Development and Validation of Software Solutions to Eliminate Bottlenecks in the Application of High-Throughput Quantitative Immunohistochemistry

A thesis submitted for the degree of Ph.D. by
Catherine Conway, B.Sc.
January 2008

The research work in this thesis was carried out under the supervision of

Dr. Donal O'Shea
Medical Informatics Group,
School of Biotechnology,
Dublin City University

and

Dr. Lorraine O'Driscoll
National Institute for Cellular Biotechnology,
Dublin City University

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____   ID No:_____

Date: _____

ACKNOWLEDGMENTS

went above and beyond the call of duty on a daily basis, and always did it will a smile on your face, thank you.

I would like to thank my fellow post grads past and present, Sean, Dan, Cormac, John, Lynne, Damien, Zelda and Jenny. Thank you all for helping me see the funny side when work went wrong. You are a fantastic bunch of people and an extremely hard act to follow. Special thanks to Zelda for being the most efficient post grad that ever lived, and for showing me how it's done, to Damien for being the calmest cookie that ever walked, to John for his constant friendship, and the many times he brought me to tears of laughter and finally to Lynne, you have a heart made from pure gold and I am eternally grateful for our friendship.

Thanks to my brother and sisters, Mary, Elaine and Thomas who have had to put up with some truly terrible presents over the past few years, in fact limiting it to the past few years maybe a bit of an understatement. You have all contributed in your own way, your constant thoughts and prayers have lifted my spirits along the way. I am extremely proud to have you all as my siblings. I also have to thank you for my wonderful nieces and nephews who have provided me with hours of light hearted relief, at times when I needed it the most. In particular, I want to thank my sister Mary; you have selflessly given me your time and constant support. This day would simply not have come without you. You truly are wonder women in my eyes.

Thanks to my boyfriend Neil, you have been at the cold face of this since the start, and suffered more than most from the high and lows of doing this PhD. Thank you for your patience, understanding and for always having faith in me, even when I had lost faith in myself. You have carried me when I was at my lowest and I appreciate you more than will ever know.

Finally, I wish to thank the two people that have made the completion of this thesis possible my parents, Sean and Annetta Conway. The example you both have set, of hard work and determination has been my inspiration throughout my life. No parents could do more for their children. You have earned this day, every bit as much as I have. My proudest moment is knowing what joy the completion of this thesis is

giving you both. Words can not do justice to the input you two have had, so all that I can simply say is, thank you both.

TABLE OF CONTENTS

ABSTRACT

Title: Development and Validation of software solutions to eliminate bottlenecks in the application of high-throughput quantitative immunohistochemistry

Author: Catherine Conway

Tissue Microarrays facilitate high-throughput immunohistochemistry; however, there are key bottlenecks apparent in their analysis, particularly in the domain of performing manual reviews by eye using microscopes.

The advent of Virtual Slides has permitted the review of tissue slides across the Internet. Virtual slides also enable the creation of software solutions to assist in the review of Tissue Microarrays. The Virtual Tissue Matrix, a web application that allows the online review of digitised Tissue Microarrays, was created as part of this work. The performance of the Virtual Tissue Matrix was validated by comparing the accuracy and reproducibility of human assessment of immunohistochemically stained Tissue Microarrays using this and the conventional microscope-based approach. Results illustrated that this means of review was as accurate as microscope-based reviews. However, neither method of assessment was truly reproducible, particularly when intensity of membrane immunohistochemistry was assessed. Inherent flaws in human vision prohibit the accurate determination of membrane staining intensity, resulting in low inter- and intra-observer agreement.

As a result, an image analysis algorithm to quantify intensity of membrane-bound immunohistochemical staining was created. The feasibility of utilising computer-assisted image analysis algorithms as an alternative to human microscope-based reviews was extensively assessed across clinical and research datasets. Membrane-bound immunohistochemical staining was assessed by image analysis and compared with protein and gene expression levels when assessed by enzyme-linked immunosorbent assay, Western blots, and fluorescent *in situ* hybridisation. In addition, protein expression levels when assessed by human and image analysis were compared. Results illustrated that image analysis surpassed the performance of human analysis when identifying prognostic outcome of urothelial cell carcinoma of the bladder and as good as humans in assessment of HER-2 status of invasive carcinoma. Image analysis was capable of higher-throughput, greater objectivity, and reproducibility in quantifying membrane staining intensity.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

°C	Degrees Celsius
ACIS	Automated Cellular Imaging Systems
Akt	Protein kinase
API	Association of Pathology Informatics
AQUA	Automated quantitative analysis
ATA-I	Automated Tissue Arrayer I
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
Bcl-2	B cell lymphoma 2
cDNA	Complementary Deoxyribonucleic acid
CI	Confidence interval
CO_2	Carbon Dioxide
Cox-2	Cyclooxygenase 2
CSV	Comma separated value
DAB	3',3-daminobenzidine tetrahydrochloride
DAPI	Diamidino-2-phenylindole
DB	Database
Df	Degrees of freedom
DLL	Dynamic link library
DNA	Deoxyribonucleic acid
ECACC	European Collection of Cell Cultures
EGFR	Epidermal growth factor receptor
ELISA	Enzyme linked immunosorbant assay
ER	Estrogen receptor
FCS	Fetal calf serum
FDA	Food and Drug Administration
FFPE	Formalin-fixed-paraffin-embedded
FISH	Fluorescent *in situ* hybridisation

GB	Gigabyte
H & E	Haematoxylin and Eosin
HER-2	Human Epidermal Growth Factor 2, also know as ErbB – 2,
HSL	Hue, Saturation, Luminance
HTML	Hypertext markup language
IHC	Immunohistochemistry
IMS	Industrial methylated spirits
ISH	*In situ* hybridisation
LED	Light emitting diode
MAPK	Mitogen activated protein kinases
Mb	Megabyte
Mbps	Megabits per second
MDR	Multidrug resistance
mRNA	Messenger ribonucleic acid
MTA-I	Manual Tissue Arrayer I
MTA-II	Manual Tissue Arrayer II
NA	Sodium
NCI	National Cancer Institute
NCTCC	National Cell and Tissue Culture Centre
NEAA	Non essential amino acids
NICB	National Institute for Cellular Biotechnology
P53	Protein 53
PBS	Phosphate buffered saline
pEGFR	Phosphorylated epidermal growth factor receptor
PR	Progesterone receptor
pRB	Phosphorylated Retinoblastoma protein
RAM	Random access memory
RCSI	Royal College of Surgeons Ireland
RESA	Rapid exponential subtraction algorithm

