in Colour. (n.d.). Color Management & Printing: Understanding the Process. Digital Photography Tutorials. Retrieved May 9, 2020, from https://www.cambridgeincolour.com/tutorials.htm

Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S., & Fuchs, T. J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nature Medicine, 25(8), 1301–1309. https://doi.org/10.1038/s41591-019-0508-1

Camparo, P., Egevad, L., Algaba, F., Berney, D. M., Boccon-Gibod, L., Compérat, E., Evans, A. J., Grobholz, R., Kristiansen, G., Langner, C., Lopez-Beltran, A., Montironi, R., Oliveira, P., Vainer, B., & Varma, M. (2012). Utility of whole slide imaging and virtual microscopy in prostate pathology. APMIS: Acta Pathologica,

Microbiologica, et Immunologica Scandinavica, 120(4), 298–304. https://doi.org/10.1111/j.1600-0463.2011.02872.x

Chlipala, E., Elin, J., Eichhorn, O., Krishnamurti, M., Long, R. E., Sabata, B., Smith, M. (n.d.) Archival and Retrieval in Digital Pathology Systems. Digital Pathology Association. Retrieved May 5, 2020, from https://digitalpathologyassociation.org/_data/files/Archival_and_Retrieval_in_Digital_Pathology_Systems.pdf

Clunie, D. A., Dennison, D. K., Cram, D., Persons, K. R., Bronkalla, M. D., & Primo, H. R. (2016). Technical Challenges of Enterprise Imaging: HIMSS-SIIM Collaborative White Paper. Journal of Digital Imaging, 29(5), 583–614. https://doi.org/10.1007/s10278-016-9899-4

Conway, C. M. & Hewitt, S. M. (2017). Image analysis applications. In L. Pantanowitz & A. V. Parwani (Eds.), Digital Pathology. American Society for Clinical Pathology.

Cullen, D., Wheatley, L., Orchard, G. (2012). Molecular diagnostics in histopathology. In G. Orchard & B. Nation (Eds.), Histopathology. Oxford University Press.

Danielsen, H. E., Kildal, W., & Sudbø, J. (2000). [Digital image analysis in pathology—Exemplified in prostatic cancer]. Tidsskrift for Den Norske Laegeforening: Tidsskrift for Praktisk Medicin, Ny Raekke, 120(4), 479–488.

Danks, D., & London, A. J. (2017). Algorithmic bias in autonomous systems. Proceedings of the 26th International Joint Conference on Artificial Intelligence, 4691–4697.

Delahunt, B., Egevad, L., Samaratunga, H., Srigley, J. R., Cheng, L., Clouston, D., Furusato, B., Kench, J., Leite, K. R. M., MacLennan, G. T., Moch, H., Pan, C.-C., Ro, J., Tsuzuki, T., van der Kwast, T., Wheeler, T., & Yaxley, J. W. (2020). Intraductal carcinoma of the prostate is not a diagnostic entity. Histopathology. https://doi.org/10.1111/his.14260

Dias, A., Claudino, W., Sinha, R., Perez, C. A., & Jain, D. (2016). Human epidermal growth factor antagonists and cardiotoxicity - A short review of the problem and preventative measures. Critical Reviews in Oncology/Hematology, 104, 42–51. https://doi.org/10.1016/j.critrevonc.2016.04.015

DICOM. (2010, August 24). Digital Imaging and Communications in Medicine (DICOM) Supplement 145: Whole Slide Microscopic Image IOD and SOP Classes. Retrieved May 4, 2020, from ftp://medical.nema.org/medical/dicom/final/sup145_ft.pdf

Digital Pathology Association. (n.d.). Glossary of Terms. Retrieved April 27, 2020, from https://digitalpathologyassociation.org/glossary-of-terms_1

Eble, J. N., Sauter, G., Epstein, J. I., Sesterhenn, I. A. (Eds.). (2004). Pathology and Genetics of Tumors of the Urinary System and Male Genital Organs. International Agency for Research on Cancer.

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. Histopathology, 19(5), 403–410. https://doi.org/10.1111/j.1365-2559.1991.tb00229.x

Epstein, J. I. (2000). Gleason score 2-4 adenocarcinoma of the prostate on needle biopsy: A diagnosis that should not be made. The American Journal of Surgical Pathology, 24(4), 477–478. https://doi.org/10.1097/00000478-200004000-00001

Epstein, J. I., Allsbrook, W. C., Amin, M. B., Egevad, L. L., & ISUP Grading Committee. (2005). The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. The American Journal of

Surgical Pathology, 29(9), 1228–1242. https://doi.org/10.1097/01.pas.0000173646.99337.b1

Epstein, J. I. & Lotan, T. L. (2015). The Lower Urinary Tract and Male Genital System. In V. Kumar, A. K. Abbas, J. C. Aster (Eds.), Robbins & Cotran pathologic basis of disease (9th ed.). Retrieved June 6, 2020, from https://ebookcentral.proquest.com

Epstein, J. I., Egevad, L., Amin, M. B., Delahunt, B., Srigley, J. R., Humphrey, P. A., & Grading Committee. (2016). The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. The American Journal of Surgical Pathology, 40(2), 244–252. https://doi.org/10.1097/PAS.0000000000000530

Epstein, J. I., Amin, M. B., Fine, S. W., Algaba, F., Aron, M., Baydar, D. E., Beltran, A. L., Brimo, F., Cheville, J. C., Colecchia, M., Comperat, E., da Cunha, I. W., Delprado, W., DeMarzo, A. M., Giannico, G. A., Gordetsky, J. B., Guo, C. C., Hansel, D. E., Hirsch, M. S., … Trpkov, K. (2020). The 2019 Genitourinary Pathology Society (GUPS) White Paper on Contemporary Grading of Prostate Cancer. Archives of Pathology & Laboratory Medicine. https://doi.org/10.5858/arpa.2020-0015-RA

Feller, J., Fitzgerald, B., Hissam, S. A., Lakhani, K. R. (2005). Introduction. In Perspectives on Free and Open Source Software. The MIT Press; eBook Academic Collection (EBSCOhost).

Fereidouni, F., Harmany, Z. T., Tian, M., Todd, A., Kintner, J. A., McPherson, J. D., Borowsky, A. D., Lechpammer, M., Bishop, J., Demos, S. G., & Levenson, R. (2017). Microscopy with ultraviolet surface excitation for rapid slide-free histology. Nature Biomedical Engineering, 1, 957–966.

Ferreira, R., Moon, B., Humphries, J., Sussman, A., Saltz, J., Miller, R., & Demarzo, A. (1997). The Virtual Microscope. Proceedings: A Conference of the American Medical Informatics Association. AMIA Fall Symposium, 449–453.

Fiete, R. D. (2010a). The Imaging Chain and Applications. In Modeling the Imaging Chain of Digital Cameras. Spie Press.

Fiete, R. D. (2010b). Digital Sensors. In Modeling the Imaging Chain of Digital Cameras. Spie Press.

García-Rojo, M. (2016). International Clinical Guidelines for the Adoption of Digital Pathology: A Review of Technical Aspects. Pathobiology: Journal of Immunopathology, Molecular and Cellular Biology, 83(2–3), 99–109. https://doi.org/10.1159/000441192

Gleason, D. F. (1966). Classification of prostatic carcinomas. Cancer Chemotherapy Reports, 50(3), 125–128.

Gleason, D. F. (1992). Histologic grading of prostate cancer: A perspective. Human Pathology, 23(3), 273–279. https://doi.org/10.1016/0046-8177(92)90108-f

Gress, D. M., Edge, S. B., Greene, F. L., Washington, M. K., Asare, E. A., Brierley, J. D., Byrd, D. R., Compton,C. C., Milburn Jessup, J., Winchester, D. P., Amin, M. B., Gershenwald, J. E. (2017). Principles of Cancer Staging. In M. B. Amin, S. Edge, F. Greene, D. R. Byrd, R. K. Brookland, M. K. Washington, J. E. Gershenwald, C. C. Compton, K. R. Hess, D. C. Sullivan, J. M. Jessup, J. D. Brierley, L. E. Gaspar, R. L. Schilsky, C. M. Balch, D. P. Winchester, E. A. Asare, M. Madera, D. M. Gress, L. R. Meyer. (Eds.), AJCC Cancer Staging Manual (8th ed.). Springer International Publishing.

Griffin, J., & Treanor, D. (2017). Digital pathology in clinical use: Where are we now and what is holding us back? Histopathology, 70(1), 134–145. https://doi.org/10.1111/his.12993

Grimes, G. J., McClellan, S. A., Goldman, J., Vaughn, G. L., Conner, D. A., Kujawski, E., McDonald, J., Winokur, T., & Fleming, W. (1997). Applications of virtual reality technology in pathology. Studies in HealthTechnology and Informatics, 39, 319–327.

Grundium. (n.d.). Grundium Ocus portable whole slide imaging microscope. Retrieved May 9, 2020, from https://www.grundium.com/

Gurcan, M. N. (2017). Image analysis fundamentals. In L. Pantanowitz & A. V. Parwani (Eds.), Digital Pathology. American Society for Clinical Pathology.

Haidekker, M. A. (2010). Advanced biomedical image analysis. Retrieved May 8, 2020, from https://ebookcentral.proquest.com

Halmi, N. S. (1978). Immunostaining of growth hormone and prolactin in paraffin-embedded and stored or previously stained materials. The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society, 26(7), 486–495. https://doi.org/10.1177/26.7.80416

Hajdu, S. I. (2005). Rudolph Virchow, Pathologist, Armed Revolutionist, Politician, and Anthropologist. Annals of Clinical & Laboratory Science, 35(2), 203–205.

Hicks, D. G. (2011). Immunohistochemistry in the diagnostic evaluation of breast lesions. Applied Immunohistochemistry & Molecular Morphology: AIMM, 19(6), 501–505. https://doi.org/10.1097/PAI.0b013e31822c8a48

Holmström, O. (2020). Point-of-care diagnostics with digital microscopy and artificial intelligence. [Doctoral dissertation, University of Helsinki]. Helda.

Holmström, O., Linder, N., Moilanen, H., Suutala, A., Nordling, S., Ståhls, A., Lundin, M., Diwan, V., & Lundin, J. (2019). Detection of breast cancer lymph node metastases in frozen sections with a point-of-care low-cost microscope scanner. PloS One, 14(3), e0208366. https://doi.org/10.1371/journal.pone.0208366

Humphrey, P. A., Amin, M. B., Berney, D. M., Billis, A., Cao, D., Cheng, L., Delahunt, B., Egevad, L., Epstein, J. I., Fine, S. W., Grignon, D. J., Kristiansen, G., Lopez-Beltran, A., Magi-Galluzzi, C., Netto, G. J., Rubin, M. A., Samaratunga, H., Srigley, J. R., True, L. D., Tsuzuki, T., Van der Kwast, T. (2016). Tumors of the prostate: Acinar adenocarcinoma. In H. Moch, P. A. Humphrey, T. M Ulbright, V. E. Reuter (Eds.). WHO Classification of Tumors of the Urinary System and Male Genital Organs. (4th ed.). International Agency for Research on Cancer.

Häger, S. (2016, May 23). How to build a business case to justify the investment in digital pathology. Sectra Medical. Retrieved May 5, 2020, from https://medical.sectra.com/resources/how-to-build-a-business-case-to-justify-the-investment-in-digital-pathology/

Hänninen, T. & Liljeberg, K. (2019). Bioanalyytikon toteuttama dissekointi patologian laboratoriossa. [Bachelor's thesis, Tampere University of Applied Sciences]. Theseus.

International Agency for Research on Cancer. (2018, September 12). Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018. Press release N° 263. Retrieved April 23, 2020, from https://www.who.int/cancer/PRGlobocanFinal.pdf

International Agency for Research on Cancer. (2020). Global Cancer Observatory. Retrieved April 23, 2020, from https://gco.iarc.fr/

International Agency for Research on Cancer. (2021). WHO Classification of Tumors Online. Retrieved May 23, 2021, from https://tumorclassification.iarc.who.int/

ISO. (2000). ISO/IEC 15444-1:2000. Information technology — JPEG 2000 image coding system — Part 1: Core coding system. International Organization for Standardization. Retrieved May 7, 2020, from https://www.iso.org/standard/27687.html

Jardim, M., & Peres, I. (2014). Photographing Microscopic Preparations in the 19th Century: Techniques And Instrumentation. Retrieved April 24, 2020, from https://www.researchgate.net/publication/265292564_Photographing_Microscopic_Preparations_in_the_19th_Century_Techniques_And_Instrumentation

JPEG. (n.d.). The Joint Photographic Experts Group (JPEG). Retrieved May 7, 2020, from https://jpeg.org/

Kartasalo, K., Bulten, W., Delahunt, B., Chen, P.-H. C., Pinckaers, H., Olsson, H., Ji, X., Mulliqi, N., Samaratunga, H., Tsuzuki, T., Lindberg, J., Rantalainen, M., Wählby, C., Litjens, G., Ruusuvuori, P., Egevad, L., & Eklund, M. (2021). Artificial Intelligence for Diagnosis and Gleason Grading of Prostate Cancer in Biopsies-Current Status and Next Steps. European Urology Focus, 7(4), 687–691.

Khalbuss, W. E., Pantanowitz, L., & Parwani, A. V. (2011). Digital Imaging in Cytopathology. Pathology Research International, 2011. https://doi.org/10.4061/2011/264683

Kelty C. (2005). Free Science. In J. Feller, B. Fitzgerald, S. A. Hissam, K. R. Lakhani (Eds.), Perspectives on Free and Open Source Software. The MIT Press; eBook Academic Collection (EBSCOhost).

Kreutzfeldt, J., Rozeboom, B., Dey, N., & De, P. (2020). The trastuzumab era: Current and upcoming targeted HER2+ breast cancer therapies. American Journal of Cancer Research, 10(4), 1045–1067.

Krupinski, E. A., Bhattacharyya, A. K., Weinstein, R. S. (2016). Telepathology and Digital Pathology Research. In K. J. Kaplan & L. K. F. Rao (Eds.), Digital Pathology: Historical Perspectives, Current Concepts & Future Applications. Springer.

Kumar V., Abbas, A. K, Aster, J. C. (2015). Neoplasia. In Robbins & Cotran pathologic basis of disease (9th ed.). Retrieved June 6, 2020, from https://ebookcentral.proquest.com

Kweldam, C. F., van Leenders, G. J., & van der Kwast, T. (2019). Grading of prostate cancer: A work in progress. Histopathology, 74(1), 146–160. https://doi.org/10.1111/his.13767

Lacaze, P., & Lacroix, J. (2014). Information Storage and the State of the Art of Electronic Memories. In Non-volatile memories. Retrieved May 5, 2020, from https://ebookcentral.proquest.com

Leong, F. J. W.-M., Brady, M., & McGee, J. O. (2003). Correction of uneven illumination (vignetting) in digital microscopy images. Journal of Clinical Pathology, 56(8), 619–621. https://doi.org/10.1136/jcp.56.8.619

Lester, S. C. (2015). The Breast. In V. Kumar, A. K. Abbas, J. C. Aster (Eds.), Robbins & Cotran pathologic basis of disease (9th ed.). Retrieved June 6, 2020, from https://ebookcentral.proquest.com

Liu, Y., Kohlberger, T., Norouzi, M., Dahl, G. E., Smith, J. L., Mohtashamian, A., Olson, N., Peng, L. H., Hipp, J. D., & Stumpe, M. C. (2019). Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for

Pathologists. Archives of Pathology & Laboratory Medicine, 143(7), 859–868. https://doi.org/10.5858/arpa.2018-0147-OA

Lucas, R., McMichael, T., Smith, W., Armstrong, B. (2006). Solar Ultraviolet Radiation: Global burden of disease from solar ultraviolet radiation. (Prüss-Üstün, A., Zeeb, H., Mathers, C., Repacholi, M., Eds.) World Health Organization

Lundin, M., Lundin, J., Helin, H., & Isola, J. (2004a). A digital atlas of breast histopathology: An application of web based virtual microscopy. Journal of Clinical Pathology, 57(12), 1288–1291. https://doi.org/10.1136/jcp.2004.018739

Lundin, M., Lundin, J., & Isola, J. (2004b). Virtual microscopy. Journal of Clinical Pathology, 57(12), 1250–1251. https://doi.org/10.1136/jcp.2004.019919

Mahmoud, M. S., & Xia, Y. (2019). Clound Computing. In Networked Control Systems: Cloud Control and Secure Control. Elsevier Science & Technology. Retrieved January 17, 2021, from http://ebookcentral.proquest.com/lib/tampere/detail.action?docID=5693517

Marée, R. (2019). Open Practices and Resources for Collaborative Digital Pathology. Frontiers in Medicine, 6, 255. https://doi.org/10.3389/fmed.2019.00255

Mareschal D. & Althaus N. (2009). Connectionism. In M.D. Binder, N. Hirokawa, U. Windhorst (Eds.), Encyclopedia of Neuroscience. Springer.

Mendelsohn, M. L., Kolman, W. A., Perry, B., & Prewitt, J. M. (1965). Morphological analysis of cells and chromosomes by digital computer. Methods of Information in Medicine, 4(4), 163–167.

Metter, D. M., Colgan, T. J., Leung, S. T., Timmons, C. F., & Park, J. Y. (2019). Trends in the US and Canadian Pathologist Workforces From 2007 to 2017. JAMA Network Open, 2(5), e194337-e194337. https://doi.org/10.1001/jamanetworkopen.2019.4337

Meyer Instruments. (n.d.). uScope Digital Whole Slide Scanners. Retrieved May 9, 2020, from https://www.meyerinst.com/brand/uscope/

Meyers, P. H., Nice, C. M., Becker, H. C., Nettleton, W. J., Sweeney, J. W., & Meckstroth, G. R. (1964). Automated Computer Analysis of Radiographic Images. Radiology, 83, 1029–1034. https://doi.org/10.1148/83.6.1029

Moch, H., Humphrey, P. A., Ulbright, T. M., Reuter, V. E. (Eds.). (2016). WHO Classification of Tumors of the Urinary System and Male Genital Organs. (4th ed.). International Agency for Research on Cancer.

Moinfar, F. (Ed.). (2007). Immunohistochemistry (Selected Topics). In Essentials of Diagnostic Breast Pathology: A Practical Approach. Springer.https://doi.org/10.1007/978-3-540-45120-4_18

Molin, J., Lundström, C., & Fjeld, M. (2015). A comparative study of input devices for digital slide navigation. Journal of Pathology Informatics, 6, 7. https://doi.org/10.4103/2153-3539.151894

Montalto, M. C., McKay, R. R., & Filkins, R. J. (2011). Autofocus methods of whole slide imaging systems and the introduction of a second-generation independent dual sensor scanning method. Journal of Pathology Informatics, 2, 44. https://doi.org/10.4103/2153-3539.86282

Morrison, A. O., & Gardner, J. M. (2015). Microscopic Image Photography Techniques of the Past, Present, and Future. Archives of Pathology & Laboratory Medicine, 139(12), 1558–1564. https://doi.org/10.5858/arpa.2014-0315-RA

Mozilla and individual contributors (2021). Image file type and format guide—Web media technologies | MDN. (N.d.). Retrieved July 11 2021, from https://developer.mozilla.org/en-US/docs/Web/Media/Formats/Image_types

Mukherjee, S. (2010). The emperor of all maladies: A biography of cancer. New York: Scribner.

Murphy, D. B., Davidson M. W. (2013a). Diffraction and Spatial Resolution. In D. B. Murphy & M. Davidson, Fundamentals of Light Microscopy and Electronic Imaging. 2nd ed. New Jersey: John Wiley & Sons.

Murphy, D. B., Davidson M. W. (2013b) Fundamentals of Digital Imaging. In D. B. Murphy & M. Davidson,Fundamentals of Light Microscopy and Electronic Imaging. 2nd ed. New Jersey: John Wiley & Sons.