RGB	Red, Green, Blue
RIA	Radioimmunoassay
RNA	Ribonucleic acid
RPM	Revolutions per minute
RR	Relative Risk
SQL	Structured Query Language
TARP	Tissue Array Research Program
TBS	Tris buffered Saline
TMA	Tissue Microarrays
TNM	Tumour, Node, Metastasis
UCB	Urothelial cell carcinoma of the bladder
VTM	Virtual Tissue Matrix

CHAPTER 1: OVERVIEW OF TISSUE MICROARRAYS, MICROSCOPE AND COMPUTER-ASSISTED REVIEW OF IMMUNOHISTOCHEMICALLY STAINED TISSUE

1.1   Introduction

Proteins are vital parts of living organisms, as they control the main components of the physiological metabolic pathways of the cell. The majority of diagnostic and drug targets are comprised of specific receptors, enzymes or ion channels, all of which are protein localised either on or within the diseased cell or within the surrounding cells. Although traditional methods of investigations have yielded abundant information about individual proteins, they have been less successful at providing an integrated understanding of biological systems. By studying many components simultaneously the mechanism by which proteins interact with each other, as well as with other molecules, to control complex processes within cells, tissues and even whole organisms can be determined (MacBeath, 2002). Therefore, examination of protein expression within sub-cellular compartmentalisation and cellular distribution on a large number of specimens is of particular importance in proteomic research (Tzankov et al., 2005).

Numerous platforms are available for the examination of protein expression. Methods include; 2-D Electrophoresis/Mass spectrometry, Surface Enhanced Laser Desorption /Ionization, ICAT, and Protein Arrays. However, these high-throughput methods have prohibitive limitations including; relying on serum samples with known protein sequences, (2-D Electrophoresis/Mass Spectrometry), dependence on instruments that are difficult to calibrate and which produce lot-to-lot variability of reagents (Surface enhanced laser desorption / ionisation), high costs, technical difficulties and relative scarcity of the method (ICAT), stability of arrays is uncertain and difficulties with labelling methods (Protein arrays) (Hewitt et al., 2004). In addition, such methods are disadvantaged by the fact that actual tissue is disintegrated before analysis is performed, therefore, identification of the cell type and localisation expression of the protein is not possible (Simon et al., 2003).

Table 1.1-1 illustrates numerous methods of tissue analysis (Camozzi and Razvi, 2004). Lysate arrays allow the analysis of entire protein profiles in single or multiple signalling pathways; they are so sensitive that measurements are reliably done on one cell equivalent and with very small reagent volumes, which allows the measurement

of protein profiles in a cost and time-efficient way. However, cell type and protein localisation can not be identified when using Lysate arrays. Using blot technology is possible to identify DNA (Southern blot), RNA (Northern blot) and Protein expression levels (Western blot). However, throughput of analysis is low.

Immunohistochemistry (IHC) is a method of evaluating protein expression within cells by localising specific antigens in tissues based on antigen-antibody recognition. Immunohistochemistry is the assay of choice when performing protein expression profiling, as disintegration of the cell prior to protein expression profiling is not required. In addition, immunohistochemistry is a widely used and versatile technique. However, validations of newly identified potential target candidates require analysis of hundreds of individual tissue specimens. Conventional immunohistochemical analysis of tissue specimens on the molecular level is labour intensive, time-consuming and semi-quantitative at best and, in the long run, will not allow scientists to keep up with the rate at which new targets for tissue investigation are identified (Hoos and Cordon-Cardo, 2001). Therefore, new technologies that facilitate, high-throughput, standardisation, localisation identification and reagent economy are required for protein expression profiling.

Table 1.1-1: Comparison of different methods of tissue analysis. ++++ excellent, +++ very good. ++ good, + possible, -- limited, --- not possible (Camozzi and Razvi, 2004).

Methods	Whole Section IHC	Lysate Array	Western blot	Northern blot	Southern blot
Compartment-specific antibodies	+++	---	---	---	---
DNA Copy Number Analysis	+++	---	---	---	++
RNA Expression Analysis (Frozen tissue)	++	---	---	+++	---
Protein Expression Analysis	+++	+++	+++	---	---
Detection of RNA splice variants	---	---	---	+++	---
Detection of Protein Isoforms	+	++	+++	---	---
Mutation Analysis	---	---	---	---	---
Standardisation of Analysis	--	++++	+	+	+
Quantification Analysis	+	++	+	+	+
Reagent Economy	--	+++	+	+	+
Analysis Speed	---	+++	--	--	--

1.2   Tissue Microarrays

The development of high-throughput analytical platforms utilising nucleic acids for analysis of cells and tissues has resulted in the demand for higher throughput platforms for the confirmation of these findings. Analysis of tissue for the expression of genes, either at the nucleic acid level or at the protein level, remains the gold standard validation of experimental data obtained by other methods. The two most common methods of validation have been *in situ* hybridisation (ISH) and immunohistochemistry on tissue with interpretation of the histomorphology to discern the complexity of expression patterns that cannot be determined from methods that rely on the extraction of biomolecules (Hewitt, 2006). Recent advances in molecular biology have centred on increases in throughput and quantification of biologic phenomena. No longer is experimental design focused on one gene or one protein, but rather on tens to hundreds of genes, proteins or tissue on analytical platform (MacBeath, 2002).

The use of Tissue Microarrays (TMAs) has become a common tool to both validate and generalize the results of microarray experiments. Prior to Tissue Microarray (TMA) development, the validation of microarray data had become a bottleneck for both basic and clinical research (Hewitt, 2006). Tissue Microarrays provide high-throughput analysis of histomorphologic examination of tissue for *in situ* hybridisation and immunohistochemistry, by means of arranging multiple tissue samples in a uniform structure on the surface of a glass slide. Tissue Microarrays allow for large numbers of tissue samples to be analysed simultaneously based on fluorescence *in situ* hybridisation (FISH) for genetic rearrangements, RNA *in situ* hybridisation for genetic expression, or immunohistochemistry for protein overexpression. Tissue Microarray technology was developed by Kononen et al. (1998) in order to facilitate gene expression and copy number surveys of large cohorts of tumours. Due to the nature of Tissue Microarray construction which allows multiple sections to be obtained from a single Tissue Microarray block. Rapid analysis of hundreds of molecular markers on the same cohort of specimens is now possible.

Utilising Tissue Microarrays disease-specific expression of a gene can be identified and, therefore, a possible diagnostic test and therapeutic target can be developed (Simon et al., 2003). Tissue Microarrays provide substantial value in rapidly translating genomic and proteomic information into clinical applications (Torhorst et al., 2001). When initially created, Tissue Microarrays were envisioned to make a dramatic impact on basic cancer research and anatomic pathology (Moch et al., 2001); in almost ten years since Tissue Microarrays development this hypothesis has been realised through various studies.

1.3   Evolution of Tissue Microarrays

Historically, immunoassays are the predominant method utilised for protein quantification. As early as 1929, antibodies were used in serology to precipitate antigens for subsequent quantification. Analytical immunoassay technology was greatly advanced with the introduction of the radioimmunoassay (RIA) by Yarlow and Berson in 1959 and enzyme-linked immunosorbant assay (ELISA) by Engvall and Perlman in 1971. In late 1980s, Roger Ekins argued that extremely sensitive quantitative assays could be developed using microspots of antibodies on solid supports and that multianalyte assays could be carried out using spatially separated arrays of such spots. It was not until the late 1990s that the benefits of deoxyribonucleic acid (DNA) microarray technology were well recognised and that equipment to manufacturer microarrays became more accessible, that attention shifted to quantitative proteomics (MacBeath, 2002).

 Due to advances in high density, high-throughout complementary DNA microarray (cDNA) and oligonucleotide technologies and the completion of the sequencing of the human genome, the rate of discovery of new genes that potentially could be used in determining disease diagnosis, prognosis and response to specific therapeutics regimens increased rapidly. Validations of newly identified potential target candidates require analysis of hundreds of individual tissue specimens. Tissue-based molecular analysis has traditionally been dependent on the serial examinations of candidate genes expression at the DNA, messenger ribonucleic acid (mRNA) or protein level by FISH, chromagen based *in situ* hybridisation or immunohistochemistry performed on individual whole block tissue sections. The analysis of hundreds of individual biopsies is highly labour intensive. More over, inherent variability in experimental conditions from batch-to-batch of specimens analysed is unavoidable. A serious drawback is that human tissue resources are finite in amount, and typical paraffin tissue block will be exhausted after only a few hundred sections (Wang et al., 2002a). In response to the need for high-throughput analytical analysis of large number of tissue samples, a predecessor of Tissue Microarray technology was developed.

The term "multitissue" or "sausage block" is used to describe the first attempt to create high density tumour arrays. Battifora et al. (1986) described the "sausage block" methods in which 1 mm thick rods of tissue were wrapped carefully in a sheet of small intestine, this was then embedded in a paraffin wax block from which sections were cut (Battifora, 1986, Shergill et al., 2004) and stained for analysis. Later this technique was updated to the "checkerboard" configuration, which allowed the examination of multiple samples under a more uniform structure; however, it was limited by its inability to create individual rods (Battifora and Mehta, 1990). These original efforts were limited in the orientation and density of specimens. Only with the development of instruments that could place cores of tissue in a recipient block did the utility of a multitissue platform come about (Hewitt, 2006). In 1998, Kononen et al., published the utility of Tissue Microarrays for high-throughput molecular profiling of tumour specimens (Kononen et al., 1998). In contrast to previous efforts to create high density tumour assays, Tissue Microarrays had an extremely orderly and uniform structure which is amenable to automation, is more agreeable when performing reviews, and results in a higher order of magnitude of samples per slide than was previously possible.

1.4 Applications of Tissue Microarrays

Tissue Microarrays were originally developed for high-throughput molecular profiling of tumour specimens and confirmation of data from DNA microarrays (Kononen et al., 1998). Subsequently, Tissue Microarrays have expanded to function as a staring point for biomarker discovery and validation and in support of molecular epidemiology. In addition to biomarker validation, Tissue Microarrays have numerous applications and are extensively used in biomedical research. It has been suggested that, in the future, Tissue Microarrays will be used in clinical diagnostics as a cost-reduction method (Braunschweig et al., 2004).

The major applications of Tissue Microarrays can be broadly divided into the following categories:

1.4.1 Translational Research

To date the majority of analysis applied to Tissue Microarrays have been associated with clinical translational research, particularly the confirmation of tumour markers identified by gene expression profiling methods in association with cancer research (Wang et al., 2002a). Prior to the advent of Tissue Microarrays, the development of a potential biomarker was accomplished by repetitive validation on different cohorts of specimen to demonstrate its application, utilising whole section analysis. These different cohorts start with samples taken from the original microarray experiment and expand to samples in the same laboratory, to samples from different laboratories, until the biomarker is validated on a population based cohort (Hewitt, 2006). Tissue Microarrays offer the opportunity to profile the expression of a gene product in a broad range of tissues on a single slide (Braunschweig et al., 2004).

With the advent of targeted therapeutics, it is essential to have knowledge of associated expression pathways. Previous chemotherapeutic agents (for example cisplatin) were non specific and targeted replicating cells rather than non-replicating cells. Recent developments in chemotherapeutics agents have meant the developments of drugs that target particular proteins or pathways, for example Gleevac, Herceptin and Tarceva. Utilising predictive Tissue Microarrays it is

possible to investigate what tumour expresses a particular target, therefore aiding in constructing clinical trials (Braunschweig et al., 2004).