Muskett, D. (2012). From specimen to slide. In G. Orchard & B. Nation (Eds), Histopathology. Oxford University Press.

Mälkiä, H. (2015, February 11). Pula patologeista vaikea koko Suomessa – pahimmillaan viivästyttää jo leikkauksia. Yle Uutiset. Retrieved April 28, 2020, from https://yle.fi/uutiset/3-7794676

Nagpal, K., Foote, D., Tan, F., Liu, Y., Chen, P.-H. C., Steiner, D. F., Manoj, N., Olson, N., Smith, J. L., Mohtashamian, A., Peterson, B., Amin, M. B., Evans, A. J., Sweet, J. W., Cheung, C., van der Kwast, T., Sangoi, A. R., Zhou, M., Allan, R., … Mermel, C. H. (2020). Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens. JAMA Oncology. https://doi.org/10.1001/jamaoncol.2020.2485

National Cancer Institute. (n.d.). NCI Dictionary of Cancer Terms. Retrieved May 8, 2020, from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/digital-image-analysis

Nature. (n.d.). Milestones timeline: Nature Milestones in Light Microscopy. Retrieved April 24, 2020, fromhttps://www.nature.com/milestones/milelight/timeline.html

Nauhria, S., & Ramdass, P. V. A. K. (2019). Randomized cross-over study and a qualitative analysis comparing virtual microscopy and light microscopy for learning undergraduate histopathology. Indian Journal of Pathology & Microbiology, 62(1), 84–90. https://doi.org/10.4103/IJPM.IJPM_241_18

Netapplications.com. (n.d.). Operating System Market Share. Market Share Statistics for Internet Technologies. Retrieved May 4, 2020, fromhttps://www.netmarketshare.com/operating-system-market-share.aspx

Nordin, I. (1988). Teknologins rationalitet: En teori om teknikens struktur och dynamik. Timbro.

Orchard, G., & Javed, A. (2012). Immunohistochemistry in diagnostic histopathology. In G. Orchard & B. Nation (Eds), Histopathology. Oxford University Press.

Overney, N., & Overney, G. (2011). The History of Photomicrography. Retrieved April 24, 2020, from http://www.microscopy-uk.org.uk/mag/artmar10/history_photomicrography_ed3.pdf

Pantanowitz, L., & Parwani, A. V. (Eds.). (2017). Digital Pathology. American Society for Clinical Pathology.

Pantanowitz, L. (2017). Whole slide imaging. In L. Pantanowitz & A. V. Parwani (Eds.), Digital Pathology. American Society for Clinical Pathology.

Pantanowitz, L., Dickinson, K., Evans, A. J., Hassell, L. A., Henricks, W. H., Lennerz, J. K., Lowe, A., Parwani, A. V., Riben, M., Smith, C. D., Tuthill, J. M., Weinstein, R. S., Wilbur, D. C., Krupinski, E. A., & Bernard, J. (2014). American Telemedicine

Association clinical guidelines for telepathology. Journal of Pathology Informatics, 5(1), 39. https://doi.org/10.4103/2153-3539.143329

Pantanowitz, L., Parwani, A. V., Birdsong, G. G., Kurtycz, D. F. I. (2017). Digital Imaging Process. In L. Pantanowitz & A. V. Parwani (Eds.), Digital Pathology. American Society for Clinical Pathology.

Pearce, J. M. (2020). Economic savings for scientific free and open source technology: A review. HardwareX, 8, e00139. https://doi.org/10.1016/j.ohx.2020.e00139

Peercy, P. S. (2000). The drive to miniaturization. Nature, 406(6799), 1023–1026. https://doi.org/10.1038/35023223

Pernick, Nat. IHC procedure. PathologyOutlines.com website. (2020, July 28). Retrieved October 12, 2020, from https://www.pathologyoutlines.com/topic/methodsihcprocedure.html

Rakha, E. A., Reis-Filho, J. S., Baehner, F., Dabbs, D. J., Decker, T., Eusebi, V., Fox, S. B., Ichihara, S., Jacquemier, J., Lakhani, S. R., Palacios, J., Richardson, A. L., Schnitt, S. J., Schmitt, F. C., Tan, P.-H., Tse, G. M., Badve, S., & Ellis, I. O. (2010). Breast cancer prognostic classification in the molecular era: The role of histological grade. Breast Cancer Research, 12(4), 207. https://doi.org/10.1186/bcr2607

Rakha, E. A., Allison, K. H., Ellis, I. O., Horii, R., Masuda, S., Penault-Llorca, F., Tsuda, H., Vincent-Salomon, A. (2019). Invasive breast carcinoma: General overview. In WHO Classification of Tumors Editorial Board (Eds.). Breast Tumors. (5th ed.). International Agency for Research on Cancer.

Saco, A., Bombi, J. A., Garcia, A., Ramírez, J., & Ordi, J. (2016). Current Status of Whole slide Imaging in Education. Pathobiology, 83(2–3), 79–88. https://doi.org/10.1159/000442391

Sakamoto, T., Furukawa, T., Lami, K., Pham, H. H. N., Uegami, W., Kuroda, K., Kawai, M., Sakanashi, H., Cooper, L. A. D., Bychkov, A., & Fukuoka, J. (2020). A narrative review of digital pathology and artificial intelligence: Focusing on lung cancer. Translational Lung Cancer Research, 9(5). https://doi.org/10.21037/tlcr-20-591

Saltz, J. H. (2000). Digital pathology — The big picture. Human Pathology, 31(7), 779–780.

Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., & Drummen, G. P. C. (2019). Super-resolution microscopy demystified. Nature Cell Biology, 21(1), 72–84. https://doi.org/10.1038/s41556-018-0251-8

Sellaro, T. L., Filkins, R., Hoffman, C., Fine, J. L., Ho, J., Parwani, A. V., Pantanowitz, L., & Montalto, M. (2013). Relationship between magnification and resolution in digital pathology systems. Journal of Pathology Informatics, 4, 21. https://doi.org/10.4103/2153-3539.116866

Shah, R. B., & Zhou, M. (2012). Immunohistochemistry in Prostate Biopsy Evaluation. In R. B. Shah & M. Zhou (Toim.), Prostate Biopsy Interpretation: An Illustrated Guide. Springer. https://doi.org/10.1007/978-3-642-21369-4_4

Sinard, J. H. (2017). Static digital imaging: basics & clinical use. In L. Pantanowitz & A. V. Parwani (Eds.), Digital Pathology. American Society for Clinical Pathology.

Smith, S. C., Gandhi, J. S., Moch, H., Aron, M., Compérat, E., Paner, G. P., McKenney, J. K., & Amin, M. B. (2020). Similarities and Differences in the 2019 ISUP and GUPS Recommendations on Prostate Cancer Grading: A Guide for Practicing Pathologists. Advances in Anatomic Pathology. https://doi.org/10.1097/PAP.0000000000000287

Sopyllo, K., Erickson, A. M., & Mirtti, T. (2021). Grading Evolution and Contemporary Prognostic Biomarkers of Clinically Significant Prostate Cancer. Cancers, 13(4), 628. https://doi.org/10.3390/cancers13040628

Stathonikos, N., Veta, M., Huisman, A., & van Diest, P. J. (2013). Going fully digital: Perspective of a Dutch academic pathology lab. Journal of Pathology Informatics, 4, 15. https://doi.org/10.4103/2153-3539.114206

Stathonikos, N., Nguyen, T. Q., & van Diest, P. J. (2021). Rocky road to digital diagnostics: Implementation issues and exhilarating experiences. Journal of Clinical Pathology, 74(7), 415–420. https://doi.org/10.1136/jclinpath-2020-206715

Steckhan, D., Bergen, T., Wittenberg, T., & Rupp, S. (2008). Efficient large scale image stitching for virtual microscopy. Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2008, 4019–4023. https://doi.org/10.1109/IEMBS.2008.4650091

Ström, P., Kartasalo, K., Olsson, H., Solorzano, L., Delahunt, B., Berney, D. M., Bostwick, D. G., Evans, A. J., Grignon, D. J., Humphrey, P. A., Iczkowski, K. A., Kench, J. G., Kristiansen, G., van der Kwast, T. H., Leite, K. R. M., McKenney, J. K., Oxley, J., Pan, C.-C., Samaratunga, H., … Eklund, M. (2020). Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: A population-based, diagnostic study. The Lancet. Oncology, 21(2), 222–232. https://doi.org/10.1016/S1470-2045(19)30738-7

Sucaet, Y., & Waelput, W. (2014a). Digital Pathology's Past to Present. In Digital Pathology. Springer International Publishing. https://doi.org/10.1007/978-3-319-08780-1

Sucaet, Y., & Waelput, W. (2014b). Hardware and Software. In Digital Pathology. Springer International Publishing. https://doi.org/10.1007/978-3-319-08780-1

Szymiczek, A., Lone, A., & Akbari, M. R. (2020). Molecular intrinsic versus clinical subtyping in breast cancer: A comprehensive review. Clinical Genetics. https://doi.org/10.1111/cge.13900

Tadrous P. J. (2010). On the concept of objectivity in digital image analysis in pathology. Pathology, 42(3), 207–211. https://doi.org/10.3109/00313021003641758

Tannas, L. E., Jr. (2020). Electronic display. AccessScience. Retrieved May 4, 2020, from https://doi-org.libproxy.tuni.fi/10.1036/1097-8542.225300

Taylor, V. J., Barnes, P. J., Godwin, S. C., & Bethune, G. C. (2021). Assessment of HER2 using the 2018 ASCO/CAP guideline update for invasive breast cancer: A critical look at cases classified as HER2 2+ by immunohistochemistry. Virchows Archiv: An International Journal of Pathology. https://doi.org/10.1007/s00428-021-03034-4

Tecuci, G. (2012). Artificial intelligence. Wiley Interdisciplinary Reviews: Computational Statistics, 4(2), 168–180. https://doi.org/10.1002/wics.200

The International Society of Urological Pathology. (n.d.). Prostate cancer grading - Grading of prostatic adenocarcinoma: ISUP grade 1 – 5. Retrieved May 23, 2021, from https://isupweb.org/pib/

The Royal College of Pathologists. (2018). Meeting pathology demand: Histopathology workforce census. Retrieved April 28, 2020, from https://www.rcpath.org/uploads/assets/952a934d-2ec3-48c9-a8e6e00fcdca700f/Meeting-Pathology-Demand-Histopathology-Workforce-Census-2018.pdf

Tizhoosh, H. R., & Pantanowitz, L. (2018). Artificial Intelligence and Digital Pathology: Challenges and Opportunities. Journal of Pathology Informatics, 9, 38. https://doi.org/10.4103/jpi.jpi_53_18

Tolonen, T., Näpänkangas, J., Isola, J. (2021). Virtuaalimikroskopia ja digitaalinen patologia. In M. Mäkinen, O. Carpén, V-M. Kosma, V-P. Lehto, T. Paavonen, F. Stenbäck (Eds.), Patologia (Revised ed.), Duodecim. https://www.oppiportti.fi/op/pat00911/do

Tonttila, P. (2020). Multiparametric magnetic resonance imaging for the detection and characterisation of prostate cancer. [Doctoral dissertation, University of Oulu]. Jultika.

Tuominen, V. J., Ruotoistenmäki, S., Viitanen, A., Jumppanen, M., & 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 Research: BCR, 12(4), R56. https://doi.org/10.1186/bcr2615

Tuominen, V. (2012). Virtual Microscopy: Design and Implementation of Novel Software Applications for Diagnostic Pathology. [Doctoral dissertation, University of Tampere]. Trepo.

Tsang, M. W., & Kovarik, C. L. (2010). Global access to dermatopathology services: Physician survey of availability and needs in sub-Saharan Africa. Journal of the American Academy of Dermatology, 63(2), 346–348. https://doi.org/10.1016/j.jaad.2009.09.038

Tufts Center for the Study of Drug Development. (2019). Cancer drug approvals grew from 4% of U.S. total in the 1980s to 27% in 2010–18. Retrieved April 23, 2020, fromhttps://csdd.tufts.edu/s/summary-septoct-2019.pdf

U.S. Food and Drug Administration. (2016, April 20). Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices – Guidance for Industry and Food and Drug Administration Staff. Retrieved April 30, 2020, from https://www.fda.gov/media/90791/download

U.S. National Library of Medicine. MeSH Browser. (n.d.). Retrieved April 24, 2020, fromhttps://meshb.nlm.nih.gov/record/ui?name=telepathology

Valkonen, M., Isola, J., Ylinen, O., Muhonen, V., Saxlin, A., Tolonen, T., Nykter, M., & Ruusuvuori, P. (2020). Cytokeratin-Supervised Deep Learning for Automatic Recognition of Epithelial Cells in Breast Cancers Stained for ER, PR, and Ki-67. IEEE Transactions on Medical Imaging, 39(2), 534–542. https://doi.org/10.1109/TMI.2019.2933656

van den Tweel, J. G., & Taylor, C. R. (2010). A brief history of pathology. Virchows Archiv, 457(1), 3–10. https://doi.org/10.1007/s00428-010-0934-4

Varga, V. S., Molnár, B., & Virág, T. (2012). Automated high throughput whole slide imaging using area sensors, flash light illumination and solid state light engine. Studies in Health Technology and Informatics, 179, 187–202.

Yagi, Y., & Gilbertson, J. R. (2008). The importance of optical optimization in whole slide imaging (WSI) and digital pathology imaging. Diagnostic Pathology, 3 Suppl 1, S1.https://doi.org/10.1186/1746-1596-3-S1-S1

Walter, C. (2005, August 1). Kryder's Law. Scientific American. https://doi.org/10.1038/scientificamerican0805-32

van Leenders, G. J. L. H., van der Kwast, T. H., Grignon, D. J., Evans, A. J., Kristiansen, G., Kweldam, C. F., Litjens, G., McKenney, J. K., Melamed, J., Mottet, N., Paner, G. P., Samaratunga, H., Schoots, I. G., Simko, J. P., Tsuzuki, T., Varma, M.,

Warren, A. Y., Wheeler, T. M., Williamson, S. R., Iczkowski, K. A. ISUP Grading Workshop Panel Members. (2020). The 2019 International Society of Urological Pathology (ISUP) Consensus Conference on Grading of Prostatic Carcinoma. The American Journal of Surgical Pathology. https://doi.org/10.1097/PAS.0000000000001497

Jahn, S. W., Plass, M., & Moinfar, F. (2020). Digital Pathology: Advantages, Limitations and Emerging Perspectives. Journal of Clinical Medicine, 9(11), E3697. https://doi.org/10.3390/jcm9113697

Weinstein, R. S. (1986). Prospects for telepathology. Human Pathology, 17(5), 433–434. https://doi.org/10.1016/s0046-8177(86)80028-4

Weinstein, R.S., Bloom, K. J., Rozek, L. S. (1987). Telepathology: system design and specifications. SPIE Visual Communications and Image Processing, 845, 404–407.

Weinstein, R. S. (2005). Innovations in medical imaging and virtual microscopy. Human Pathology, 36(4), 317–319. https://doi.org/10.1016/j.humpath.2005.03.007

Weinstein, R. S., Graham, A. R., Lian, F., Braunhut, B. L., Barker, G. R., Krupinski, E. A., & Bhattacharyya, A. K. (2012). Reconciliation of diverse telepathology system designs. Historic issues and implications for emerging markets and new applications. APMIS, 120(4), 256–275. https://doi.org/10.1111/j.1600-0463.2011.02866.x

WHO Classification of Tumors Editorial Board (Eds.). (2019). Breast Tumors. (5th ed.). International Agency for Research on Cancer.

Wilson, M. (2016, December 2). Microscope Resolution: Concepts, Factors and Calculation – Airy Discs, Abbe's Diffraction Limit and the Rayleigh Criterion. Leica Microsystems Learn & Share. Retrieved April 30, 2020, from https://www.leica-microsystems.com/science-lab/microscope-resolution-concepts-factors-and-calculation/

Wolff, A. C., Hammond, M. E. H., Allison, K. H., Harvey, B. E., Mangu, P. B., Bartlett, J. M. S., Bilous, M., Ellis, I. O., Fitzgibbons, P., Hanna, W., Jenkins, R. B., Press, M. F., Spears, P. A., Vance, G. H., Viale, G., McShane, L. M., & Dowsett, M. (2018). Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. Archives of Pathology & Laboratory Medicine, 142(11), 1364–1382. https://doi.org/10.5858/arpa.2018-0902-SA

Wolff, A. C., Hammond, M. E. H., Hicks, D. G., Dowsett, M., McShane, L. M., Allison, K. H., Allred, D. C., Bartlett, J. M. S., Bilous, M., Fitzgibbons, P., Hanna, W., Jenkins, R. B., Mangu, P. B., Paik, S., Perez, E. A., Press, M. F., Spears, P. A., Vance, G. H., Viale, G., Hayes, D. F., American Society of Clinical Oncology, College of American Pathologists. (2013). Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 31(31), 3997–4013. https://doi.org/10.1200/JCO.2013.50.9984

Wolff, A. C., Hammond, M. E. H., Schwartz, J. N., Hagerty, K. L., Allred, D. C., Cote, R. J., Dowsett, M., Fitzgibbons, P. L., Hanna, W. M., Langer, A., McShane, L. M., Paik, S., Pegram, M. D., Perez, E. A., Press, M. F., Rhodes, A., Sturgeon, C., Taube, S. E., Tubbs, R., Vance, G. H., van de Vijver, M., Wheeler, T. M., Hayes, D. F, American Society of Clinical Oncology, College of American Pathologists. (2007). American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast

cancer. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 25(1), 118–145. https://doi.org/10.1200/JCO.2006.09.2775

Wolfram, S. (2002). Processes of Perception and Analysis. In A New Kind of Science. Wolfram Media. Retrieved May 7, 2020, from https://www.wolframscience.com/nks/

World Health Organization. (2018a). Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Disease burden and mortality estimates. Retrieved April 23, 2020, from https://www.who.int/healthinfo/global_burden_disease/estimates/en/

World Health Organization. (2018b, September 12). Cancer. Retrieved April 23, 2020, from https://www.who.int/news-room/fact-sheets/detail/cancer

Wrenn, S. M., Callas, P. W., & Abu-Jaish, W. (2017). Histopathological examination of specimen following cholecystectomy: Are we accepting resect and discard? Surgical Endoscopy, 31(2), 586–593.

PUBLICATIONS

Publication I Web-based virtual microscopy in teaching and standardizing Gleason grading. Helin H, Lundin M, Lundin J, Martikainen P, Tammela T, Helin H, van der Kwast T, Isola J. Hum Pathol. 2005 Apr;36(4):381-6. doi:10.1016/j.humpath.2005.01.020

Publication II Virtual microscopy in prostate histopathology: simultaneous viewing of biopsies stained sequentially with hematoxylin and eosin, and alpha-methylacyl-coenzyme A racemase/p63 immunohistochemistry. Helin HO, Lundin ME, Laakso M, Lundin J, Helin HJ, Isola J. J Urol. 2006 Feb;175(2):495-9. doi:10.1016/S0022-5347(05)00164-3

Publication III Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry. Helin HO, Tuominen VJ, Ylinen O, Helin HJ, Isola J. Virchows Arch. 2015 Oct 22. [Epub ahead of print] doi:10.1007/s00428-015-1868-7

Publication IV Optimized JPEG 2000 Compression for Efficient Storage of Histopathological Whole slide Images.Helin H, Tolonen T, Ylinen O, Tolonen P, Näpänkangas J, Isola J.J Pathol Inform. 2018 May 25;9:20. doi: 10.4103/jpi.jpi_69_17. ECollection 2018.