1.4.2   Quality Control

Tissue Microarrays can be used as important controls in clinical laboratories (Braunschweig et al., 2004). Tissue Microarray sections can be used as internal or external controls in tandem with surgical specimens to serve as positive and negative quality controls for diagnostic immunohistochemistry and *in situ* tests (Hoos et al., 2001). Therefore, Tissue Microarrays will have underwent similar conditions as the whole section and can act as "total assay" controls validating sensitivity and specificity of the antibodies utilised, the tissue fixation methods and the antigen retrieval methods (Moch et al., 2001, Gulmann et al., 2004).

Tissue Microarrays have been used to validate inter- and intra-laboratory quality control (Hoos et al., 2001). Miniature Tissue Microarrays have been incorporated on whole sections slides as a means of validating quality control in the assessment of HercepTest™, which is an immunohistochemical assay utilised to evaluate HER-2/neu overexpression in breast tumours (Gulmann et al., 2004). The inclusion of cell line preparations containing multiple samples of known HER-2 status, characterised by FISH and immunohistochemistry, is recommended in HER-2 assessment (Ellis et al., 2004).

Currently, HercepTest™ is the only Food and Drug Administration (FDA) approved immunohistochemical assay used to determine HER-2 protein overexpression in breast cancer tissue. The HercepTest™ kit includes cell line Tissue Microarrays, which are used to qualify the procedure and reagents quality. The inclusion of normal tissue within Tissue Microarrays also acts as internal controls, and can identify excessive antigen retrieval performed in immunohistochemical analysis (Ellis et al., 2004). Tissue Microarrays have also been used to evaluate inter-laboratory reproducibility when performing immunocytochemistry and FISH analysis to evaluate HER-2 overexpression and amplification (Kay et al., 2004).

1.5   Advantages of Tissue Microarrays

There are numerous advantages of using Tissue Microarrays compared with whole section analysis. The major benefits are as follows:

1.5.1   Standardisation

Tissue Microarray introduces standardisation of protocols into histopathology over and above what is possible with whole sections. As a result, Tissue Microarrays hugely reduce inter- and intra-laboratory variations due to standardisation of laboratory procedure (Tzankov et al., 2005). With whole section analysis inherent variability in experimental conditions from batch-to-batch of specimens analysed is unavoidable. However, all tissue specimens arrayed on one Tissue Microarray are analysed in an identical fashion. Antigen retrieval reagent concentrations, incubation times with primary and secondary antibodies, temperatures and wash conditions are identical for each core within a Tissue Microarray, resulting in an unprecedented level of standardisation which is unattainable utilising whole section techniques (Shergill et al., 2004).

1.5.2   High-throughput

Tissue Microarrays greatly increases throughput of tissue analysis. Analysis of prognostic and predictive markers had traditionally been performed by testing one marker at a time (Torhorst et al., 2001). However, utilising a single Tissue Microarray block containing 1,000 cores which can create 200 slides, as many as 200,000 individual assays can be performed (Shergill et al., 2004). Therefore, Tissue Microarrays allow serial selection analysis of multiple markers from the same molecular pathway in a large number of tissue samples, facilitating direct comparison of alterations of multiple molecular targets in virtually identical histologically highly conserved tumour regions, all of which are under uniform experimental conditions (Wang et al., 2002a).

1.5.3   Cost and time

Tissue Microarrays decrease the cost of reagents and the time taken to review the same number of patients when performed using whole section analysis. Cost and time are critical factors when comparing whole section and Tissue Microarray analysis. Milanes-Yearsley et al. (2002) extensively investigated the cost and time involved in creating, immunohistochemically staining and reviewing 4 Tissue Microarrays compared with 150 whole sections. Purchase of materials and reagents for immunohistochemistry was 37.1 times more expensive for 150 whole section slides compared with 4 Tissue Microarrays. Moreover, the greatest impact in reagent savings is typically seen with probes for ISH (Braunschweig et al., 2004). However, including technical time it is estimated that the cost of creating 4 replica Tissue Microarrays is $\geq$ \$5000, this cost is justified by the fact that potentially hundreds of Tissue Microarrays can be produced from one block.   The creation of Tissue Microarrays proved to be extremely labour intensive. To prepare, build and section 4 Tissue Microarrays consisting of 150 cores required 32 hours, in contrast with 8 hours required to section 150 slides from a single whole section block. However, to perform immunohistochemistry on 4 Tissue Microarrays required three hours, in comparison with 16 hours to perform immunohistochemistry on 150 whole tissue sections. The review of the 4 Tissue Microarrays required considerably less time than that required for 150 whole tissue sections. Immunohistochemical review required 1-2 minutes per case within the Tissue Microarrays, compared with 5-10 minutes per whole section tissue. Therefore, when comparing whole sections and Tissue Microarrays, histologists and pathologists spend more time in selecting blocks and constructing Tissue Microarrays compared to whole sections. However the technologist spend considerably less time immunohistochemically staining Tissue Microarrays and the pathologists spent less time reviewing Tissue Microarrays, compared to whole sections (Milanes-Yearsley et al., 2002).

It is not always cost effective to construct a Tissue Microarray when less than 25 specimens are involved. One exception to this model is the creation of cell line Tissue Microarrays where the inclusion of 5 or more cell lines can be justified. This is largely due to the fact that it is not necessary to review and stain a section of the

embedded cells before secting begins. Also the cost of the Tissue Microarray arrayer must be taken into account (Braunschweig et al., 2004).

1.5.4 Conservation of tissue

Human tissue for research purposes can be hard to obtain. Utilising Tissue Microarrays, sampling and combining tissue from different sources dramatically extend its utility. From a traditional block of tissue, between 300 and 500 whole section slides can be obtained. In comparison, if the same block can be repeated in between 5 and 50 Tissue Microarrays, results in upwards of 25,000 samples of that original block (Braunschweig et al., 2004). This "tissue expansion" capability of Tissue Microarray provides for the maximum conservation and utilisation of limited valuable, and often irreplaceable tissue (Wang et al., 2002a).

The availability of fresh tissue is often a constricting factor in genomic and proteomic research. However, large resources of FFPE tissue are usually catalogued in research hospitals. Therefore, retrospective studies with long follow-up information are possible. For example, the Radiation Therapy Oncology Group (National Cancer Institute, National Institute of Health, USA) have a resource of 12,262 blocks of prostate, brain, head and neck, gastrointestinal, gynaecological, bladder and lung tumours available to qualified investigators (Milanes-Yearsley et al., 2002). The small diameter of Tissue Microarray cores result in minimal damage to the original donor block which are often considered vital resources (Shergill et al., 2004). Therefore, subsequent to sampling, the remaining donor block can be retained for future whole section or Tissue Microarray analysis. This has resulted in pathologists giving samples more readily to researchers, whilst still retaining the tissue blocks for diagnostic purposes (Moch et al., 2001).

1.6   Limitations of Tissue Microarrays

Despite the numerous advantages of Tissue Microarrays, there are also difficulties associated with this technology. The majority of their limitations are associated with all types of histopathology applications, for example edge effect, fixation problems, loss of antigenicity, but these issues are often amplified when using Tissue Microarrays due to the small sampling size. The most prevalent limitations of Tissue Microarrays are as follows:

1.6.1   Tissue Heterogeneity

When Tissue Microarray technology was first created in 1998, it was expected that tissue heterogeneity would negatively influence the predictive power of Tissue Microarray results (Torhorst et al., 2001). Tissue Microarrays have inherently small tissue core samples which were suspected not to be fully representative of the heterogeneous parent tissue from which they were derived. How representative Tissue Microarrays are of the original specimen depends on the degree of heterogeneity of the tissue under investigation, the number of replica cores taken from the whole sections and the expression pattern of the marker involved (Wang et al., 2002a). It is also important to note that Tissue Microarrays have been developed to examine tumour populations and not individual tumours (Moch et al., 2001).

Torhorst et al. (2001) assessed the impact of tumour heterogeneity by comparing the frequency of ER and PR positivity in 553 whole section breast tissue samples with 4 Tissue Microarrays constructed from the same tissue. The frequency of ER positivity was virtually the same in the whole section and a single Tissue Microarray. By comparing a Tissue Microarray with the corresponding whole section agreement was slightly lower for PR (88.0%) and 3 x 0.6 mm replica cores were required to achieve the same level of positivity as the whole section.

Milanes-Yearsley et al. (2002) examined the correlation between whole sections and 4 replica multitumour Tissue Microarrays consisting of 150 x 0.6 mm cores for the expression of p53. Visual estimations illustrated that by comparing a Tissue Microarray with the corresponding whole sections, 86-90% agreement was achieved

and agreement increased to 97-100% when 4 Tissue Microarrays with duplicate cores were compared with the corresponding whole sections. The impact the type of tumour has on agreement was also assessed. Agreement was found to be 84.0% for brain tumour, 89.7% for prostate tumour, and 90.0% for head and neck tumours.

Hoos et al. (2001) illustrated that, with complex phenotypes, it is crucial to evaluate the number of cores required. Hoos et al. (2001) compared protein expression for Ki-67, p53 and retinoblastoma protein (pRB) for a group of 59 human fibroblastic tumours by whole sections and Tissue Microarray analysis. Tissue Microarrays were composed of 3 x 0.6 mm cores per specimen. Results indicated that tumour heterogeneity can lead to lower concordance rates if more complex phenotypes are analysed. Overall 3 cores per specimen resulted in 96.0-98.0% concordance with whole sections. However, concordance decreased to 91.0% for the more complex pRB phenotype.

Camp et al. (2000) investigated the phenotypic profiles of ER, PR and HER-2 expression of 38 breast carcinomas by immunohistochemically, comparing 1 to 10 0.6 mm cores per specimen with whole sections. Results illustrated greater than 95.0% concordance between Tissue Microarrays and whole sections when 2 cores per specimen were used, concordance increased to 99.5% with 5 cores per specimen.

Literature clearly indicates that Tissue Microarrays can sufficiently represent the donor tumours in order to establish associations between molecular alterations and clinical endpoints (Torhorst et al., 2001). The general opinion is that at least 3 x 0.6 mm cores are required from the donor tissue in order for Tissue Microarrays to be truly representative. However, due to the homogeneous character of cell line tissue pellets, one core per cell line is often enough within a Tissue Microarray constructed solely from cell lines. Due to the loss of cores during sectioning, it is recommended that cell lines are arrayed in duplicate form to avoid loss of the specimen (Hoos and Cordon-Cardo, 2001).

1.6.2 Fixation Problems

A major cause of variation in the reproducibility of immunohistochemical staining is induced by tissue fixation. Almost all paraffin-embedded tissue used in Tissue Microarrays is formalin fixed. Formalin-fixing is amongst the most popular, due to low cost, ease of preparation and because it preserves morphologic detail with few artifacts. In addition formalin-fixed tissue is also commonly used in histopathology laboratories. Neutral-buffered formalin (10%) which contains 4% formaldehyde is the most common type of formalin used in laboratories (Werner et al., 2000). Formaldehyde works under a broad variety of conditions, is stable, functions effectively over a five fold or more range of concentrations and is useable with almost any tissue. Formaldehyde is not a coagulate fixative; therefore, tissue does not contain clumps of coagulated materials nor is cellular detail distorted by the formation of coagulum (Fox et al., 1985). In some applications, 70% ethanol may be used rather than formalin, and this is more commonly used in cell lines, animal and xenograft arrays (Braunschweig et al., 2004).

The most important molecular change induced by formaldehyde is the formation of cross-links between proteins, or between protein and nucleic acids, involving hydroxymethylene bridges. Cross linkage by formaldehyde is a slow process and usually takes between 24 to 48 hours to be completed. Numerous problems are associated with incorrect fixation times and delayed fixation. Delayed fixation results in an increasing proteolytic degradation. Depending on the antigen, this may cause loss of immunoreactivity and therefore irreversible weak or absent immunohistochemical staining. Formalin fixation results in a variably reversible loss of immunoreactivity by its masking or damaging some epitopes. Epitope retrieval methods can be used to unmask the epitope. For some epitopes, the duration of formalin fixation is critical. Too short of a fixation time will lead to cross-linking only at the periphery of the tissue block; toward the centre of the block coagulation by alcohol during tissue dehydration may occur or the centre may remain raw and not fixed. This may cause more intense staining on the periphery or the centre depending on which antibody is used. Too long of a fixation time may result in excess cross-linkage formation which may lead to weak or absent staining (Werner et al., 2000).