PUBLICATION I

Web-based virtual microscopy in teaching and standardizing Gleason grading

Henrik Helin, Mikael Lundin, Johan Lundin, Paula Martikainen, Teuvo Tammela, Heikki Helin, Theo van der Kwast, Jorma Isola

Human Pathology (2005) 36, 381-386
https://doi.org/10.1016/j.humpath.2005.01.020

Human Pathology (2005) **36**, 381–386

ELSEVIER

Human
PATHOLOGY

www.elsevier.com/locate/humpath

Web-based virtual microscopy in teaching and standardizing Gleason grading☆

Henrik Helin[a,1], Mikael Lundin MD[d,1], Johan Lundin MD, PhD[d], Paula Martikainen MD, PhD[b], Teuvo Tammela MD, PhD[c], Heikki Helin MD, PhD[b,e], Theo van der Kwast MD, PhD[f], Jorma Isola MD, PhD[a,*]

[a]*Institute of Medical Technology, Tampere University and University Hospital, 33014 Tampere, Finland*
[b]*Department of Pathology, Tampere University Hospital, Tampere 33520, Finland*
[c]*Department of Urology, Tampere University Hospital, Tampere 33520, Finland*
[d]*Biomedical Informatics Group, Department of Oncology, University of Helsinki, Helsinki 00290, Finland*
[e]*Division of Pathology, HUSLAB, Helsinki University Hospital, Helsinki 00029, Finland*
[f]*Department of Pathology, Josephine Nefkens Institute, Erasmus Medical Center, 3000 Rotterdam, The Netherlands*

Received 12 January 2005; accepted 19 January 2005

Keywords:
Gleason grading;
Prostatic neoplasms/pathology;
Virtual microscopy

Summary Gleason grading forms the basis of prognostic and therapeutic assessment in prostatic carcinoma despite its subjective nature and substantial interobserver variation. The accuracy of Gleason grading can be improved by the use of educational tools such as reference images. However, conventional microscopy images are of limited educational value because it is neither possible to view the sample at different magnifications nor to navigate into different areas of the specimen. This limitation can be overcome by the use of virtual microscopy, which allows viewing entire digitized microscope slides. We created an interactive Web site (http://www.webmicroscope.net/gleason) featuring a comprehensive set of prostatic needle biopsies as virtual slides, which can be viewed with a standard Web browser (Internet Explorer or Netscape). To evaluate the validity of Web-based virtual microscopy for Gleason grading, an experienced uropathologist (TK) scored a series of 62 biopsies from the original glass slides and 6 weeks later from virtual slides on the Web site using an ordinary desktop computer. The intraobserver agreement was excellent, with identical Gleason scores found in 48 of the 62 cases ($\kappa = 0.73$). The 14 remaining scores differed only by 1 point on the Gleason scale (2-10). The virtual slides were viewed by 2 other uropathologists (PM and HH), with interobserver κ coefficients ranging from 0.55 to 0.62, which is within the range of previously reported studies using glass slides. The 3 uropathologists' Gleason scores were included as reference scores on the Web site, which now serves as a publicly open platform for self-testing and learning of Gleason grading.

☆ This study was supported by grants from the Scientific foundation of Instrumentarium Inc, Finska Läkaresällskapet, Biomedicum foundation, Medicinska Understödsföreningen Liv och Hälsa, Svenska Kulturfonden, Finnish Cancer Foundation, and Sigfrid Juselius Foundation.

* Corresponding author.
E-mail address: jorma.isola@uta.fi (J. Isola).

[1] These authors contributed equally to this work.

doi:10.1016/j.humpath.2005.01.020

We conclude that Web-based virtual microscopy is a promising new tool for teaching and standardizing Gleason grading.

1. Introduction

The Gleason grading system is the most widely used and accepted grading system for prostatic adenocarcinoma [1-5]. The Gleason score of a prostate needle biopsy can have important clinical implications, in part determining whether a patient is a candidate for radical prostatectomy or other therapies. To ensure objective therapeutic decisions, Gleason grading should be as accurate and uniform as possible. The Gleason grading system is based on evaluating the architectural pattern of the malignant glands observed at a relatively low magnification. The 2 most prevalent patterns are assigned grades from 1 to 5, with 1 being the most differentiated and 5 the least differentiated. The sum of the 2 grades, called the Gleason score, thus obtains values from 2 to 10.

As can be expected from a method based on subjective visual judgment, Gleason scores are subject to a significant degree of interobserver variability. Since the introduction of Gleason grading in 1974 [1], a number of studies on the accuracy and variability of the grading system have been conducted [6-13]. Most of the studies have included no preceding consensus training, and these studies uniformly show that the interobserver variability of Gleason grading is considerable. Prior use of educational tools such as reference images or Web tutorials can clearly improve the grading accuracy and reproducibility [13-16]. However, current image reference systems and Web tutorials are based on conventional microscopy images limited to only one magnification level and a preselected specimen area. The act of changing magnification when using conventional light microscopy has been simulated with zoomable images of tissue microarray spots, but still each sample represents only a small preselected area of a specimen [17,18].

The limitations imposed by images representing small preselected areas can be overcome by the means of virtual microscopy. An entire microscope slide can be digitized at a

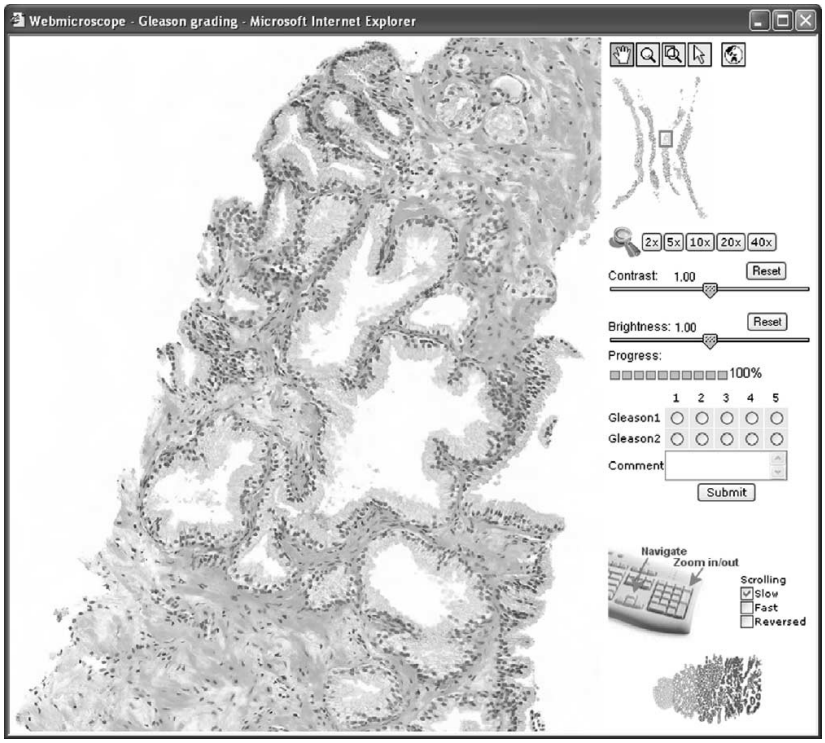


Fig. 1 Screenshot of the virtual slide viewing interface. The small image in the upper right is the overview window, showing the current position and extent of the large window.

high resolution, which allows the user to view any part of the specimen at any magnification. We have recently shown that by using advanced techniques for transferring image data over the internet, Web-based virtual microscopy is a powerful platform for developing various educational applications in histopathology [19]. Here we present an interactive virtual microscopy Web site featuring a comprehensive set of representative, digitized, prostatic needle biopsies. A pathologist can score a set of virtual slides and compare the Gleason scores with those given by 3 expert uropathologists. To validate virtual microscopy in Gleason grading, we studied the concordance of Gleason scores assigned from conventional glass slides to those assigned from virtual slides, as well as the concordance of virtual slide Gleason scores assigned by the 3 experienced uropathologists.

2. Materials and methods

2.1. Patient material

From the hospital records of the Tampere University Hospital, 62 consecutive patients who underwent radical prostatectomy preceded by a needle biopsy were identified. For each patient, a representative, archived, hematoxylin and eosin–stained glass slide needle biopsy was selected to be digitized.

2.2. Virtual slide acquisition and image processing

A Zeiss Axioskop2 MOT microscope (Zeiss Gmbh, Göttingen, Germany) was equipped with a 40× NeoFluar oil objective and Märzhauser motorized specimen stage (Märzhauser, Wetzlar, Germany), which holds 8 standard microscope slides at a time. Images were captured at a 0.26-μm resolution with a CCD camera (Zeiss Axiocam HR; sensor size 6.8 × 7.6 mm; capture resolution 1300 × 1030 pixels). The camera was attached to the microscope with a 0.63× magnifying phototube. Image acquisition was controlled by the KS400 software (ver 3.0, Zeiss) running on a standard MS Windows workstation. A custom macro command script running within the KS400 controlled for sequential autofocus, image acquisition, shading correction, and stage movement consecutively for 8 slides on the specimen stage. The acquired image files were digitally sharpened and stitched into a single montage file, which was compressed into a wavelet-type image file (enhanced compressed wavelet [ECW] format) using the ERMapper software (Earth Resource Mapping Pty, West Perth, Australia) [19].

Table 1 Intraobserver variation of Gleason scores assessed by an experienced uropathologist using glass slides and Web-based virtual microscopy

Gleason scores (virtual microscopy)	Gleason scores (conventional microscopy)			Total
	<7	7	>7	
<7	32	4	0	36
7	1	14	5	20
>7	0	1	5	6
Total	33	19	10	62

Weighted κ = 0.75 (95% CI, 0.63-0.88).

Table 2 Interobserver variation of Gleason scores assessed by 3 experienced uropathologists using Wweb-based virtual microscopy

	Gleason score	<7	7	>7	Interobserver agreement[a]
Uropathologist 1		Uropathologist 2			
	<7	23	5	0	0.66
	7	5	21	1	(0.50-0.82)
	>7	0	3	4	
Uropathologist 3		Uropathologist 2			
	<7	25	11	0	0.59
	7	3	16	1	(0.42-0.77)
	>7	0	2	4	
Uropathologist 3		Uropathologist 1			
	<7	25	10	1	0.61
	7	3	16	1	(0.43-0.78)
	>7	0	1	5	

[a] Values are expressed as weighted κ (95% CI).

2.3. Virtual microscopy image server and network protocol

The compressed virtual slides were uploaded to our Web server running the Image Web Server software (Earth Resource Mapping Pty) [19]. Situated in the National Library of Health Sciences in Helsinki, the image server is connected to the Finnish University and Research Network (www.funet.fi), providing a high-capacity connection to the global internet. Although the virtual slides can be viewed even with a very slow internet connection, we have previously defined 1 to 2 Mbit as a sufficient connection speed for smooth Web viewing [19]. A connection speed test function can be found on the Web site, accompanied by a world map featuring measured connection speeds to the image server from around the world. Virtual slides on the Web site can be viewed within a standard Web browser (MS Internet Explorer current versions of Netscape and Firefox, and Safari on the Macintosh OS X). A small (600 KB) plug-in for the browser is automatically downloaded and installed upon first use. The client computer requirements are modest (300-MHz standard PC, 1-GHz processor speed recommended).

2.4. Experimental design and statistics

The virtual slides were viewed from the Web site with the pathologists' own office computers and screens and with their ordinary internet connection. For both the

intraobserver and the interobserver variability assessments, the primary as well as the secondary Gleason grades were recorded and the slides were classified according to the sum of the grades into less than 7, 7, or more than 7. In the agreement analysis, κ coefficients were calculated with linear weights. The κ coefficients calculated from the raw scores, as well as the coefficients calculated from the categorized scores, were reported. For self-learning purpose, the Web site also contains consensus Gleason grades, defined as the Gleason grade given by 2 or, when available, by all 3 uropathologists. When all 3 uropathologists gave a different primary or secondary grade (in 3 cases), the middle score was taken as consensus.

3. Results

3.1. The interactive Web site

A public Web site was constructed for the purpose of the study (http://www.webmicroscope.net/gleason). The Web site serves as a platform for self-testing and learning of Gleason grading. The user can choose whether to randomly view slides from the pool of 62 slides or from a selected smaller set of biopsies representing an educationally more meaningful distribution of Gleason grades. The user can assign each biopsy a Gleason score and compare it with the scores given by expert uropathologists. Grading can be

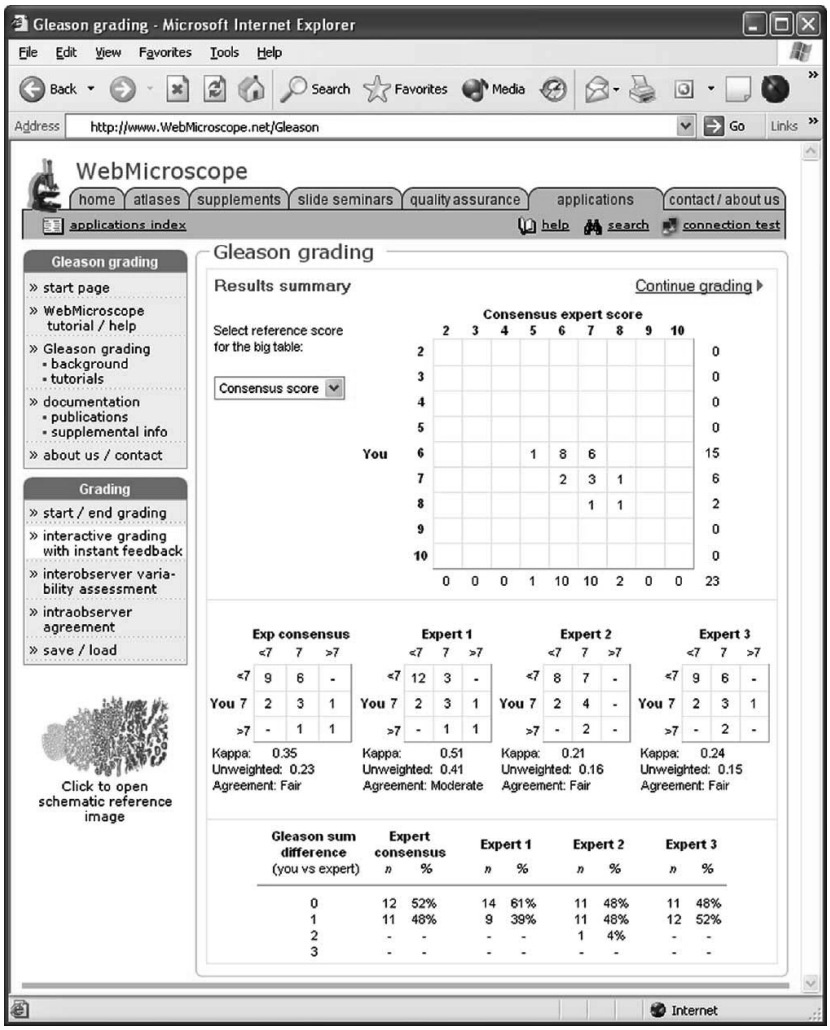


Fig. 2 Screenshot of the summary statistics page of the interactive Web site after a Gleason grading session of 30 prostatic biopsies. The interobserver variation is tabulated against each of the 3 expert uropathologists (HH, PM, and TK) and the consensus Gleason score formed from the experts' scores.

conducted in an interactive learning mode with instant feedback, or the user can choose to assess his or her agreement with the expert uropathologists after finishing a grading session. Special functions on the Web site include an option to save a scoring set to later return and score the same slides again in random order and thus assess one's own intraobserver agreement.

The virtual slides can be viewed within a standard Web browser (Fig. 1). By clicking on a thumbnail image, the corresponding virtual slide opens in a separate browser window. The current interface includes mouse and cursor navigation, stepless zooming with the mouse to any magnification level, incremented keyboard zooming, and fixed microscope-resembling zoom levels. An area of interest can be zoomed in by drawing a box with the mouse. A small overview window shows the current view area and allows fast "dragging" navigation. The contrast and brightness of the virtual slide can be adjusted in the browser.

3.2. Validation of Web-based virtual microscopy for Gleason grading

One experienced uropathologist (TK) viewed the 62 biopsies from the original glass slides with a conventional light microscope and 6 weeks later with his ordinary office computer, screen, and internet connection from the virtual slides on the Web site in a random order and without specimen codes. The intraobserver agreement (glass slides versus virtual slides) was excellent (weighted $\kappa = 0.73$; 95% confidence interval [CI], 0.61-0.85) in that 48 (77%) of the 62 scores were identical and the 14 remaining scores differed only by 1 point on the Gleason scale from 2 to 10. When the data were categorized as in Table 1, weighted κ was 0.75 (95% CI, 0.63-0.88). Recorded from the Web server log, the average time needed to grade a biopsy from a virtual slide was 3 minutes.

3.3. Interobserver variability of Gleason grading with Web-based virtual microscopy

All 62 virtual slides were independently assessed from the Web site by 3 expert uropathologists (HH, PM, and TK) with their own ordinary office computer, screen, and internet connection. No efforts were made to standardize the Gleason scoring for the purpose of this study. The Gleason scores of the 3 pathologists were compared pairwise, that is, pathologist 1 (HH) versus pathologist 2 (PM), pathologist 2 versus pathologist 3 (TK), and pathologist 1 versus pathologist 3 (Table 2). The corresponding weighted κ coefficients were 0.66 (95% CI, 0.50-0.82), 0.59 (95% CI, 0.42-0.77), and 0.61 (95% CI, 0.43-0.78), whereas the coefficients calculated from the uncategorized data were 0.62 (95% CI, 0.47-0.78), 0.58 (95% CI, 0.42-0.75), and 0.55 (95% CI, 0.39-0.70), reflecting moderate to substantial interobserver agreement. A screen shot of the summary statistics page after a virtual microscopy Gleason grading session is shown in Fig. 2.

4. Discussion

Interobserver comparisons of histopathologic diagnoses have traditionally been made by preparing replicate sets of microscope slides and distributing them to the participating pathologists. In the case of small needle biopsies, this is not possible, leaving the use of micrographs as the only option for distributing identical samples to a large number of pathologists. Because of recent advances in slide digitization techniques and within image viewing and internet data transfer technology [20,21], Web-based virtual microscopy has emerged as an important new tool, especially in histopathology education [19]. The main advantage of virtual microscopy as compared with conventional printed or digital photomicrographs is the possibility to view the entire specimens, that is, any part of a whole slide at any magnification. This is particularly important in Gleason grading, where assessing the architectural growth pattern of the malignant glands is decisive for the Gleason grade. The slides are viewed at relatively low magnification, but the evaluation also requires inspection of cellular details at high magnification.

The interactive virtual microscopy Web site presented in this study was evaluated by 3 experienced uropathologists (HH, PM, and TK). With their own ordinary office computer, monitor, and internet connection, all found that the image quality and viewing speed over the internet are sufficient for Gleason grading to be done in the context of teaching and/or interlaboratory quality control. To achieve a satisfactory image quality, the slides were digitized with a 40× oil microscope objective at a resolution of 0.3 μm/pixel. A high digitizing resolution generates huge amounts of image data and requires an effective virtual slide viewing system to achieve a sufficient viewing speed over the internet. The present Gleason grading Web site is based on advanced image server and viewing technology adopted from the aerial and satellite imaging industry [19]. The average time one uropathologist (TK) spent to assess the Gleason score for a biopsy was 3 minutes, reflecting a high image viewing speed over the internet. In practice, the image refresh delay is negligible and does not cause any noticeable inconvenience in viewing, provided that the internet connection speed is adequate. We have previously defined 1 to 2 Mbit as an adequate connection speed [19]. This can be accomplished with a fast home internet connection (xDSL), and with a 2-Mbit or faster connection, virtual microscopy is very smooth. At the time of writing, satisfactory connection speeds have been verified from around Europe and North America.