1.6.3   Loss of Antigenicity

As described in section 1.12 sectioning and storing multiple Tissue Microarray sections at one time conserves tissue by avoiding the unnecessary loss of tissue which occurs every time a blocked is faced; however, this assumes that immunoreactivity is stable over time. Literature reports the loss of staining intensity in paraffin-embedded whole tissue sections over time, when probed for several antibodies (Wester et al., 2000, Bertheau et al., 1998, Jacobs et al., 1996). The decline of staining intensity within whole tissue sections and Tissue Microarrays prepared from paraffin sections may result in false negative findings. It is assumed that this problem may be accentuated in Tissue Microarrays, due to the limited tissue representation of each donor case (Fergenbaum et al., 2004).

Fergenbaum et al. (2004) assessed the loss of immunoreactivity of two Tissue Microarrays originating from a single Tissue Microarray block. The Tissue Microarrays were constructed from FFPE invasive breast carcinomas. A single Tissue Microarray was created and stored for 6 months before immunohistochemistry was performed. Another Tissue Microarray was created and immunohistochemically stained on the same day as creation, for PR, ER and HER-2. Protein expression was assessed by microscope-based reviews. The Tissue Microarray which was stored for 6 months before staining exhibited a loss in immunostaining for ER, PR and HER-2 compared with the Tissue Microarray sectioned and stained on the same day. The loss of immunostaining was at its greatest for the HER-2. In conjunction with increased length of storage time of sections, other factors are known to influence the rate of loss of antigenicity. Increased storage temperatures have been found to reduce antigenicity in samples stained for PR and HER-2 (Wester et al., 2000, Bertheau et al., 1998); however, reducing the cold storage does not eliminate loss of antigenicity. Nuclear and cytoplasmic specific stains are less affected by loss of antigenicity than membrane specific stains (van den Broek and van de Vijver, 2000). Oxidation may be a factor in loss of antigenicity. Currently the storage of unstained Tissue Microarrays under nitrogen and shipping in vacuum packed bags is being investigated as the optimal methods to reduce the loss of antigenicity. However, as with all immunohistochemistry, evaluation of staining intensity is a comparison factor and, as

long as all tissues under analysis have a uniform loss of antigenicity, analysis is still considered to be valid, as would be the case for tissue within Tissue Microarrays (Fergenbaum et al., 2004).

1.7   Commercial vs. Public Arrays

It is now possible to obtain "off the shelf" Tissue Microarrays from commercial vendors, academics, government sponsored or consortium based providers. However, commercial vendors are the most expensive, as the alternative providers charges are either nominal or to cover their costs, and often these providers obtain the tissue through donation (Fergenbaum et al., 2004). Commercial arrays charge by the slide and are prohibitively expensive for most applications. The cost of the arrayer is quite modest in comparison (Hoos et al., 2001). Generally commercial Tissue Microarrays offer larger cohorts of tissue but lack annotation and are limited in the number of samples that impact statistical significance. As the annotation provided increases so to does the cost of the Tissue Microarray. Frequently commercial arrays are limited to providing 5 to 10 slides per array block. If slides are re-requested at a later date it is assumed they will be from a different block. Commercially available Tissue Microarray composed of normal and tumour rodent model systems, cell lines, and xenografts are widely available. The general rule of thumb is, if a laboratory requires less than 20 Tissue Microarray slides a year, purchasing the slide would be the best option. If less than 100 slides a year are required, setting up collaboration with a more experienced laboratory is the best alternative. Non-commercial Tissue Microarray providers typically do not offer the same level of quality control, especially with regard to the design of the Tissue Microarray. However, there are often ethical concerns when purchasing a commercial Tissue Microarray. In some instances, institutional ethic review boards have required that they approve the use of Tissue Microarrays obtained from commercial sources even when these Tissue Microarrays have already received broad approval from the originating organizations (Hewitt, 2006).

1.8   Design of Tissue Microarrays

Tissue Microarray blocks should be designed in a logical fashion. The orientation of the specimens on the array is crucial because confusion about their location can threaten the evaluation of the experiment. Tissue Microarrays often include two orientation spots in specific positions outside the geometric margins of the usually square of rectangular arrays to orient the entire Tissue Microarray (Hoos and Cordon-Cardo, 2001).

The general guidelines for designing Tissue Microarrays vary greatly in literature. Some experts recommend randomised designs, whereby cores from different tumours and features are mixed together in a random fashion. Scattering normal tissue throughout the array to ensure even immunohistochemical staining has been reported. Also the use of different normal tissue to identify different rows based on the tissue morphology may be relevant. For example, distinct morphologic tissue such as bone, liver or skeletal muscle have been used (Hoos and Cordon-Cardo, 2001). However, other experts argue that these approaches can confuse pathologists, as some pathologists often rely on some aspects of histopathology to score the tumour, and would benefit from similar tumour located in close proximity to each other on the Tissue Microarray (Fergenbaum et al., 2004).

Tissue loss during sectioning and staining is a common problem with Tissue Microarrays. In addition "edge effect" which is the over staining of artifacts at the tissue borders is a well known occurrence in immunohistochemistry, which occurs most frequently on the periphery of Tissue Microarrays. To minimise the effect this may have on immunohistochemistry and the loss of tissue during sectioning it is recommended to frame the arrayed cancer specimens with rows of normal tissue (Hoos and Cordon-Cardo, 2001).