The accuracy of assessing Gleason grading from virtual slides has to the best of our knowledge not previously been reported. To evaluate the validity of Web-based virtual microscopy for Gleason grading, we conducted an experiment where an expert uropathologist (TK) evaluated the same samples twice. He first viewed the original glass slides, and after a 6-week break, he scored the corresponding virtual slides directly from the Web site. In

only 14 of the 62 biopsies did the score differ by 1 point, the rest of the scores being identical. The results indicate that Gleason grading conducted with Web-based virtual microscopy corresponds well to scoring from glass slides. Furthermore, none of the 3 expert pathologists felt that the image quality of the virtual slides had any significant impact on the accuracy of Gleason grading. As expected, the interobserver variability of the 3 expert uropathologists was within the range of previously reported studies using glass slides and conventional microscopy [6-13]. As no effort was made to standardize the pathologists' Gleason grades beforehand, the results emphasize that the concordance and accuracy of Gleason grading is not perfect even among expert uropathologists and that new tools for standardization are needed at all levels.

We have previously shown that Web-based virtual microscopy is useful in the education of histopathologic entities by publishing a public virtual slide atlas of breast histopathology (www.webmicroscope.net/breastatlas) [19]. The present study demonstrates that Web-based virtual microscopy is well suited for interactive education, that is, for self-testing and learning Gleason grading of prostate carcinoma. We conclude that Web-based virtual microscopy is an effective method to evaluate interobserver variability and that it is a promising new tool for teaching and standardizing Gleason grading.

References

[1] Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol 1974;111:58-64.

[2] Epstein JI. Pathology of adenocarcinoma of the prostate. In: Walsh PC, Retik AB, Vaughan ED, Wein AJ, editors. Campbell's Urology. 7th ed. Pennsylvania: W.B. Saunders Company; 1998. p. 2497-505.

[3] DeMarzo AM, Nelson WG, Isaacs WB, Epstein JI. Pathological and molecular aspects of prostate cancer. Lancet 2003;361:955-64.

[4] Murphy GP, Busch C, Abrahamsson PA, et al. Histopathology of localized prostate cancer. Consensus conference on diagnosis and prognostic parameters in localized prostate cancer. Stockholm, Sweden, May 12-13, 1993. Scand J Urol Nephrol Suppl 1994;162:7-42.

[5] Bostwick DG. Grading prostate cancer. Am J Clin Pathol 1994;102: 38-56.

[6] Allsbrook Jr WC, Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: General pathologists. Hum Pathol 2001;32:81-8.

[7] Allsbrook Jr WC, Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: Urologic pathologists. Hum Pathol 2001;32:74-80.

[8] de las Morenas A, Siroky MB, Merriam J, et al. Prostatic adenocarcinoma: reproducibility and correlation with clinical stages of four grading systems. Hum Pathol 1988;19:595-7.

[9] Svanholm H, Mygind H. Prostatic carcinoma. Reproducibility of histologic grading. Acta Pathol Microbiol Immunol Scand 1985;93: 67-71.

[10] di Loreto C, Fitzpatrick B, Underhill S, et al. Correlation between visual clues, objective histologic features, and interobserver agreement in prostate cancer. Am J Clin Pathol 1991;96:70-5.

[11] Lessells AM, Burnett RA, Howatson SR, et al. Observer variability in the histopathological reporting of needle biopsy specimens or the prostate. Hum Pathol 1997;28:646-9.

[12] Glaessgen A, Hamberg H, Pihl CG, et al. Interobserver reproducibility of modified Gleason score in radical prostatectomy specimens. Virchows Arch 2004;445:17-21.

[13] Mikami Y, Manabe T, Epstein JI, et al. Accuracy of gleason grading by practicing pathologists and the impact of education on improving agreement. Hum Pathol 2003;34:658-65.

[14] Egevad L. Reproducibility of Gleason grading of prostate cancer can be improved by the use of reference images. Urology 2001;57:291-5.

[15] Kronz JD, Silberman MA, Allsbrook Jr WC, et al. A Web-based tutorial improves practicing pathologists' Gleason grading of images of prostate carcinoma specimens obtained by needle biopsy. Cancer 2000;89:1818-23.

[16] Kronz JD, Silberman MA, Allsbrook Jr WC, et al. Pathology residents' use of a Web-based tutorial to improve Gleason grading of prostate carcinoma on needle biopsies. Hum Pathol 2000;31: 1044-50.

[17] Bova GS, Parmigiani G, Epstein JI, et al. Web-based tissue microarray image data analysis: initial validation testing through prostate cancer Gleason grading. Hum Pathol 2001;32:417-27.

[18] de la Taille A, Viellefond A, Berger N, et al. Evaluation of the interobserver reproducibility of Gleason grading of prostatic adenocarcinoma using tissue microarrays. Hum Pathol 2003;34: 444-9.

[19] Lundin M, Lundin J, Helin H, Isola J. A digital atlas of breast histopathology: an application of Web-based virtual microscopy. J Clin Pathol 2004;57:1288-91.

[20] Leong FJ, McGee JO. Automated complete slide digitization: a medium for simultaneous viewing by multiple pathologists. J Pathol 2001;195:508-14.

[21] Glatz-Krieger K, Glatz D, Mihatsch MJ. Virtual slides: high-quality demand, physical limitations, and affordability. Hum Pathol 2003; 34:968-74.

PUBLICATION II

Virtual Microscopy in Prostate Histopathology: Simultaneous Viewing of Biopsies Stained Sequentially With Hematoxylin and Eosin, and α-Methylacyl-Coenzyme A Racemase/p63 Immunohistochemistry

Henrik O. Helin, Mikael E. Lundin, Mervi Laakso, Johan Lundin, Heikki J. Helin and Jorma Isola

Journal of Urology (2006) 175(2), 495-499
https://doi.org/10.1016/s0022-5347(05)00164-3

Virtual Microscopy in Prostate Histopathology: Simultaneous Viewing of Biopsies Stained Sequentially With Hematoxylin and Eosin, and α-Methylacyl-Coenzyme A Racemase/p63 Immunohistochemistry

Henrik O. Helin, Mikael E. Lundin, Mervi Laakso, Johan Lundin, Heikki J. Helin and Jorma Isola*

From the Institute of Medical Technology (HOH, JI), University of Tampere, Tampere, Biomedical Informatics Group (MEL, JL), Department of Oncology, University of Helsinki and Division of Pathology (HJH), HUSLAB, Helsinki University Central Hospital, Helsinki and Department of Pathology (ML), Seinäjoki Central Hospital, Seinäjoki, Finland

Purpose: Histopathological diagnosis of small focus carcinomas in prostatic needle biopsies is often assisted by IHC. To make a definitive diagnosis the pathologist must compare IHC findings with hematoxylin and eosin stained tissue morphology. We introduce what is to our knowledge a new application of virtual microscopy, in which hematoxylin and eosin, and IHC stains done sequentially on the same microscope slide can be simultaneously displayed and compared on a computer screen.

Materials and Methods: A total of 30 hematoxylin and eosin stained prostatic needle biopsies were scanned with a computer controlled microscope. The slides were destained and then immunostained with a cocktail of AMACR and p63 antibodies, which labels the nuclei of nonmalignant basal cells (p63) and the cytoplasm of neoplastic glandular cells suspicious for malignancy (AMACR). The slides were then scanned again and the pairs of virtual slides were aligned for synchronized viewing.

Results: The presented technique was found helpful when suspicious lesions were small and when examining the immunoprofile of specimens was warranted, in addition to examining hematoxylin and eosin stained tissue morphology. The usefulness of our approach based on virtual microscopy can be evaluated on the website http://www.webmicroscope.net/AMACRp63, which also serves as an educational tool for self-learning the correlation between hematoxylin and eosin stained tissue morphology, and AMACR/p63 IHC in prostate biopsies.

Conclusions: The technology for simultaneously viewing sequentially hematoxylin and eosin and IHC stained prostate biopsies can be readily used for educational purposes, as exemplified by our website, and along with the availability of rapid virtual slide scanners it can also be used for clinical diagnostics.

Key Words: prostate, prostatic neoplasms, biopsy, alpha-methylacyl-CoA racemase, telepathology

Histopathological diagnosis of prostatic needle biopsies can be difficult due to benign mimics of cancer, eg post-atrophic hyperplasia, mucinous metaplasia and atypical adenomatous hyperplasia, PIN and borderline changes, often termed ASAP.[1–3] To better discriminate these lesions from small focus carcinomas pathologists often use IHC, in addition to standard hematoxylin and eosin staining. A central histopathological criterion is basal cells, which distinguishes benign prostate glands and PIN from cancerous glands. Basal cells can be difficult to identify with a standard morphological stain but with IHC using antibodies against high molecular weight cytokeratins[4-6] and the transcription factor p63[7–10] their presence or absence can be ascertained. The usefulness of basal cell IHC as an aid in the interpretation of prostate biopsies is well established.[4–10]

Since the transcription factor p63 is a nuclear protein, it can be used simultaneously with the cytoplasmic marker AMACR (clone P504S, Dako Cytomation, Copenhagen, Denmark), which was recently discovered to be up-regulated in prostate cancer.[11] An increasing number of studies have described the usefulness of AMACR IHC for the differential diagnostics of prostate carcinoma.[12] More recently IHC using a cocktail of AMACR and p63 antibodies has pinpointed more specifically the nature of suspicious glands in prostatic biopsies.[13,14] Positive cytoplasmic AMACR reactivity occurring without nuclear p63, ie glands lacking basal cells, gives the most comprehensive immunophenotypic evidence of carcinoma.[13,14] PIN is characterized by AMACR positivity in the presence of p63 positive basal cells.[13,14]

Although the usefulness of AMACR and basal cell markers has been firmly documented, the concordance between IHC and a diagnosis based on hematoxylin and eosin staining is not always perfect.[1,2] Therefore, the histopathological diagnosis of prostate biopsies cannot be based on IHC alone. Hematoxylin and eosin staining allows the examination of

Submitted for publication February 17, 2005.

Supported by grants from the Scientific Foundation of Instrumentarium, Inc., Finska Läkaresällskapet, Biomedicum Foundation, Medicinska Understödsföreningen Liv och Hälsa, Svenska Kulturfonden, Finnish Cancer Foundation and Sigfrid Juselius Foundation.

* Correspondence: Institute of Medical Technology, 33014 University of Tampere, Tampere, Finland (telephone: +358-3-2156729; FAX: +358-3-2158923; e-mail: jorma.isola@uta.fi).

cellular morphology, especially for nucleoli, which are a central cellular criterion for carcinoma. Thus, it is well established that hematoxylin and eosin, and IHC complement each other and must be evaluated in parallel. Unfortunately suspicious lesions in prostatic needle biopsies are often small and not always present on deeper sections cut from the paraffin block for IHC.[15] A solution to the problem is to routinely prepare intervening unstained slides from prostate biopsy blocks. The extra slides are stored and used for IHC should the morphological stain (hematoxylin and eosin) not provide a definitive diagnosis.[16] Another approach is to destain a previously stained hematoxylin and eosin slide, followed by IHC.[15,17] Due to destaining hematoxylin and eosin morphology is no longer available for comparison with IHC unless a digital imaging system is used.[15]

We describe an application of virtual microscopy,[18–20] which allows simultaneous on-screen visualization of a hematoxylin and eosin and AMACR/p63 double IHC staining done sequentially on the same slide. The entire biopsies can be viewed at any magnification in a transparent overlay mode in a single viewing window or side by side in 2 synchronized viewing windows. This method has several practical advantages, namely in histopathology education, in second opinion and quality control situations, and in clinical diagnostics when the morphology and immunophenotype of small lesions must be compared.

MATERIALS AND METHODS

For the study we selected 23 cases with formalin fixed, paraffin embedded prostatic needle biopsies available in the archive at the Pathology Department at Central Hospital Seinäjoki, Finland. Specimens were selected that favored diagnostically challenging diagnoses, such as small focus carcinoma, PIN, ASAP, proliferative inflammatory atrophy or suspicion for carcinoma without a more detailed description. All diagnoses were based on the original hematoxylin and eosin staining only. Common diagnostic entities were also included.

Ordinary 3 to 4 μm tissue sections were cut from the paraffin blocks on charged SuperFrost™ Plus slides to avoid detaching tissue sections from the slides. The slides were stained routinely with hematoxylin and eosin, and digitized using a virtual microscopy slide scanning system, as described. After slide scanning the coverslips were removed by soaking the slides in xylene until the coverslips detached. The slides were then washed with absolute ethanol, air dried and immersed in antigen retrieval buffer, composed of 0.5M tris and 1 mM ethylenediaminetetraacetic acid, pH 9. Antigen retrieval was done in an autoclave at 105C for 5 minutes, followed by a 20-minute cooling period. This procedure also removes hematoxylin and eosin staining completely. After rinsing and endogenous peroxidase quenching primary antibody incubation was done using a cocktail of antibodies to p63 (clone 4A4+Y4A3, Novocastra, Newcastle, United Kingdom), dilution 1:200, and to AMACR (clone P504S), dilution 1:200, for 30 minutes at room temperature. Antibodies were detected using a PowerVision+™ reagent kit according to manufacturer instructions. The slides were immersed for 10 minutes in diaminobenzidine and enhanced with 0.5% copper sulfate for 5 minutes. The slides were counterstained in hematoxylin, dehydrated in graded ethanols, cleared in xylene and coverslipped. The now IHC stained slides were then digitized a second time. Figure 1 shows the slide staining and scanning protocol.

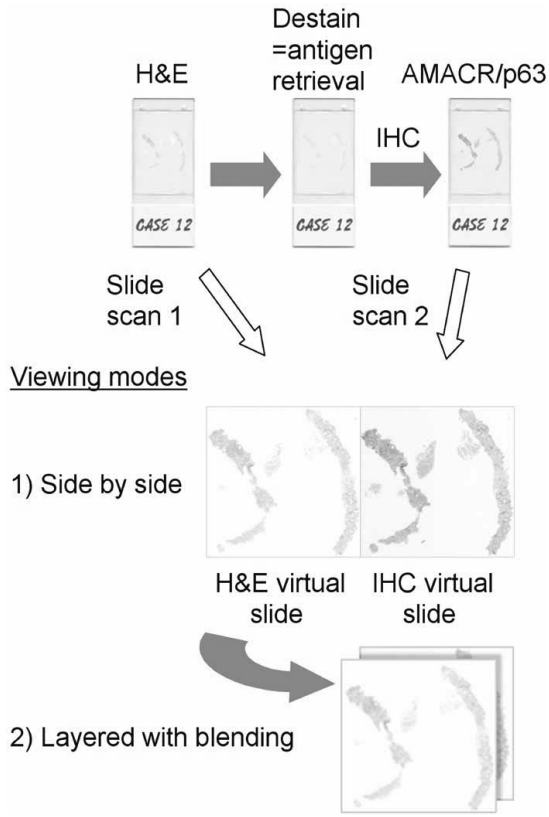


FIG. 1. Preparation and viewing modes of virtual slide pair. H & E stained glass slide is scanned using computer controlled microscope. Same slide is destained as result of antigen retrieval procedure and immunostained with cocktail of AMACR and p63 antibodies. IHC stained slide is scanned again, and H & E and IHC virtual slides are processed for synchronized viewing layered with blending on top of each other or side by side.

Virtual microscopy. An Axioskop2 MOT microscope (Zeiss GmBH, Göttingen, Germany) was equipped with a NeoFluar® oil 40× objective (numerical aperture 1.3) and motorized specimen stage. A contiguous array of digital images covering the entire biopsy was captured at 0.3 μm per pixel resolution with a color sensor camera (capture resolution 1,300 × 1,030 pixels in 3 color scanning mode). The image capture process, ie stage movement, autofocus, shading correction and image capture, was automated using KS400 software, version 3.0 (Zeiss GmBH). The acquired image files were digitally sharpened and stitched into a single montage file, which was compressed into a wavelet-type image file (enhanced compressed wavelet format) using ER Mapper software (Earth Resource Mapping Pty, West Perth, Australia). The compressed virtual slides were uploaded to a web server running Image Web Server software (Earth Resource Mapping Pty, West Perth, Australia). The virtual slide pairs (hematoxylin and eosin, and IHC) were

FIG. 2. Screen shot of 2 slide virtual microscopy interface demonstrates prostatic biopsy stained with H & E (left) and AMACR/p63 immunostaining (right). Zooming (magnification change) and navigation in sample are synchronized, ie every movement occurs in H & E and IHC viewing windows simultaneously and similarly. AMACR positive and p63 negative glands suspicious for carcinoma. Inset (upper right) shows overview image of current position and extent of large window. Circle indicates ROI.

exactly aligned for synchronized viewing by keeping the hematoxylin and eosin virtual slide as a reference and adjusting the position of the corresponding IHC virtual slide (http://www.webmicroscope.net/AMACRp63/methods).

Virtual slides on the website can be viewed in a standard web browser (Microsoft® Internet Explorer or Mozilla® Firefox™) on any Windows® platform. A small (600 kb) plug-in for the browser is automatically downloaded and installed at the first viewing session. Client computer requirements are modest, that is a 1 GHz standard personal computer is recommended.

RESULTS

Routinely hematoxylin and eosin stained slides were first digitized using an established virtual microscopy scanning technique. The coverslips were removed and immunohistochemical staining for AMACR and p63 was performed using a standard protocol. The antigen retrieval procedure removed the hematoxylin and eosin staining completely and the quality of immunostaining on destained slides was similar to that on ordinary unstained material. Removing the coverslip and performing AMACR/P63 immunostaining requires approximately 3 hours, followed by slide scanning, which requires 0.5 to 1 hour.

Figure 1 shows the slide staining and scanning protocol. A total of 30 biopsies from 23 patients were scanned and a publicly open website containing the virtual slides was created (http://www.webmicroscope.net/AMACRp63). The virtual slides representing pairs of hematoxylin and eosin, and IHC stains can be viewed within a standard web browser (Internet Explorer or Mozilla® Firefox™). Figure 2 shows a screen shot.

By clicking on a biopsy thumbnail image on the website a new browser window is opened and the corresponding IHC, and hematoxylin and eosin virtual slides are loaded into a single viewing window but in separate layers. The viewing interface includes basic virtual microscopy functions, such as mouse and cursor navigation, and zooming to any magnification level. A small overview window is always present in the margin, showing the current view area and allowing rapid dragging navigation. The contrast and brightness of the virtual slide can be adjusted in the browser. Initially only hematoxylin and eosin staining is visible in the viewing window but hematoxylin and eosin, and IHC image blending can be controlled using a mouse operated slider bar. Almost perfect alignment of hematoxylin and eosin, and IHC staining was achieved, allowing comparison of hematoxylin and eosin morphology and immunophenotype even at a single cell level. As an alternative to blending hematoxylin and eosin, and IHC images, the user can choose to view hematoxylin and eosin, and IHC virtual slides side by side in 2 viewing windows (fig. 2). Also, in this mode zooming (magnification change) and navigation within the sample are synchronized, ie every movement takes place in the hematoxylin and eosin, and IHC viewing windows simultaneously and similarly.

The website not only demonstrates the simultaneous viewing of sequentially stained virtual slides, but also serves as an educational tool for self-learning the correlation between prostate histopathology and the interpretation of AMACR/p63 immunostaining. To improve the pedagogic value an experienced pathologist (HJH) defined and annotated ROIs in slides with foci of particular diagnostic interest. ROIs are indicated on the screen by circles. They represent carcinomatous and suspicious glands, PIN, ASAP and post-atrophic hyperplasia. ROI indicators and the annotations are initially hidden but they can be turned on by the user. By clicking on an annotation hyperlink in the right margin the corresponding ROI is automatically zoomed into the viewing window and a textual description is displayed. A general comment on each slide is also available but not the definitive diagnoses of the case, which would require inspection of all biopsies. Depending on observer experience with prostate pathology the nature of most of these lesions can be diagnosed or suspected on hematoxylin and eosin staining. However, the aid of parallel AMACR/p63 immunostaining is already apparent after a short time of use.