1.9   Types of Tissue Microarrays

Since Tissue Microarray development, numerous types of Tissue Microarrays have evolved. Tissue Microarrays types are dependant on their function and composition. Currently there are numerous types of Tissue Microarrays, outlined as follows:

1.9.1   Multitumour Tissue Microarrays

Tissue Microarrays can be composed of numerous tumour types, which allow large numbers of tumours to be screened simultaneously for molecular abnormalities of interest (Shergill et al., 2004). Multitumour Tissue Microarrays have been used for a range of applications.

1.9.2   Progression Tissue Microarrays

Morphological and molecular changes through the different stages of progression of one particular tumour type can be assessed in a tumour progression Tissue Microarray. Tumours can be extracted from one or multiple sources. An example of a prostate cancer progression Tissue Microarray would consist of normal prostate, benign prostatic hyperplasia, prostatic intraepithelial neoplasia and different stages of prostate cancer from localised disease to metastatic cancer (Shergill et al., 2004).

1.9.3   Prognostic Tissue Microarrays

Prognostic Tissue Microarrays are highly sought after, as they contain tumours from patients whose clinical outcome are known and, therefore, are used to identify potential prognostic biomarkers (Moch et al., 2001). The correlation of Tissue Microarray derived data with clinical follow-up data to assess prognosis is of significant importance to clinicians and their patients (Shergill et al., 2004).

1.9.4   Predictive Tissue Microarrays

Tissue Microarrays constructed from tumours "before and after" disease-specific treatment are used to investigate predictive markers with respect to response to a specific cancer therapy (Braunschweig et al., 2004).

1.9.5   Cell Microarrays

Tissue Microarrays can be constructed solely from cell lines (CMAs). CMAs are known as a "western on a slide", and are often used to test effects of new chemotherapeutics. Many methods exist for embedding the cell lines and for constructing the arrays; however, this is usually determined by the purpose of the CMA and the number of cells available. An example of a Tissue Microarray composed solely from cell lines is the NI60 created by TARP in NIH (Braunschweig et al., 2004).

1.9.6   Xenograft Tissue Microarrays

Xenograft tumours are routinely used in the construction of Tissue Microarrays and are referred to as XMAs (Hewitt, 2006). Optimal use of xenograft models for drug testing requires the use of panels of xenografts that closely mimic the biological characteristics of there respective primary tumours. Single xenograft models cannot capture the variability of the corresponding cancer. For paediatric cancers, xenograft model systems have been used for drug screening with some success. Due to the limited amount of tissues available for specific types of cancer in paediatric oncology, it is important to prioritise new agents for clinical trials (Whiteford et al., 2007). Utilising a panel of xenograft models to construct Tissue Microarrays it is possible to conserve precious tissue and perform large scale molecular characterisation.   XMAs have also been used in adult cancer analysis, such as prostate cancer (Glinsky et al., 2003, Bani et al., 2004).

1.9.7   Frozen Tissue Microarrays

It is appealing to many investigators to construct Tissue Microarray from frozen tissue rather than formalin-fixed-paraffin-embedded (FFPE) tissue, as it is possible to apply *in situ* hybridisation and antibodies for immunocytochemistry that do not work in FFPE material. Optimal cutting temperature (OCT) arrays or frozen arrays are used in cases where certain antibodies do not work in formalin-fixed-paraffin-embedded tissue. These arrays contain larger, but fewer, cores (Braunschweig et al., 2004).   However, the limitations of this approach usually outweigh the benefits. Frozen Tissue Microarrays are uneconomical with tissue requiring larger cores of

tissue, than FFPE Tissue Microarrays. The capacity to target selection of tissue can be impaired and the histology on the resultant arrays is typically of limited quality (Hewitt, 2006).

1.9.8   Bone marrow Tissue Microarrays

Until recently very little information was available about the applicability of the Tissue Microarrays for the analysis of bone marrow biopsies, because of the specific difficulties in constructing arrays generated from small tissue cylinders of bone structure. Recent studies have shown the construction of Bone marrow Tissue Microarray (BM-TMAs) is hampered by the presence of bony trabeculae in the bone marrow biopsies and the heterogeneity of the samples used. This problem was reduced by firstly carefully making representative areas rich in leukaemic blasts in the bone marrow biopsies selected for Tissue Microarrays construction and, secondly, by punching these selected areas exactly. Therefore, the punched bone marrow tissue very rarely turned out to be non-representative, due to necrotic or normal non-neoplastic bone marrow tissue (< 5%) (Zimpfer et al., 2007).

1.10   Selection of Tissue

A major factor in the loss of satisfactory samples is the quality of tissue utilised. Tissue Microarrays can be no better than the tissue they are built from, and often are not as good as the original tissue (Hewitt, 2006). The majority of human tissues used in Tissue Microarrays are usually sourced from surgical specimens, either from removed organs, biopsies or autopsies. The use of autopsy material in Tissue Microarrays is limited due to the significant risk of anoxic or autolytic defects which can impact immunohistochemistry or ISH results. Autopsy tissue is usually only used when there is no alternative, for example, in the case of normal human brain tissue (Braunschweig et al., 2004). The quality of the tissue must be noted, for example prostatic carcinoma preoperatively treated with radiation may consist only of small numbers of remaining tumour cells (Hoos and Cordon-Cardo, 2001).

In addition to the original quality of the tissue when removed from the body, the subsequent handling of the tissue also alters the quality. The tissue should be frozen or fixed directly (See section 1.6.2) after surgery in order to ensure optimal protein or nucleic acid preservation (Braunschweig et al., 2004). If optimal fixation is not performed tissue quality can be hugely affected. If donor blocks are over-fixed, the tissue tends to be dry and run a greater risk of cracking and breaking into fragments than optimally fixed tissue (Hoos and Cordon-Cardo, 2001).

Another consideration when selecting human or animal tissue is the ethical regulation. Restrictions vary greatly depending on the type of tissue used and the source of tissue. For human tissue, the ethical regulations vary greatly from country to country. Animal tissue has been widely used for the construction of Tissue Microarrays, (including mouse and rat) as well as other animals used as models of human disease such as canine models of human cancer (Hewitt, 2006). For animal tissue, it is essential that the protocols for handling the animals are followed. Tissue Microarray are commonly constructed from anonymized tissue, which usually lacks clinical follow-up data. Tissue Microarrays composed of anonymized tissue are typically used as the tissue is less expensive to purchase, approval from the ethical committees is easier, and approval from the ethical offices of the individual end users

is not required. Tissue Microarrays with clinical and epidemiology information greatly expand the potential benefits of the Tissue Microarray; however, the cost of their creation is great and this type of Tissue Microarray is usually not commercially available (Braunschweig et al., 2004).