DISCUSSION

Immunostains for high molecular weight cytokeratin, p63 and AMACR are frequently used to aid in the interpretation of prostatic biopsies containing glands suspicious for malignancy. Despite the obvious advantages of IHC it is not 100% sensitive or specific to provide a definitive diagnosis of prostatic biopsies alone.[1,2] Unfortunately biopsies often contain only a few suspicious glands, which may have been cut through and, therefore, are no longer present on deeper sections cut from the paraffin block for IHC.[15] For this reason obtaining immunohistochemical information on such a lesion can be difficult. Our approach is based on virtual microscopy. An entire hematoxylin and eosin stained biopsy is digitized, destained and then after immunostaining the same slide is scanned a second time. Using virtual microscopy the suspicious glands identified by hematoxylin and eosin staining, and their AMACR/p63 immunophenotype can be compared on a computer screen more easily than with any other approach. To our knowledge this is the first demonstration of the use of whole slide virtual microscopy for real-time comparison of 2 stains made sequentially on the same slide.

Feedback from clinical pathologists who tested the simultaneous hematoxylin and eosin, and IHC viewing was exclusively positive. If an analogous comparison is done using stored intervening unstained sections for IHC,[16] the hematoxylin and eosin, and IHC slides must be changed back and forth in the microscope, and relocating the cells or gland of interest can be difficult and time-consuming. Another approach is destaining the hematoxylin and eosin stained slide to be used for IHC.[15] This procedure is similar to ours but since the hematoxylin and eosin staining is removed, a real-time comparison of tissue morphology with IHC is not possible. A modification of this protocol, termed tissue protection immunohistochemistry, was recently introduced.[17] In this method all except 1 tissue section on the slide are covered with a liquid cover glass embedding medium, which protects the sections from destaining. The uncovered section is immunostained and can be compared with the adjacent hematoxylin and eosin stained sections on the same slide.[17] Although the procedure has proved usable, it has the disadvantage that most immunostaining robots (autostainers) cannot perform IHC on slides covered partially with a staining protectant. Furthermore, only adjacent sections and not exactly the same section on the slide can be compared. In our virtual microscopy based system the laboratory protocols need not be altered from those used routinely for IHC of unstained slides and the entire biopsy is available for real-time comparison. Since the alignment of the hematoxylin and eosin, and IHC staining is done by adjusting the on-screen positioning of the virtual slides, the presented system for synchronized viewing should be applicable to any slide scanner with a compatible output image format. Along with the availability of rapid slide scanners hematoxylin and eosin stained slides can routinely be scanned within minutes, followed by IHC staining, which typically requires 2 to 4 hours, a second virtual slide scanning, requiring minutes, and processing of the virtual slide pair for on-screen viewing on the Internet, requiring 0.5 to 1 hour. The pathologist can then view the hematoxylin and eosin, and IHC virtual slide pair on the office computer screen via the Intranet (a local computer network).

Overall with recent advances in microscope slide scanning techniques and Internet based image transfer technology virtual microscopy has emerged as an important new tool, especially for histopathology education and quality control.[18] The main advantage of virtual microscopy compared to conventional printed or digital photomicrographs is the possibility of viewing entire specimens, ie any part of a whole slide at any magnification. This is particularly important for the evaluation of prostatic biopsies, when assessing the architectural growth pattern of the glands is decisive for the diagnosis. Slides are first inspected at relatively low magnification but the diagnosis requires the evaluation of cellular details at high magnification. We have previously applied virtual microscopy to assess observer variation in Gleason grading.[19] An excellent agreement between Gleason grades assessed from conventional glass slides and grades assessed from virtual slides indicated that virtual slides of prostatic biopsies can be used without compromising grading accuracy.[19] The web based viewing technology allows the distribution of specimens to an unlimited number of pathologists for educational and quality control purposes as well as for a second opinion in problem cases. The application presented adds simultaneous viewing of 2 slides as a new functionality that cannot be done with a conventional microscope.

CONCLUSIONS

The presented technique for simultaneous viewing of hematoxylin and eosin staining, and immunostaining in prostate needle biopsies can readily be applied in pathology education, and in second opinion and quality control situations. With the availability of new, more rapid slide scanners the technique could also be used in clinical diagnostics. The usefulness of the virtual microscopy based approach can be evaluated on our publicly open website (http://www.webmicroscope.net/AMACRp63). In addition to prostate biopsy pathology, the technology described can be applied in any field of histopathology in which a comparison of tissue morphology and immunophenotype at high resolution is of value.

Abbreviations and Acronyms

AMACR	=	α-methylacyl-coenzyme A racemase
ASAP	=	atypical small acinar proliferation
IHC	=	immunohistochemistry
PIN	=	prostatic intraepithelial neoplasia
ROI	=	region of interest

REFERENCES

1. Epstein, J. I., Algaba, F., Allsbrook, W. C., Jr., Bastacky, S., Boccon-Gibond, L., De Marzo, A. M. et al: Acinar adenocarcinoma. In: WHO Classification of Tumours: Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs. Edited by J. N. Eble, G. Sauter, J. I. Epstein and I. A. Sesterhenn. Lyon: IARC Press, 2004
2. Epstein, J. I.: Diagnosis and reporting of limited adenocarcinoma of the prostate on needle biopsy. Mod Pathol, **17:** 307, 2004
3. Van der Kwast, T. H., Lopes, C., Martikainen, P. M., Pihl, C. G., Santonja, C., Neetens, I. et al: Report of the Pathology Committee: false-positive and false-negative diagnoses of prostate cancer. BJU Int, suppl., **92:** 62, 2003
4. Hedrick, L. and Epstein, J. I.: Use of keratin 903 as an adjunct in the diagnosis of prostate carcinoma. Am J Surg Pathol, **13:** 389, 1989
5. O'Malley, F. P., Grignon, D. J. and Shum, D. T.: Usefulness of immunoperoxidase staining with high-molecular-weight cytokeratin in the differential diagnosis of small-acinar lesions of the prostate gland. Virchows Arch A Pathol Anat Histopathol, **417:** 191, 1990
6. Shin, M., Fujita, M. Q., Yasunaga, Y., Miki, T., Okuyama, A. and Aozasa, K.: Utility of immunohistochemical detection of high molecular weight cytokeratin for differential diagnosis of proliferative conditions of the prostate. Int J Urol, **5:** 237, 1998
7. Signoretti, S., Waltregny, D., Dilks, J., Isaac, B., Lin, D., Garraway, L. et al: p63 is a prostate basal cell marker and is required for prostate development. Am J Pathol, **157:** 1769, 2000
8. Parsons, J. K., Gage, W. R., Nelson, W. G. and De Marzo, A. M.: p63 protein expression is rare in prostate adenocarcinoma: implications for cancer diagnosis and carcinogenesis. Urology, **58:** 619, 2001

9. Shah, R. B., Zhou, M., LeBlanc, M., Snyder, M. and Rubin, M. A.: Comparison of the basal cell-specific markers, 34betaE12 and p63, in the diagnosis of prostate cancer. Am J Surg Pathol, **26:** 1161, 2002
10. Weinstein, M. H., Signoretti, S. and Loda, M.: Diagnostic utility of immunohistochemical staining for p63, a sensitive marker of prostatic basal cells. Mod Pathol, **15:** 1302, 2002
11. Xu, J., Stolk, J. A., Zhang, X., Silva, S. J., Houghton, R. L., Matsumura, M. et al: Identification of differentially expressed genes in human prostate cancer using subtraction and microarray. Cancer Res, **60:** 1677, 2000
12. Evans, A. J.: Alpha-methylacyl CoA racemase (P504S): overview and potential uses in diagnostic pathology as applied to prostate needle biopsies. J Clin Pathol, **56:** 892, 2003
13. Tacha, D. E. and Miller, R. T.: Use of p63/P504S monoclonal antibody cocktail in immunohistochemical staining of prostate tissue. Appl Immunohistochem Mol Morphol, **12:** 75, 2004
14. Molinie, V., Fromont, G., Sibony, M., Vieillefond, A., Vassiliu, V., Cochand-Priollet, B. et al: Diagnostic utility of a p63/alpha-methyl-CoA-racemase (p504s) cocktail in atypical foci in the prostate. Mod Pathol, **17:** 1180, 2004
15. Dardik, M. and Epstein, J. I.: Efficacy of restaining prostate needle biopsies with high-molecular weight cytokeratin. Hum Pathol, **31:** 1155, 2000
16. Green, R. and Epstein, J. I.: Use of intervening unstained slides for immunohistochemical stains for high molecular weight cytokeratin on prostate needle biopsies. Am J Surg Pathol, **23:** 567, 1999
17. Kubier, P. and Miller, R. T.: Tissue protection immunohistochemistry: a useful adjunct in the interpretation of prostate biopsy specimens and other selected cases in which immunostains are needed on minute lesions. Am J Clin Pathol, **117:** 194, 2002
18. Lundin, M., Lundin, J. and Isola, J.: Virtual microscopy. J Clin Pathol, **57:** 1250, 2004
19. Helin, H., Lundin, M., Lundin, J., Martikainen, P., Tammela, T., Helin, H. et al: Web-based virtual microscopy in teaching and standardizing Gleason grading. Hum Pathol, **36:** 381, 2005
20. Lundin, M., Lundin, J., Helin, H. and Isola, J. A.: digital atlas of breast histopathology: an application of web-based virtual microscopy. J Clin Pathol, **57:** 1288, 2004

PUBLICATION III

Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry

Henrik O. Helin, Vilppu J. Tuominen, Onni Ylinen, Heikki J. Helin, Jorma Isola

Virchows Archiv (2016) 468(2), 191-198
https://doi.org/10.1007/s00428-015-1868-7

Virchows Arch (2016) 468:191–198
DOI 10.1007/s00428-015-1868-7

ORIGINAL ARTICLE

Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry

Henrik O. Helin[1] · Vilppu J. Tuominen[1] · Onni Ylinen[1] · Heikki J. Helin[2] · Jorma Isola[1]

Received: 17 July 2015 / Revised: 22 September 2015 / Accepted: 12 October 2015 / Published online: 22 October 2015

Abstract Evaluation of human epidermal growth factor receptor 2 (HER2) immunohistochemistry (IHC) is subject to inter-observer variation and lack of reproducibility. Digital image analysis (DIA) has been shown to improve the consistency and accuracy of the evaluation and its use is encouraged in current testing guidelines. We studied whether digital image analysis using a free software application (ImmunoMembrane) can assist in interpreting HER2 IHC in equivocal 2+ cases. We also compared digital photomicrographs with whole-slide images (WSI) as material for ImmunoMembrane DIA. We stained 750 surgical resection specimens of invasive breast cancers immunohistochemically for HER2 and analysed staining with ImmunoMembrane. The ImmunoMembrane DIA scores were compared with the originally responsible pathologists' visual scores, a researcher's visual scores and in situ hybridisation (ISH) results. The originally responsible pathologists reported 9.1 % positive 3+ IHC scores, for the researcher this was 8.4 % and for ImmunoMembrane 9.5 %. Equivocal 2+ scores were 34 % for the pathologists, 43.7 % for the researcher and 10.1 % for ImmunoMembrane. Negative 0/1+ scores were 57.6 % for the pathologists, 46.8 % for the researcher and 80.8 % for ImmunoMembrane. There were six false positive cases, which were classified as 3+ by ImmunoMembrane and negative by ISH. Six cases were false negative defined as 0/1+ by IHC and positive by ISH. ImmunoMembrane DIA using digital photomicrographs and WSI showed almost perfect agreement. In conclusion, digital image analysis by ImmunoMembrane can help to resolve a majority of equivocal 2+ cases in HER2 IHC, which reduces the need for ISH testing.

Keywords Computer-assisted image analysis · Digital pathology · Whole-slide imaging · HER2 · Immunohistochemistry · Breast cancer

Introduction

The human epidermal growth factor receptor 2 (HER2, ERBB2) oncogene protein is overexpressed in approximately 15 % of primary breast cancers [1–4]. The HER2 status of a tumour provides both prognostic and predictive information and is required for patients to qualify for chemotherapy with anti-HER2 drugs such as trastuzumab, lapatinib and pertuzumab [5, 6]. HER2 testing is standard of care in the histopathological diagnosis of breast and gastric cancers and is done mostly by immunohistochemistry (IHC) and in situ hybridisation (ISH) [7, 8].

The optimal strategy for HER2 testing has been under debate for over a decade and there is no consensus as to which testing algorithm is the "gold standard". ISH (both fluorescent and bright-field, to detect HER2 gene amplification) is generally considered accurate and reliable although high reagent costs and labour-intensiveness limit its use in most laboratories to a secondary test to confirm equivocal immunohistochemistry results [7, 9].

HER2 immunohistochemistry is straightforward and can be performed in all modern diagnostic pathology laboratories. Several reagent kits, such as HercepTest™ (Dako, Denmark), PATHWAY™ (Ventana Medical Systems) and Oracle™ (Leica Biosystems) approved for clinical use by the U.S. Food and Drug Administration (FDA), have been analytically

✉ Jorma Isola
jorma.isola@uta.fi

[1] BioMediTech/Cancer Biology, University of Tampere, 33014 Tampere, Finland

[2] HUSLAB, Division of Pathology and Genetics, Helsinki University Central Hospital, P.O. Box 400, 00029 HUS, Finland

validated and can be used in existing automated immunostaining devices. However, evaluation of the HER2-stained slides is subject to substantial interobserver variation and lack of reproducibility [7]. In order to improve the accuracy of HER2 testing, the American Society of Clinical Oncology and the College of American Pathologists published [7] and subsequently updated [10] guideline recommendations for determining HER2 status of breast cancer (ASCO/CAP guidelines). According to current recommendations, equivocal results by IHC (staining of 2+ in the four-tier scale of 0 to 3+) require additional confirmation by a validated assay for HER2 gene amplification [10].

Some studies report false scores in up to 14 % of cases [11, 12]. When viewing hundreds of slides (e.g. for a scientific study), most pathologists find it relatively easy to maintain a reproducible visual scale for their 0/1+, 2+ and 3+ scores, whereas in a diagnostic pathology setting, many general pathologists review only a small number of breast cancer cases per week. A pathologist might not see a true 2+ case in weeks to months, making it difficult to score borderline cases reproducibly. While an incorrect HER2 score must be avoided, most pathologists stay on the safe side, give a 2+ score and submit the case for ISH testing. This reduces the cost- and time-effectiveness of the two-step testing algorithm and some authors argue for using FISH rather than IHC as the primary assay for HER2 testing [13].

Digital image analysis (DIA) has been shown to improve the consistency and accuracy of HER2 evaluation by IHC [14, 15] and its use is encouraged in the ASCO/CAP guidelines for cases with 1–2+ IHC staining [10]. We have previously introduced the ImmunoMembrane software for digital image analysis of HER2 IHC using photomicrographs [16]. ImmunoMembrane (http://jvsmicroscope.uta.fi/software) is both a free web application and an open-source plug-in for the public domain image analysis tool ImageJ [17]. Furthermore, ImmunoMembrane has been introduced recently in a web-based whole-slide image viewing system (JVSwebserver, Jilab Inc., Tampere, Finland) which allows its use directly on whole-slide images (WSI). ImmunoMembrane analyses the completeness and intensity of the cell membrane staining reaction, based on the IHC interpretation criteria of the ASCO/CAP guidelines [7, 10]. This approach of analysing the anatomical localisation of the staining reaction rather than just the immunoreaction intensity has been shown to correlate well with gene amplification in HER2 diagnostics [18]. Similar image analysis applications of HER2 have been described by both academic research groups [15, 19–22] and commercial vendors (reviewed elsewhere [23, 24]).

We conducted the current study to test whether ImmunoMembrane can assist in interpreting the HER2 status of cases visually classified as equivocal (2+) by IHC. We also investigated the agreement between digital photomicrographs and WSI as material for ImmunoMembrane DIA.

Materials and methods

A database search was conducted at the Department of Pathology, Helsinki University Central Hospital, Finland, in order to identify invasive breast cancer cases tested for HER2 from the period of 1 January 2010 to 1 July 2011, resulting in 1249 cases. The slides for the corresponding surgical resection specimens were retrieved from the archives, and starting from the earliest case, 750 consecutive cases (one slide per case) were included in the study. The 750 cases represented the period of 1 January 2010 through 23 May 2011, in which period 1186 breast cancer cases were histologically diagnosed. The missing 436 cases were not found in the archives at the time of retrieval and were distributed along the whole period of time taken for the study. The largest number of consecutive cases missing was 28, representing a period of 14 days. The specimens had been routinely fixed for a period of 24–48 h in neutral formalin and embedded in paraffin.

Immunohistochemistry had been performed according to the manufacturer's instructions using the BenchMark XT automated staining system (Ventana Medical Systems, Tucson, AZ) with PATHWAY anti-HER2/neu (4B5) rabbit monoclonal antibody at a dilution of 6 μg/ml and Ventana ultraView Universal DAB Detection Kit (both Ventana Medical Systems). The slides were counterstained using Ventana Hematoxylin II (Ventana Medical Systems) as part of the automated staining procedure.

The HER2 status of the specimens was established by immunohistochemical staining, and positive (3+) and equivocal (2+) cases were further subjected to in situ hybridization to classify them into positive and negative with regard to HER2 gene amplification. A total of 30 pathologists including residents and specialists scored the cases. Six of the specialists were experienced breast pathologists and the residents routinely consulted senior pathologists before giving their scores.

In situ hybridization had been carried out in the BenchMark XT automated staining system using the INFORM HER2 DNA and the INFORM Chromosome 17 probes and the ultraVIEW SISH Detection Kit (all Ventana Medical Systems) according to the manufacturer's instructions. Consecutive sections of the specimens were hybridised with the probes for HER2 and chromosome 17, respectively, and the specimens' HER2 gene status was classified as amplified if the ratio of HER2 to chromosome 17 was over 2.2. Both the ISH and IHC assays have been subjected to external quality assessment by the Nordic Immunohistochemical Quality Control [25].

For this study, the immunohistochemically stained slides were digitally photomicrographed for image analysis and visually scored as positive (3+), equivocal (2+) or negative (0/1+) by the researcher (HOH) without knowledge of the original pathologist's score. Before the photomicrography and visual scoring, the researcher received training in HER2

IHC evaluation from an experienced breast pathologist with whom he also underwent interobserver testing on 35 consecutive cases from the study material. The cases were scored visually as positive (3+), equivocal (2+) or negative (0/1+), without knowledge of the original scores. This yielded a substantial interobserver agreement with a weighted kappa value of 0.71 (95 % confidence interval 0.49–0.94) calculated using linear weighting.

Digital photography was carried out with a standard Olympus microscope equipped with an Mshot MC-30 3.3-megapixel CCD microscope camera (Micro-shot Technology, Guangzhou, China) using a 1X phototube. A minimum of four non-overlapping jpg images with a resolution of 2048 × 1536 pixels was captured per slide using a ×10 objective lens. The photographed areas were chosen to represent the different HER2 staining patterns of the cancer. In cases where the cancer was small and four non-overlapping images could not be captured, the maximum number of non-overlapping images was captured instead. In cases with a large cancer and a wide range of staining patterns, more than four images were captured. A blank field image and an image of the on-slide positive control tissue (3+) were captured for each photography session.

The images were analysed with the ImmunoMembrane software (http://jvsmicroscope.uta.fi/immunomembrane/) using the advanced mode. This permits the user to define custom cutoff values for the IM score (0–20 points: 0–10 points for membrane completeness and 0–10 points for membrane intensity) the software produces and uses to classify staining into 0/1+, 2+ and 3+. We chose cutoff values of 4 and 10 points (i.e. 0–3 points being classified as 0/1+, 4–9 points as 2+ and 10–20 points as 3+) based on our previous experience and empirical testing. The blank field image was used to correct for the microscope illumination and colour balance whereas the positive control image was used for reference contrast and intensity. A stage micrometre was used to calculate the image scale in pixels per micrometre.

The statistical analyses were performed using MedCalc for Windows, version 15.6.1 (MedCalc Software, Ostend, Belgium).

Comparison of photomicrographs and WSI as material for digital image analysis

The records of a HER2 diagnostics reference laboratory (Jilab Inc., Tampere, Finland) were searched retrospectively from May 2015 so as to identify invasive breast cancer cases immunohistochemically stained for HER2. Ten consecutive cases each of IHC staining patterns, 0/1+, 2+ and 3+ according to the original pathologist's report, were identified. The 30 cases, one glass slide per case, were scanned as whole-slide images using the Objective Imaging Surveyor with Turboscan (Objective Imaging Ltd., Cambridge, UK), using a ×20 Plan Apo microscope objective (scanning resolution 0.23 μm per pixel). Whole-slide images were stored as JPEG 2000 images.

The researcher analysed the virtual slides using ImmunoMembrane directly in the whole-slide viewing interface. This was done by defining a polygonal region of interest (ROI) which was analysed by ImmunoMembrane included in the WSI viewer software. A minimum of four areas were analysed per WSI. The ImmunoMembrane DIA results were compared to the original pathologists' results which had been obtained with the help of ImmunoMembrane using digital photomicrographs. In all cases, the HER2 IHC score in the original pathology report matched the ImmunoMembrane score obtained using photomicrographs.

Results

In the original database search result of 1249 cases, 157 cases (12.6 %) were positive by ISH and of the 750 case subset in our current study 74 (9.9 %). In situ hybridisation had been performed in 314 (41.9 %) cases (originally 2+ and 3+ by IHC). The frequencies for a positive (3+) IHC staining for the original pathologist, the researcher and ImmunoMembrane were 8.4 % (63 cases), 9.5 % (71 cases) and 9.1 % (68 cases), respectively. Equivocal (2+) staining was reported by the pathologist, researcher and ImmunoMembrane in 34 % (255 cases), 43.7 % (328 cases) and 10.1 % (76 cases), respectively. Negative staining (0/1+) showed, for the pathologist, researcher and ImmunoMembrane, frequencies of 57.6 % (432 cases), 46.8 % (351 cases) and 80.8 % (606 cases), respectively. These data are summarised in Table 1.

Of the 750 cases, 6 (0.8 %) were false positives in ImmunoMembrane DIA, defined as cases 3+ positive by IHC but negative by ISH. The corresponding number of false positives was for the researcher 8 (1.1 %) and for the pathologist zero. Similarly, there were six (0.8 %) false negative cases by ImmunoMembrane DIA (0/1+ negative by IHC while positive by ISH). The number of false negatives was zero for the researcher whereas it could not be defined for the pathologist because the laboratory performed ISH only on cases originally scored as 2+ or 3+ by IHC. A cross tabulation comparing the IHC scores of the pathologist, researcher and ImmunoMembrane DIA with in situ hybridisation is presented in Table 2. Examples of images from the ImmunoMembrane analyses are shown in Fig. 1.

Analysis of outliers

We categorised the six false positive and six false negative cases (each representing 0.8 % of the total 750 cases) in ImmunoMembrane DIA into four categories: truly discrepant IHC and ISH, borderline IM-score, heterogenous staining, and low-contrast staining. Of the six

Table 1 Frequencies of the HER2 IHC scores in percent, number of cases in parentheses; total number of cases 750

	Pathologist visual score	Researcher visual score	ImmunoMembrane DIA score
Negative (0/1+)	57.6 (432)	46.8 (351)	80.8 (606)
Equivocal (2+)	34 (255)	43.7 (328)	10.1 (76)
Positive (3+)	8.4 (63)	9.5 (71)	9.1 (68)
Total	100 (750)	100 (750)	100 (750)

DIA digital image analysis

false positive cases, only two exhibited true discrepancy between IHC and ISH after review by an experienced breast pathologist, who scored the cases as 3+ positive in line with ImmunoMembrane while ISH was negative. In both cases, the original pathologist had scored the cases 2+ and the researcher 3+.

Two false positive cases showed a borderline IM score of 10 (in the scale of 0–20) which according to our cutoff value qualifies them as positive, although they were negative by ISH. In both cases, both the original pathologist and the researcher had scored the cases 2+.

Two false positive cases were found to exhibit heterogenous staining when reviewed by an experienced breast pathologist. In these cases, some of the photomicrographs were classified by ImmunoMembrane in agreement with ISH whereas a greater number of images showed a negative score, which rendered the overall score in disagreement with ISH. One of the cases was scored 2+ by the original pathologist and 3+ by the researcher whereas the other case was scored 2+ by both.

Of the six false negative cases, one was found to be discrepant between IHC and ISH after review by an experienced breast pathologist (0/1+ IHC-negative in both the pathologist and ImmunoMembrane DIA but positive by ISH). The case was scored 2+ by both the original pathologist and the researcher. Heterogenous staining was observed in three false negative cases. All three cases were scored 2+ by both the original pathologist and the researcher. Two false negative cases exhibited low-contrast staining when reviewed by an experienced breast pathologist. Both cases were scored 2+ by the original pathologist and the researcher.

ImmunoMembrane analysis using digital photomicrographs and whole-slide images

Table 3 presents a cross tabulation of HER2 IHC scores rendered by a pathologist with the help of ImmunoMembrane using digital photomicrographs and by the researcher using ImmunoMembrane directly on scanned whole-slide images. Of the total 30 cases, two were discordant, the first having been scored 2+ using ImmunoMembrane with photomicrographs and 3+ using WSI and the second vice versa (3+ using WSI and 2+ using photomicrographs). Interobserver variability, calculated using linearly weighted kappa statistic, was almost perfect at 0.92 (95 % confidence interval 0.82–1.0). Figure 2 shows ImmunoMembrane being used directly in a WSI viewer.

Discussion

In this study we, show that ImmunoMembrane [16], a free and publicly available cross-platform compatible ImageJ plug-in and web application for digital image analysis of HER2 IHC (http://jvsmicroscope.uta.fi/immunomembrane/), can assist in interpreting the status of cases visually classified as equivocal (2+). In our material of 750 surgical resection specimens, the proportion of immunohistochemical cases classified as equivocal by the original pathologist was 34 % (255 cases), which is on the upper end of the range reported in the literature. In a systematic review from 2007, the mean frequency of an equivocal IHC score in 17 studies was 23.3 % (with a range of 2 to 87.5 %) [26]. In another systematic review from 2009, the mean frequency of an equivocal IHC score in 10 studies was 18.9 % with a range of 7.9 to 53.2 % [27].

Table 2 Cross tabulation comparing HER2 ISH (presence/absence of amplification) with the IHC score given by the pathologist, researcher and ImmunoMembrane digital image analysis (DIA)

	Pathologist visual IHC			Researcher visual IHC			ImmunoMembrane DIA IHC			Total
	0/1+	2+	3+	0/1+	2+	3+	0/1+	2+	3+	
No ISH	427	5	4	322	110	4	424	7	5	436
ISH−	5	235	0	29	203	8	176	58	6	240
ISH+	0	15	59	0	15	59	6	11	57	74
Total	432	255	63	351	328	71	606	76	68	750

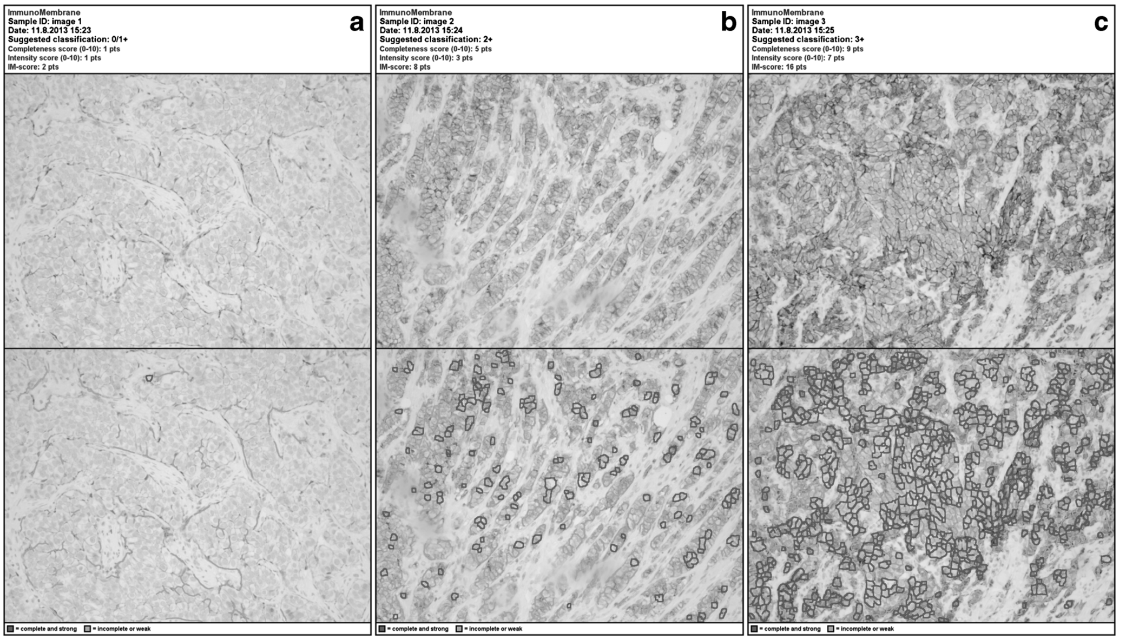


Fig. 1 **a**–**c** Three result images from ImmunoMembrane digital image analysis using photomicrographs. All three cases were originally scored as equivocal (2+) by a pathologist. The first image (**a**) receives an IM score of 2 points and is thus classified as negative (0/1+), the second image (**b**) 8 points (classified as equivocal, 2+) and the third image (**c**) 16 points (classified as positive, 3+). The first two cases were classified visually as 2+ by the researcher and the last one 3+. The first two cases are negative by ISH and the last one positive

The proportion of equivocal results of 34 % (255) by the original pathologist was reduced to 10.1 % (76 cases) by ImmunoMembrane DIA. Of the 255 cases originally classified as equivocal, ImmunoMembrane was able to resolve 190 (74.5 %), 10 of which emerged as positive (3+) and 180 as negative (0/1+). In the ImmunoMembrane analysis, 65 cases remained equivocal, while ImmunoMembrane scored five and six cases as equivocal, originally scored as 3+ and 0/1+, respectively. These data are in line with previous findings of image analysis reducing the proportion of equivocal scores in HER2 immunohistochemistry [15, 20, 28, 29] although there have been opposite findings also, with image analysis increasing the proportion of equivocal results [30, 31]. Images of the result of ImmunoMembrane DIA, demonstrating its discriminative power in cases originally classified as equivocal, are presented in Fig. 1.

Table 3 Cross tabulation comparing ImmunoMembrane digital image analysis (DIA) used with photomicrographs and with whole-slide images

ImmunoMembrane DIA with photomicrographs	ImmunoMembrane DIA with whole-slide images			
	0/1+	2+	3+	Total
0/1+	10	0	0	10
2+	0	9	1	10
3+	0	1	9	10
Total	10	10	10	30

In the two-step testing algorithm proposed in the ASCO/CAP guidelines, in which an equivocal result by IHC requires additional confirmation by a validated assay for HER2 gene amplification, a reduction of the proportion of equivocal cases both increases the information value of IHC and improves the cost- and time-effectiveness of the testing.

Because of the substantial clinical, economic and safety implications of anti-HER2 therapy, assessment of the HER2 status of a tumour must be accurate and reproducible [6, 7]. In our study, we used ISH as the reference method and defined six false positive (3+ IHC-positive and ISH-negative) and six false negative cases (0/1+ IHC-negative and ISH-positive) in the ImmunoMembrane analysis. ISH data was available for 41.9 % of the cases (314/750) because of the laboratory practice of subjecting only immunohistochemically equivocal (2+) and positive (3+) cases for ISH. The frequency of both false positive and false negative cases was thus 1.9 % (6/314) of the cases with ISH data.

After carrying out the present study the laboratory reflex tested for a period of time (8 months, 1 October 2014 through 31 May 2015) all IHC cases (including IHC 0/1+) by ISH (P.

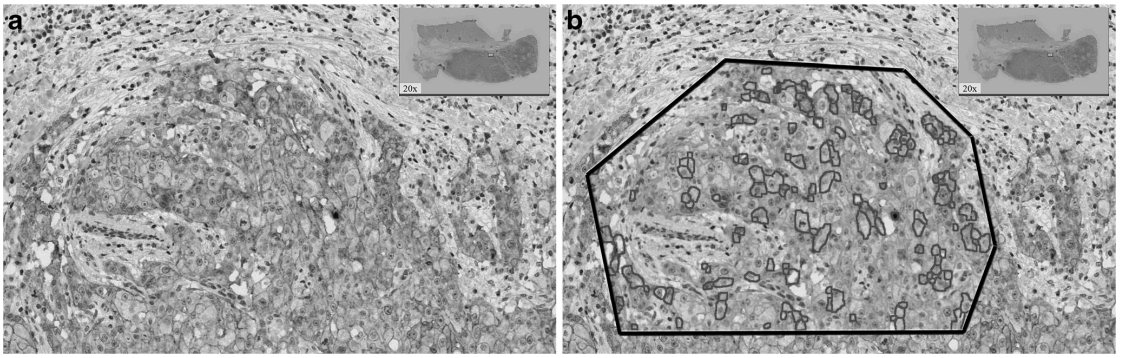


Fig. 2 **a**–**b** Two screenshots demonstrating the process of using ImmunoMembrane digital image analysis directly while viewing a whole-slide image (**a**). The region of interest is demarcated with a *polygon drawing tool*. After the analysis is complete, a pseudo-coloured result image is shown on top of the whole-slide image indicating complete and strong membrane staining in *red* and incomplete or weak staining in *green* (**b**). The depicted tissue exhibits positive 3+ staining according to ImmunoMembrane image analysis

Heikkilä, personal communication, 11 September 2015). In this set of 750 cases, 400 (53.3 %) were scored IHC 0/1+. Only two cases were found positive by ISH. Such a low false negative rate suggests that the proportion of false negatives in our study would not have been substantially affected if the ISH testing had included also the IHC-negative 0/1+ cases. We analysed one of the two false negative cases by ImmunoMembrane DIA and the result was 0/1+, in concordance with the original pathologist score.

The false positive cases by ImmunoMembrane DIA would not have affected anti-HER2 treatment eligibility of the patients because all IHC 2–3+ cases were subjected for ISH. The false negative cases would have made six patients ineligible for anti-HER2 treatment. The false negative rate of 1.9 % is nevertheless lower than the corresponding rate of 6 % reported for an FDA-approved system [29]. It is also lower than the pooled false negative rate of 11 % recently reported for visual scoring, using approved and validated in vitro diagnostic tests in a Nordic IHC quality control programme [32]. However, given the far-reaching clinical and economic consequences of inaccurate HER2 testing [32], even low false negative rates should not be neglected. In the current study, ImmunoMembrane was operated by a researcher; however, as stated in the ASCO/CAP guidelines, in clinical practice, a pathologist must confirm the image analysis result [10].

Discrepancy between IHC analysed by ImmunoMembrane and ISH can be due to several factors. In addition to inaccuracy in the ImmunoMembrane analysis, the analysed images might not be representative of the whole tumour, ISH might not be 100 % accurate either, and regarding false negative cases, some immunohistochemically negative tumours might exhibit HER2 amplification, as is well documented [26, 27, 33]. A borderline score obtained with ImmunoMembrane DIA should be interpreted with care by a pathologist. Heterogenous staining has recently been shown to play a role in discordances between pathologists in reading of HER2 IHC and can also affect the result of ImmunoMembrane image analysis [34]. We propose to analyse a large enough number of image fields to be sure to representatively cover all staining patterns of the specimen in cases with heterogenous staining. This approach is in line with what is recommended in the literature for ISH [35]. Low-contrast staining which can be due to the tissue section properties can influence ImmunoMembrane DIA even with on-slide control tissues.

Our material consisted of a sample of surgical resection specimens from a large university hospital pathology department, reflecting the daily work of diagnostic pathologists. We analysed whole sections instead of tissue microarrays or core needle biopsies in order to mimic, as closely as possible, the daily diagnostic setting. The rate of ISH-positive cases was somewhat low in our material, which can be explained by sampling error due to cases missing from the archives at the time of retrieval. The rate of ISH-positive cases in the original database search is in line with recent findings in the literature [1–4].

As described earlier [16], ImmunoMembrane is a software application for semi-quantitative classification of HER2 IHC, designed as a diagnostic aid for the trained pathologist. ImmunoMembrane does not offer true quantitation of the IHC stain (or ultimately of the amount of HER2 protein), which, in the case of diaminobenzidine (DAB)-based detection systems widely used in HER2 IHC assays, may prove difficult due to the physical properties of the chromogen [36, 37]. ImmunoMembrane performs object-based image analysis [38] by separating the objects of interest (segmenting DAB-stained cell membranes) from the background and analysing them with regard to completeness and intensity of the staining reaction. ImmunoMembrane thus mimics the visual interpretation of HER2 IHC as defined in clinical guidelines and was initially pre-calibrated to match the visual scoring of an expert pathologist. While the image analysis algorithm itself is automated, the selection of regions of interest (ROI) has to be done visually, preferably by a pathologist.

In conclusion, the accuracy of HER2 IHC can be improved in equivocal cases by the use of digital image analysis carried out by the ImmunoMembrane software. In our material, almost three out of four cases originally classified as equivocal by a pathologist were further classified into either negative or positive without loss of accuracy. ImmunoMembrane is easy to apply in clinical practice by virtue of its usability with various combinations of imaging equipment (microscope and camera) and the possibility to use it over the internet (without software download or installation). The fact that the source code of the software is open adds to its utility in the research setting [39, 40].

A new way to integrate ImmunoMembrane into routine diagnostics comes with the increasing use of whole-slide scanners. When HER2 IHC slides are scanned as whole-slide images, pathologists need not spend time acquiring photomicrographs from the microscope. Defining analysable regions of interest in the WSI viewer is fast and yields almost perfect agreement compared with photomicrographs. This is in line with the general finding of good to superior agreement between glass slide and digital slide diagnoses [41] as well as the specific finding of equivalent results in the interpretation of HER2 IHC when using glass slides and whole-slide images [42]. In summary, digital image analysis such as that of HER2 offers a functionality that conventional work with glass slides does not have. The routine use of digital image analysis on whole-slide images offers the pathologist an enhanced diagnostic tool.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Choritz H, Büsche G, Kreipe H, Study Group HER2 Monitor (2011) Quality assessment of HER2 testing by monitoring of positivity rates. Virchows Arch 459:283–289. doi:10.1007/s00428-011-1132-8
2. Rydén L, Haglund M, Bendahl PO, et al. (2009) Reproducibility of human epidermal growth factor receptor 2 analysis in primary breast cancer: a national survey performed at pathology departments in Sweden. Acta Oncol 48:860–866. doi:10.1080/02841860902862511
3. Francis GD, Dimech M, Giles L, et al. (2007) Frequency and reliability of oestrogen receptor, progesterone receptor and HER2 in breast carcinoma determined by immunohistochemistry in Australasia: results of the RCPA Quality Assurance Program. J Clin Pathol 60:1277–1283. doi:10.1136/jcp.2006.044701
4. Vogel UF (2010) Confirmation of a low HER2 positivity rate of breast carcinomas—limitations of immunohistochemistry and in situ hybridization. Diagn Pathol 5:50. doi:10.1186/1746-1596-5-50
5. Gown AM (2008) Current issues in ER and HER2 testing by IHC in breast cancer. Mod Pathol 21:S8–S15. doi:10.1038/modpathol.2008.34
6. Saini KS, Azim Jr HA, Metzger-Filho O, et al. (2011) Beyond trastuzumab: new treatment options for HER2-positive breast cancer. Breast 20:S20–S27. doi:10.1016/S0960-9776(11)70289-2
7. Wolff AC, Hammond ME, Schwartz JN, et al. (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25:118–145. doi:10.1200/JCO.2006.09.2775
8. Rüschoff J, Hanna W, Bilous M, et al. (2012) HER2 testing in gastric cancer: a practical approach. Mod Pathol 25:637–650. doi:10.1038/modpathol.2011.198
9. Moelans CB, de Weger RA, Van der Wall E, et al. (2011) Current technologies for HER2 testing in breast cancer. Crit Rev Oncol Hematol 80:380–392. doi:10.1016/j.critrevonc.2010.12.005
10. Wolff AC, Hammond ME, Hicks DG, et al. (2013) Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol 31:3997–4013. doi:10.1200/JCO.2013.50.9984
11. Reddy JC, Reimann JD, Anderson SM, et al. (2006) Concordance between central and local laboratory HER2 testing from a community-based clinical study. Clin Breast Cancer 7:153–157. doi:10.3816/CBC.2006.n.025
12. De P, Smith BR, Leyland-Jones B (2010) Human epidermal growth factor receptor 2 testing: where are we? J Clin Oncol 28:4289–4292. doi:10.1200/JCO.2010.29.5071
13. Sauter G, Lee J, Bartlett JM, et al. (2009) Guidelines for human epidermal growth factor receptor 2 testing: biologic and methodologic considerations. J Clin Oncol 27:1323–1333. doi:10.1200/JCO.2007.14.8197
14. Minot DM, Voss J, Rademacher S, et al. (2012) Image analysis of HER2 immunohistochemical staining. Reproducibility and concordance with fluorescence in situ hybridization of a laboratory-validated scoring technique. Am J Clin Pathol 137:270–276. doi:10.1309/AJCP9MKNLHQNK2ZX
15. Dobson L, Conway C, Hanley A, et al. (2010) Image analysis as an adjunct to manual HER-2 immunohistochemical review: a diagnostic tool to standardize interpretation. Histopathology 57:27–38. doi:10.1111/j.1365-2559.2010.03577.x
16. Tuominen VJ, Tolonen TT, Isola J (2012) ImmunoMembrane: a publicly available web application for digital image analysis of HER2 immunohistochemistry. Histopathology 60:758–767. doi:10.1111/j.1365-2559.2011.04142.x
17. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. doi:10.1038/nmeth.2089
18. Leong AS, Formby M, Haffajee Z, et al. (2006) Refinement of immunohistologic parameters for Her2/neu scoring validation by FISH and CISH. Appl Immunohistochem Mol Morphol 14:384–389. doi:10.1097/01.pai.0000210415.53493.d4
19. Hall BH, Ianosi-Irimie M, Javidian P, et al. (2008) Computer-assisted assessment of the human epidermal growth factor receptor 2 immunohistochemical assay in imaged histologic sections using a membrane isolation algorithm and quantitative analysis of positive controls. BMC Med Imaging 8:11. doi:10.1186/1471-2342-8-11
20. Brügmann A, Eld M, Lelkaitis G, et al. (2012) Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. Breast Cancer Res Treat 132:41–49. doi:10.1007/s10549-011-1514-2
21. Masmoudi H, Hewitt SM, Petrick N, et al. (2009) Automated quantitative assessment of HER-2/neu immunohistochemical expression in breast cancer. IEEE Trans Med Imaging 28:916–925. doi:10.1109/TMI.2009.2012901

22. Keller B, Chen W, Gavrielides MA (2012) Quantitative assessment and classification of tissue-based biomarker expression with color content analysis. Arch Pathol Lab Med 136:539–550. doi:10.5858/arpa.2011-0195-OA
23. Rojo MG, Bueno G, Slodkowska J (2009) Review of imaging solutions for integrated quantitative immunohistochemistry in the pathology daily practice. Folia Histochem Cytobiol 47:349–354. doi:10.2478/v10042-008-0114-4
24. Rojo MG, García GB, Mateos CP, et al. (2006) Critical comparison of 31 commercially available digital slide systems in pathology. Int J Surg Pathol 14:285–305. doi:10.1177/1066896906292274
25. Nordic Immunohistochemical Quality Control (NordiQC). http://www.nordiqc.org/. Accessed 1 June 2015
26. Dendukuri N, Khetani K, McIsaac M, et al. (2007) Testing for HER2-positive breast cancer: a systematic review and cost-effectiveness analysis. Cmaj 176:1429–1434. doi:10.1503/cmaj.061011
27. Cuadros M, Villegas R (2009) Systematic review of HER2 breast cancer testing. Appl Immunohistochem Mol Morphol 17:1–7. doi:10.1097/PAI.0b013e318169fc1c
28. Minot DM, Kipp BR, Root RM, et al. (2009) Automated cellular imaging system III for assessing HER2 status in breast cancer specimens: development of a standardized scoring method that correlates with FISH. Am J Clin Pathol 132:133–138. doi:10.1309/AJCPJV0SKAF2PCMY
29. Cantaloni C, Tonini RE, Eccher C, et al. (2011) Diagnostic value of automated Her2 evaluation in breast cancer: a study on 272 equivocal (score 2+) Her2 immunoreactive cases using an FDA approved system. Appl Immunohistochem Mol Morphol 19:306–312. doi:10.1097/PAI.0b013e318205b03a
30. Turashvili G, Leung S, Turbin D, et al. (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. doi:10.1186/1471-2407-9-165
31. Laurinaviciene A, Dasevicius D, Ostapenko V, et al. (2011) Membrane connectivity estimated by digital image analysis of HER2 immunohistochemistry is concordant with visual scoring and fluorescence in situ hybridization results: algorithm evaluation on breast cancer tissue microarrays. Diagn Pathol 6:87. doi:10.1186/1746-1596-6-87
32. Vyberg M, Nielsen S, Røge R, et al. (2015) Immunohistochemical expression of HER2 in breast cancer: socioeconomic impact of inaccurate tests. BMC Health Serv Res 15:352. doi:10.1186/s12913-015-1018-6
33. Atkinson R, Mollerup J, Laenkholm AV, et al. (2011) Effects of the change in cutoff values for human epidermal growth factor receptor 2 status by immunohistochemistry and fluorescence in situ hybridization: a study comparing conventional brightfield microscopy, image analysis-assisted microscopy, and interobserver variation. Arch Pathol Lab Med 135:1010–1016. doi:10.5858/2010-0462-OAR
34. Potts SJ, Krueger JS, Landis ND, et al. (2012) Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. Lab Investig 92:1342–1357. doi:10.1038/labinvest.2012.91
35. Starczynski J, Atkey N, Connelly Y, et al. (2012) HER2 gene amplification in breast cancer: a rogues' gallery of challenging diagnostic cases: UKNEQAS interpretation guidelines and research recommendations. Am J Clin Pathol 137:595–605. doi:10.1309/AJCPATBZ2JFN1QQC
36. Walker RA (2006) Quantification of immunohistochemistry—issues concerning methods, utility and semiquantitative assessment I. Histopathology 49:406–410. doi:10.1111/j.1365-2559.2006.02514.x
37. van der Loos CM (2008) Multiple immunoenzyme staining: methods and visualizations for the observation with spectral imaging. J Histochem Cytochem 56:313–328. doi:10.1369/jhc.2007.950170
38. Kayser K, Borkenfeld S, Djenouni A et al (2015) Analysis of texture and objects in microscopic images. Diagn Pathol 1:14. doi:10.17629/www.diagnosticpathology.eu-2015-1:14
39. Ince DC, Hatton L, Graham-Cumming J (2012) The case for open computer programs. Nature 482:485–488. doi:10.1038/nature10836
40. Donoho DL (2010) An invitation to reproducible computational research. Biostatistics 11:385–388. doi:10.1093/biostatistics/kxq028
41. Jara-Lazaro AR, Thamboo TP, Teh M, et al. (2010) Digital pathology: exploring its applications in diagnostic surgical pathology practice. Pathology 42:512–518. doi:10.3109/00313025.2010.508787
42. Wilbur DC, Brachtel EF, Gilbertson JR, et al. (2015) Whole slide imaging for human epidermal growth factor receptor 2 immunohistochemistry interpretation: accuracy, precision, and reproducibility studies for digital manual and paired glass slide manual interpretation. J Pathol Inform 6:22. doi:10.4103/2153-3539.157788

PUBLICATION IV

Optimized JPEG 2000 Compression for Efficient Storage of Histopathological Whole-Slide Images

Henrik Helin, Teemu Tolonen, Onni Ylinen, Petteri Tolonen, Juha Näpänkangas, Jorma Isola

Journal of Pathology Informatics (2018) 9, 20
https://doi.org/10.4103/jpi.jpi_69_17

Technical Note

Optimized JPEG 2000 Compression for Efficient Storage of Histopathological Whole-Slide Images

Henrik Helin[1], Teemu Tolonen[2], Onni Ylinen[1], Petteri Tolonen[1], Juha Näpänkangas[3], Jorma Isola[1]

[1]BioMediTech, Faculty of Medicine, Tampere University, [2]Department of Pathology, Fimlab Laboratories, Tampere University Hospital, Tampere, [3]Department of Pathology, Oulu University Hospital, University of Oulu, Oulu, Finland

Received: 25 October 2017 Accepted: 02 April 2018 Published: 25 May 2018

Abstract

Background: Whole slide images (WSIs, digitized histopathology glass slides) are large data files whose long-term storage remains a significant cost for pathology departments. Currently used WSI formats are based on lossy image compression alogrithms, either using JPEG or its more efficient successor JPEG 2000. While the advantages of the JPEG 2000 algorithm (JP2) are commonly recognized, its compression parameters have not been fully optimized for pathology WSIs. **Methods:** We defined an optimized parametrization for JPEG 2000 image compression, designated JP2-WSI, to be used specifically with histopathological WSIs. Our parametrization is based on allowing a very high degree of compression on the background part of the WSI while using a conventional amount of compression on the tissue-containing part of the image, resulting in high overall compression ratios. **Results:** When comparing the compression power of JP2-WSI to the commonly used fixed 35:1 compression ratio JPEG 2000 and the default image formats of proprietary Aperio, Hamamatsu, and 3DHISTECH scanners, JP2-WSI produced the smallest file sizes and highest overall compression ratios for all 17 slides tested. The image quality, as judged by visual inspection and peak signal-to-noise ratio (PSNR) measurements, was equal to or better than the compared image formats. The average file size by JP2-WSI amounted to 15, 9, and 16 percent, respectively, of the file sizes of the three commercial scanner vendors' proprietary file formats (3DHISTECH MRXS, Aperio SVS, and Hamamatsu NDPI). In comparison to the commonly used 35:1 compressed JPEG 2000, JP2-WSI was three times more efficient. **Conclusions:** JP2-WSI allows very efficient and cost-effective data compression for whole slide images without loss of image information required for histopathological diagnosis.

Keywords: Digital pathology, image compression, JPEG 2000, virtual microscopy, whole-slide imaging

Introduction

Whole-slide images (WSIs), representing entire digitized histopathological tissue sections, are very large image files, usually over 20 gigabytes (GB) in size without image compression.[1] Large-scale use of whole-slide imaging in a pathology department generates tens[2,3] or hundreds[4,5] of terabytes (1 terabyte = 1000 GB) of image data each year, not including storage redundancy or backup, which further increase the storage footprint.[6,7] Owing to the high costs of storing WSIs,[8] digital archiving in a clinical setting may necessitate some form of image lifecycle management, such as deleting older WSIs from hard disks, or moving them to cheaper storage media, for example, magnetic tape.[2,3,6,9] This, in turn, counteracts one of the main advantages of WSIs over glass slides, namely, ease of access. To save storage space WSIs are compressed, usually irreversibly using so-called lossy compression algorithms such as JPEG or its successor JPEG 2000.[7,10] Lossy image compression is mathematically irreversible, meaning some image information is lost during the compression. Lossless image compression, on the other hand, is reversible, and no image information is lost in the process.[1] The degree of data compression is generally expressed as a compression ratio defined as the uncompressed file size divided by the compressed file size. Although there is no consensus regarding acceptable degrees of image compression

Access this article online

Quick Response Code:

Website: www.jpathinformatics.org

DOI: 10.4103/jpi.jpi_69_17

Address for correspondence: Dr. Jorma Isola, Faculty of Medicine, Tampere University, FI-33014 Tampere, Finland. E-mail: jorma.isola@uta.fi

How to cite this article: Helin H, Tolonen T, Ylinen O, Tolonen P, Näpänkangas J, Isola J. Optimized JPEG 2000 compression for efficient storage of histopathological whole-Slide images. J Pathol Inform 2018;9:20.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2018/9/1/20/233229

for pathology WSIs,[11,12] JPEG is thought to allow 10:1–20:1 and JPEG 2000 30:1–50:1 data compression without loss of diagnostic information.[10] JPEG compression is often defined by a nonstandard compression quality level, usually expressed as a value between 0 and 100, where the bigger the value, the better the resulting image quality is, and the less compression is applied. The compression ratio achieved with a given compression quality level depends on the image content and therefore the compression quality level is not directly proportional to compression ratio. Even when using image data compression, the costs of storing WSIs remain high in pathology laboratories using WSI in routine practice.

The JPEG 2000 image compression method has been designed to allow a highly customizable way of compressing image data by user-controlled parametrization.[13] The compression parameters of the JPEG 2000 algorithm (JP2) have so far not been fully optimized for digital pathology.[11] We have previously employed a fixed compression ratio to allow fast remote viewing of WSIs over the internet using JPIP (JPEG 200 interactive protocol).[13] Newer image server software is able to read and decompress JPEG 2000 files on the fly and send image tiles to the client through hypertext transfer protocol. This allows testing a wider set of compression parameters for producing maximal file size reduction while avoiding image compression artifacts. The present study demonstrates a novel image content sensitive strategy for pathology WSI-optimized JPEG 2000 compression, designated JP2-WSI. The new compression parametrization is compared to commercial WSI formats and the commonly used JPEG and constant compression ratio JPEG 2000 in terms of file sizes and visual image quality.

Procedure

The concept of whole slide image-optimized parametrization for JPEG 2000 image compression

The prevailing method of defining JPEG 2000 image compression is to choose a fixed compression ratio for the scanned microscope slides.[7] However, due to the highly variable amount of diagnostically irrelevant background on the slides [Figure 1], fixed ratio compression has not turned out to produce optimal file size reduction for WSIs. When using fixed ratio compression, the more there is tissue in relation to empty slide area, the more details must be discarded to achieve the desired compression ratio. Thus, a fixed compression ratio will result in variable image quality on tissue-containing image areas. To avoid too low image quality on any WSIs, we have previously defined compression ratios of 25–30:1,[13] and subsequently 35:1, as suitable for JPEG 2000 compression of histopathological WSIs. When using standard JPEG compression, the image quality associated with level 80 compression out of 100 has been considered suitable for large-scale applications of WSI in pathology.[2,3] Compression quality level is not a concept that is defined in the JPEG standard and as such is not unambiguous. In our experience, level 80 JPEG compression produces typically compression ratios of about 1:20 in pathology WSIs.

We defined JP2-WSI compression to match the image quality of hematoxylin and eosin stained tissue sections scanned and stored with JPEG level 80 compression, followed by maximal compression of the empty slide area (the background). For assessing the image quality of the tissue, we used peak signal-to-noise ratio (PSNR) measurements[14] and visual inspection by two senior pathologists (TT and JI). In our validation data of seventeen routine histopathological slides, the mean PSNR values of JPEG level 80 and JP2-WSI compression were 33.4 dB and 35.8 dB (with lossless image used as reference). The corresponding mean value for fixed 35:1 ratio JPEG 2000 was 40.8 dB. In visual inspection by two senior pathologists, there was no significant loss of overall image quality associated with JP2-WSI compression compared to JPEG level 80 compression.

Aside from the variable compression ratio, the main codestream parameters for JP2-WSI were essentially the same as we have used before: Eight wavelet decomposition levels, no tiling, precinct size 256 × 256, code-block size 64 × 64, progression order resolution-position-component-layer, and one quality layer.[13]

Comparison of JP2-WSI to standard JPEG and JPEG 2000 compression methods

For evaluating JP2-WSI, we digitized a set of seventeen histopathological slides selected from a university hospital pathology archive. These slides reflect the routine workload of a general pathologist. The slides included needle biopsies (n = 10) as well as surgical sections (n = 7) and were stained with hematoxylin and eosin (n = 15) and Giemsa (n = 2). Figure 1 presents an overview of the slide set. As an additional test of diagnostically challenging WSI image quality, we digitized a gastric biopsy slide to verify the image quality of JP2-WSI in visualizing Helicobacter pylori. The slides were digitized with whole slide scanners from four different vendors, including two-line scanners (Aperio and Hamamatsu) and two tile-based scanners (3DHISTECH and Jilab). The scanner setups were as follows

1. Aperio ScanScope AT2 (Leica Biosystems, Nussloch, Germany) brightfield line scanner, Piranha Color 2k PC-30-02K80 camera (Teledyne DALSA, Ontario, Canada) with 2048 × 3 pixel resolution, pixel size 14 × 14 μm, ×20 Olympus Plan-Apo objective lens with a numerical aperture (NA) of 0.75, and scanning resolution 0.5 μm/pixel
2. Hamamatsu NanoZoomer XR (Hamamatsu Photonics, Hamamatsu, Japan) brightfield line scanner, charge-coupled device (CCD) camera with 4096 × 64 pixel resolution, pixel size 8 × 8 μm, ×20 Olympus Plan-Apo objective lens (NA 0.75) and ×1.75 relay lens, scanning resolution 0.46 μm/pixel
3. Pannoramic SCAN (3DHISTECH Ltd, Budapest, Hungary) brightfield tile-based scanner, CIS 3CCD camera with 2048 × 2048 pixel resolution, pixel size 5.5 × 5.5 μm, ×20 Carl Zeiss Plan-Apochromat objective lens (NA 0.8) and ×1 phototube, scanning resolution 0.24 μm/pixel

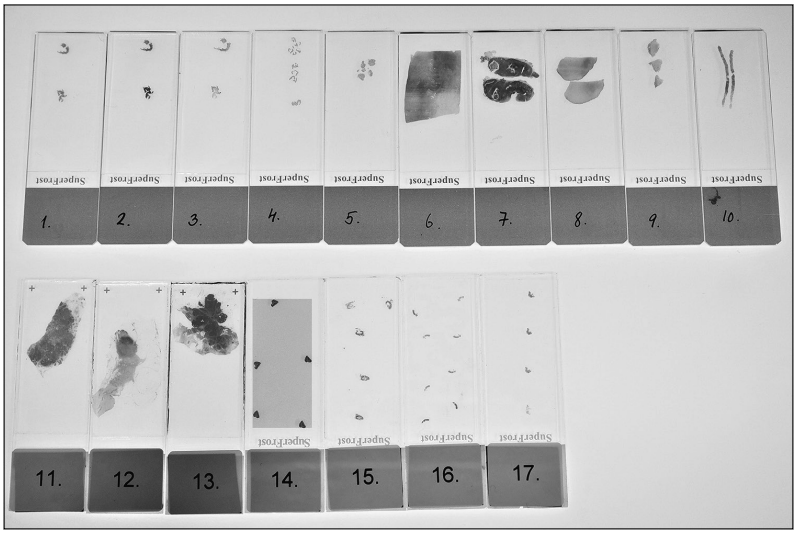


Figure 1: A macroscopic view of the seventeen routine histopathological slides used in the study. The shaded rectangle on slide fourteen demonstrates the area that makes up the whole-slide image containing both tissue and empty slide area

4. SlideStrider (Jilab Inc, Tampere, Finland) brightfield tile-based scanner, Lumenera Lt1265R CCD camera (Lumenera Corporation, Ottawa, Ontario, Canada) with 4240 × 2832 pixel resolution, pixel size 3.1 μm × 3.1 μm, ×10 (NA 0.4) and ×20 (NA 0.75) Olympus UPLSAPO objective lenses, scanning resolution 0.16–0.31 μm/pixel.

The area included in the WSI was the smallest rectangle covering all individual tissue fragments on the slide. For the SlideStrider scans, the nonscanned empty slide areas required to fill in the WSI rectangle were copied automatically from a standard empty slide image tile.

All WSIs were scanned as 24-bit RGB color images. WSIs from the Aperio, Hamamatsu, and Pannoramic scanners were saved without compression and then compressed with JP2-WSI. The same scanned WSIs were also saved using the manufacturers' proprietary file formats and their default compression schemes. Aperio SVS format used JPEG tile compression with compression level set at 70/100. Both Hamamatsu NDPI format and Pannoramic MRXS format employed JPEG tile compression with quality level 80/100. The WSIs scanned with SlideStrider were first saved losslessly and then converted to either fixed 35:1 ratio JPEG 2000 or the developed JP2-WSI compression. The compression method of the SlideStrider software is based on the Kakadu software development kit library implementation of JPEG 2000 (version 7.5, Kakadu Software Inc., NewSouth Innovations Pty Limited, Sydney, Australia).[15] The four scanners all had different sampling resolutions, Aperio 0.5 μm/pixel, Hamamatsu 0.46 μm/pixel, Pannoramic 0.24 μm/pixel, and SlideStrider 0.16–0.31 μm/pixel.

Results

Table 1 presents the pixel dimensions, the ratios of empty slide to tissue area, and the file sizes of the 17 slides digitized with the SlideStrider scanner at 0.31 μm/pixel. The uncompressed file sizes ranged from 2.6 to 30 GB. Lossless JPEG 2000 compression yielded compression ratios ranging from 3:1 to 56:1 and file sizes from 341 megabytes (MB) to 5.9 GB. The fixed ratio JPEG 2000 algorithm compressed all images to the 35:1 extent, except for two cases (slides 14 and 16), for which higher compression ratios were already achieved with the lossless algorithm. The file sizes ranged from 74 MB to 686 MB. The developed JP2-WSI compression produced overall compression ratios varying from 41:1 to 1487:1, and file sizes of 8 MB to 442 MB. As an average, using JP2-WSI, we obtained file sizes that were 33% of fixed-ratio lossy compressed JPEG 2000. Of the individual scanned histopathology test slides, JP2-WSI reduced file sizes most effectively in biopsy slides containing multiple small tissue fragments and abundant empty slide area (slides 14, 16, and 17 in our test set). The ratio of empty slide area to tissue-containing slide area showed an approximately linear relationship with the overall compression ratio achieved with JP2-WSI [Figure 2].

Figures 3 and 4 allow comparison of the visual image quality obtained with JP2-WSI compression compared to JPEG quality level 80 compression and fixed 35:1 ratio JPEG 2000 compression. In Figure 3, visually detectable differences can be seen only with zoom levels well over 100%, which represent purely digital magnification. The magnified screenshots come from WSIs with file sizes of 528 MB and 21 MB (JPEG 2000 35:1 and JP2-WSI, respectively) and 611 MB and 18 MB (JPEG 2000 35:1 and JP2-WSI, respectively). They were scanned using the SlideStrider whole-slide

Table 1: Comparison of whole-slide image file sizes produced by three different parametrizations of JPEG 2000

Slide	WSI dimensions in pixels	Ratio of empty slide to tissue area	Uncompressed file size (MB)	JP2-lossless file size (MB)	JP2-35:1 file size (MB)	JP2-WSI file size (MB)
1	19,728×71,824	9.0	4251	387	122	13
2	22,544×67,600	10.1	4572	347	131	12
3	25,360×67,600	11.5	5143	387	147	12
4	19,728×80,272	6.7	4751	593	136	17
5	25,360×33,808	3.2	2572	448	74	12
6	64,784×92,944	0.3	18,064	5545	517	442
7	67,600×67,600	1.0	13,709	3431	392	289
8	56,336×63,376	1.0	10,711	2382	306	120
9	19,728×59,152	1.9	3501	634	100	26
10	22,544×80,272	3.0	5429	928	155	38
11	64,784×92,944	1.2	18,064	3512	517	304
12	67,600×118,288	0.6	23,989	5896	686	284
13	67,600×67,600	0.2	13,709	4091	392	248
14	67,600×147,856	49.0	29,985	528	528	21
15	56,336×135,184	32.3	22,847	669	648	25
16	61,968×143,632	49.0	26,702	611	611	18
17	19,728×139,408	24.0	8251	341	236	8

WSI: Whole-slide image, MB: Megabytes, JP2-lossless: Lossless JPEG 2000 compression, JP2-35:1: Fixed 35:1 ratio JPEG 2000 compression, JP2-WSI: WSI-Optimized JPEG 2000 compression

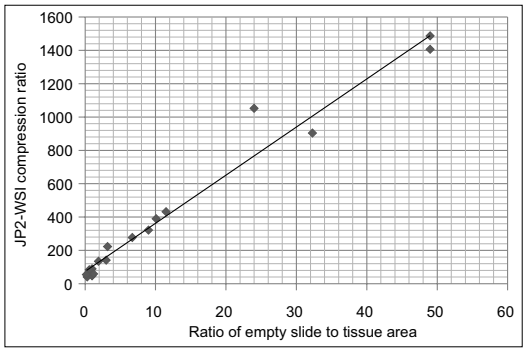


Figure 2: The compression ratio obtained by JP2-WSI plotted against the ratio of empty slide to tissue area in the set of seventeen digitized slides. The approximately linear relation signifies JP2-WSI's sensitivity to image content – the greater the ratio of empty slide to tissue area the higher the overall compression ratio that is achieved

scanner with x10 objective lens and a charge-coupled device camera with 3.1 μm pixel size resulting in 0.31 μm/pixel scanning resolution. Figure 4 presents zoomed screenshots of Helicobacteria in a gastric biopsy scanned with resolutions of 0.31 μm/pixel and 0.16 μm/pixel (Plan-Apo ×10 and ×20 objective lenses, respectively). At 0.31 μm/pixel, there are subtle visible differences in the image quality. JP2-WSI eliminates random noise resulting in a smooth or blurry image appearance, whereas JPEG produces a grainier or noisier image. With the higher optical scanning resolution we were unable to detect diagnostic differences in image quality.

Table 2 shows the file sizes resulting from digitizing the set of 17 slides with 3DHISTECH, Aperio and Hamamatsu scanners.

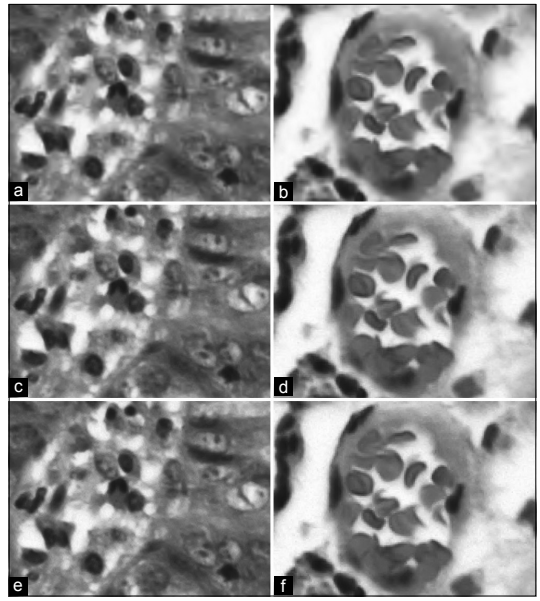


Figure 3: Comparison of image quality between JP2-WSI (a and b), JPEG quality level 80 (c and d), and fixed 35:1 ratio JPEG 2000 compression (e and f). Digitally magnified screenshots of whole-slide images 14 (a, c, e) and 16 (b, d, f)

For each scanner, three different file sizes are shown per slide: the raw uncompressed file size, the file size using the scanners default compression method, and the file size using JP2-WSI compression. The file sizes are not comparable between scanners because of different scan area dimensions and different scanning

Table 2: File sizes of the slides digitized with different scanner and compression method combinations

Slide	WSI file size in GB with compression ratio in parentheses								
	3DHISTECH			Aperio			Hamamatsu		
	Raw	Default	JP2-WSI	Raw	Default	JP2-WSI	Raw	Default	JP2-WSI
1	16.77	0.27 (63)	0.03 (643)	6.34	0.32 (20)	0.02 (270)	11.89	0.36 (33)	0.03 (446)
2	18.48	0.30 (62)	0.03 (678)	7.10	0.31 (23)	0.02 (333)	13.08	0.39 (34)	0.03 (498)
3	18.58	0.28 (66)	0.03 (709)	9.98	0.44 (23)	0.02 (479)	11.01	0.34 (32)	0.03 (432)
4	21.15	0.37 (57)	0.03 (649)	7.24	0.43 (17)	0.03 (243)	9.87	0.33 (30)	0.03 (301)
5	7.91	0.29 (28)	0.04 (221)	6.62	0.47 (14)	0.03 (205)	4.84	0.23 (21)	0.03 (143)
6	39.63	4.13 (10)	0.60 (66)	36.07	5.35 (7)	0.76 (48)	40.87	4.10 (10)	0.84 (49)
7	33.90	2.56 (13)	0.48 (70)	36.46	3.30 (11)	0.40 (91)	27.35	2.32 (12)	0.44 (62)
8	25.97	1.59 (16)	0.29 (89)	14.80	1.94 (8)	0.27 (55)	20.35	1.53 (13)	0.29 (70)
9	18.48	0.48 (38)	0.07 (271)	7.00	0.48 (15)	0.06 (116)	10.66	0.50 (22)	0.07 (151)
10	16.89	0.57 (29)	0.07 (238)	7.79	0.61 (13)	0.06 (124)	10.70	0.50 (21)	0.07 (151)
11	47.22	2.61 (18)	0.47 (100)	28.07	2.89 (10)	0.40 (70)	37.79	2.57 (15)	0.46 (83)
12	56.35	3.28 (17)	0.50 (112)	40.37	3.73 (11)	0.42 (96)	49.81	2.87 (17)	0.47 (105)
13	34.23	3.06 (11)	0.52 (66)	25.82	3.41 (8)	0.44 (59)	29.37	2.49 (12)	0.47 (62)
14	59.32	0.37 (158)	0.03 (1739)	43.14	1.48 (29)	0.03 (1289)	52.85	1.27 (42)	0.04 (1342)
15	62.33	0.51 (123)	0.04 (1641)	49.37	1.55 (32)	0.04 (1189)	49.68	1.23 (40)	0.05 (1033)
16	63.15	0.50 (127)	0.03 (2250)	37.20	1.56 (24)	0.04 (905)	49.68	1.24 (40)	0.05 (1078)
17	35.64	0.24 (146)	0.02 (2093)	13.22	0.58 (23)	0.02 (671)	16.21	0.41 (39)	0.02 (862)

Raw: Uncompressed WSI, default: The scanners default compression method (see text for details). The file sizes are not comparable between scanners due to variable scan area dimensions and scanning resolutions. WSI: Whole-slide image, GB: Gigabytes

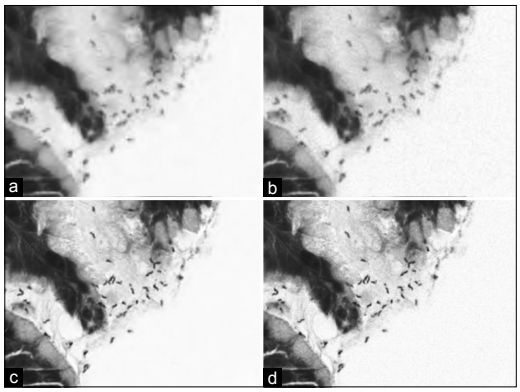


Figure 4: Effects of image compression and whole-slide image scanning resolution on detecting Helicobacteria in a gastric biopsy. JP2-WSI (a and c) and JPEG quality level 80 (b and d) compressed images scanned at 0.31 μm/pixel (a and b) and 0.16 μm/pixel (c and d) sampling resolutions

resolutions. The uncompressed file sizes ranged from 4.84 GB to 63.15 GB. 3DHISTECH default compression produced compression ratios of 10–158 while JP2-WSI compressed the same images with ratios of 66–2250. For Aperio, the compression ratios were 7–32 for default compression and 48–1289 for JP2-WSI. Hamamatsu default compression produced compression ratios of 10–42 with JP2-WSI producing compression ratios of 49–1342. JP2-WSI compression had the widest range of overall compression ratios, 66–2250, 48–1289 and 49–1342 for 3DHISTECH, Aperio, and Hamamatsu scanned images, respectively. JP2-WSI compressed images

Table 3: Mean whole-slide images file sizes produced by different scanner and compression method combinations

Compression method	Mean WSI file size (mean compression ratio)		
	3DHISTECH	Aperio	Hamamatsu
No compression	33.88	22.15	26.24
Scanner default compression	1.26 (27)	1.70 (13)	1.33 (20)
JP2-WSI compression	0.19 (176)	0.18 (122)	0.20 (131)

File sizes differ between scanners due to variable scan area dimensions and scanning resolutions. See text for the scanner default compression methods. File sizes are in gigabytes. WSI: Whole-slide image

were smallest and had the highest overall compression ratios in every case. All of the compression methods produced the highest compression ratios for the biopsy slides 14–17. These slides had the highest ratios of empty slide to tissue area [Figure 1]. Table 3 summarizes the mean file sizes and mean compression ratios of the slides digitized with 3DHISTECH, Aperio, and Hamamatsu scanners. JP2-WSI compressed file sizes were 15%, 9%, and 16% of the file sizes produced by 3DHISTECH, Aperio, and Hamamatsu default compression methods.

Discussion

Due to the costly hard disk storage of modern pathology WSIs, it is essential to optimize data compression for routine applications of digital pathology. Many commercial whole slide scanners still employ JPEG compression in their proprietary file formats, probably mainly due to its computational simplicity. Its successor, JPEG 2000, offers more powerful wavelet-based compression, and its random

data access, among other features, makes it well suited for server-based remote WSI viewing.[10,13] An optimal WSI data compression method should take into account the type of the images to be stored.[16] In the case of pathology slides, this means taking advantage of the abundant empty space on the slide. With this in mind, we designed our novel WSI-optimized JPEG 2000 parametrization, JP2-WSI, which compresses the empty slide area very heavily but retains image quality on the tissue-containing parts of the image. This not only minimizes the effect of empty slide area on WSI file size but also adjusts the overall compression ratio to the amount of image detail meaning that the image quality rather than overall compression ratio is kept constant. A standard practice in histopathology is to distribute several serial sections of small biopsies on the slide. This leads inevitably to biopsy-WSIs containing larger empty slide areas than WSIs of surgical specimens which have been scanned with tight margins. JP2-WSI was designed to handle both WSI types but was found particularly effective in reducing the file size of biopsy-WSIs.

When designing a novel compression method for WSIs, it was self-evident that the image quality may not be compromised. We designed JP2-WSI to keep the inevitable information loss associated with all lossy image compression algorithms essentially the same as that obtained with JPEG level 80 compression, which has been considered a satisfactory image quality in histopathological WSIs.[2,3] Using this definition, JP2-WSI yielded file sizes that were <20% of those obtained by proprietary JPEG 80 compression, and only 33% of conventional fixed-ratio JPEG 2000 compression. This permits significant cost savings in the routine use of WSI, where tens or even hundreds of terabytes of image data are generated each year.[2-5]

Conclusion

We developed a novel parametrization of JPEG 2000 image compression designed for histopathological WSIs. The main advantage of JP2-WSI is its sensitivity to image content. Our optimized JP2-WSI conforms to the JPEG 2000 image coding standard maintaining its open, nonproprietary nature. JPEG2000 encoding is an allowed image compression method in the DICOM standard, and it is also allowed for the WSI storage class. A commercial picture archiving and communications system has recently implemented JPEG 2000 encoded WSI storage class in full workflow (Neagen, Oulu, Finland). This demonstrates that JPEG 2000 encoding is applicable also in the DICOM context. A large-scale clinical study is underway to confirm the diagnostic accuracy of JP2-WSI-compressed WSIs compared to conventional glass slides and optical microscopes.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Pantanowitz L, Dickinson K, Evans AJ, Hassell LA, Henricks WH, Lennerz JK, *et al*. American telemedicine association clinical guidelines for telepathology. J Pathol Inform 2014;5:39.
2. Stathonikos N, Veta M, Huisman A, van Diest PJ. Going fully digital: Perspective of a Dutch academic pathology lab. J Pathol Inform 2013;4:15.
3. Huisman A, Looijen A, van den Brink SM, van Diest PJ. Creation of a fully digital pathology slide archive by high-volume tissue slide scanning. Hum Pathol 2010;41:751-7.
4. Clunie DA, Dennison DK, Cram D, Persons KR, Bronkalla MD, Primo HR, *et al*. Technical challenges of enterprise imaging: HIMSS-SIIM collaborative white paper. J Digit Imaging 2016;29:583-614.
5. Lundström C, Thorstenson S, Waltersson M, Persson A, Treanor D. Summary of 2nd nordic symposium on digital pathology. J Pathol Inform 2015;6:5.
6. Chlipala E, Elin J, Eichhorn O, Krishnamurti M, Long RE, Sabata B, *et al*. Archival and Retrieval in Digital Pathology Systems. Available from: http://www.digitalpathologyassociation.org/_data/files/Archival_and_Retrieval_in_Digital_Pathology_Syste ms.pdf. [Last accessed on 2017 Oct 21].
7. García-Rojo M. International clinical guidelines for the adoption of digital pathology: A review of technical aspects. Pathobiology 2016;83:99-109.
8. Häger S. How to build a business case to justify the investment in digital pathology. Available from: https://www.sectra.com/medical/resources/how-to-build-a-business-case-to-justify-the-investment-in-digital-pathology/. [Last accessed on 2017 Oct 21].
9. Thorstenson S, Molin J, Lundström C. Implementation of large-scale routine diagnostics using whole slide imaging in Sweden: Digital pathology experiences 2006-2013. J Pathol Inform 2014;5:14.
10. Digital Imaging and Communications in Medicine (DICOM) – Supplement 145: Whole Slide Microscopic Image IOD and SOP Classes. DICOM Standards Committee, Working Group 26, Pathology. Available from: http://www.medical.nema.org/medical/dicom/final/sup145_ft.pdf. [Last accessed on 2017 Oct 21].
11. Krupinski EA, Johnson JP, Jaw S, Graham AR, Weinstein RS. Compressing pathology whole-slide images using a human and model observer evaluation. J Pathol Inform 2012;3:17.
12. Kalinski T, Zwönitzer R, Grabellus F, Sheu SY, Sel S, Hofmann H, *et al*. Lossless compression of JPEG2000 whole slide images is not required for diagnostic virtual microscopy. Am J Clin Pathol 2011;136:889-95.
13. Tuominen VJ, Isola J. The application of JPEG2000 in virtual microscopy. J Digit Imaging 2009;22:250-8.
14. Wang Z, Bovik AC, Lu L. Why is image quality assessment so difficult. In: IEEE International Conference on Acoustics, Speech, and Signal Processing. Orlando, FL, USA: IEEE; 2002. p. 3313-6.
15. Kakadu Software. Available from: http://www.kakadusoftware.com. [Last accessed on 2017 Oct 21].
16. Muir P, Li S, Lou S, Wang D, Spakowicz DJ, Salichos L, *et al*. The real cost of sequencing: Scaling computation to keep pace with data generation. Genome Biol 2016;17:53.