1.11   Construction of Tissue Microarray Blocks

Construction of a high quality Tissue Microarray is not trivial. The cost of staff, time and equipment to construct Tissue Microarrays may suggest that collaboration is the best solution (Hewitt, 2006). The preparation for constructing Tissue Microarray blocks is extremely labour intensive. The majority of time taken to create a Tissue Microarray is spent collating tissue, reviewing biopsies, and selecting areas of interest for arraying. The actual tissue arraying processing itself is quite simple and is considered < 5% of the entire work of creating Tissue Microarrays (Simon et al., 2003). Haematoxylin and Eosin (H&E) stained full sections from the biopsy blocks must be obtained to assess morphology and to identify an area that represents the specimen; however, this can vary greatly among tumour types. For example, in colorectal cancer it is important that the target areas are small and well defined because stromal areas between the glandular structures of the tumour can be large and a core biopsy can easily miss the tumour cell regions (Hoos and Cordon-Cardo, 2001).

Construction of a Tissue Microarray requires access to specialised equipment and training (Hewitt, 2006). Low density arrays may be constructed from simple equipment, for example using spinal needles (Hidalgo et al., 2003) or a drill grinder (Vogel and Bueltmann, 2006). However, the majority of Tissue Microarrays are constructed using a manual tissue arrayer. Currently, Beecher Instruments are the leaders in the field for providing Tissue Microarray arrayers. They provide three generations of machines for constructing Tissue Microarrays, one automated machine (ATA-27) and two manual machines (MTA-I, MTA-II). MTA-I consists of a recipient block holder, bridge holder for the donor block, matched pair of donor and recipient needles available as either 0.6, 1.0, 1.5, or 2.0 mm diameter, an adjustable arm which houses the needles, and two digital micrometers used for positioning. MTA-I is illustrated in Figure 1.11-1.
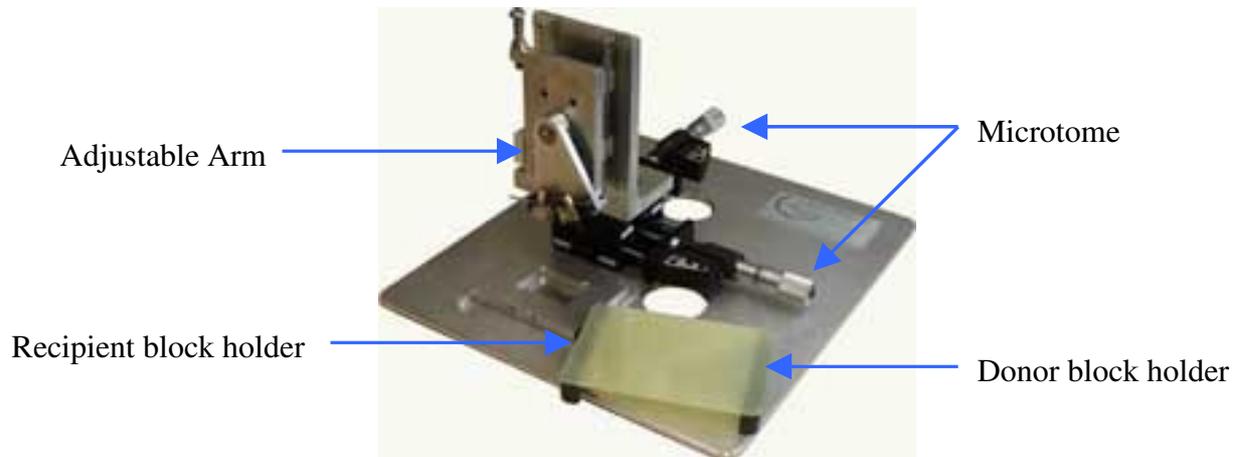
Figure 1.11-1: MTA-1- First version of Tissue Microarrayer. Adjustable arm, controls the movement of the needles. Recipient block holders houses the Tissue Microarray block. Donor block holder houses the biopsy tissue. Microtome adjusts the x-y positing of the needles. Image provided by Beecher Instruments.

Instrumentation for the construction of Tissue Microarrays is beginning to diversify, ranging from the introduction of automated arrayers to supplementary applications to assist in the construction process. MTABooster® designed by ALPHELYS, is a motorisation kit that can be retrospectively installed on the MTA-1 in order to automate the X-Y movement of the stage (ALPHELYS, France). During manual construction of tissue arrays, rotation of the micrometers may be required up to 2000 times. Utilising MTABooster® the movement of the micrometers can be automated and visualised by a LED screen. The application of fully-automated arrayers such as Beecher's the ATA-27 is largely limited to industrial and high-volume production facilities (Hewitt, 2004). An alternative to Beecher Instruments is the TMArrayer™ which is a semi-automated arrayer provided by Pathology Devices (Westminster, USA). When purchasing an arrayer cost, ease of use, speed and accuracy are the typical considerations.

The process of arraying involves removing a cylindrical core of paraffin from the recipient block. The needle used to remove the paraffin has a slightly smaller diameter than the second needle (used to remove the donor tissue). An area of interest is selected for sampling using H&E stained sections of the donor biopsy. A

cylindrical core of tissue is removed from the biopsy using the second needle, the adjustable stage is removed to reveal the recipient block directly underneath. The needle is lowered into the cylindrical cavity and the tissue is expelled. The process is visualised in Figure 1.11-2. It is critical to place every core the same depth into the recipient block. If the donor blocks are not of the same depth double stacking may be required. Paraffin must be soft enough at room temperature to punch holes without cracking the block. After the recipient block is created, it is placed in the oven at 37°C for several hours to soften the paraffin and to compound the cylinders with the paraffin, this process is known as "tempering" (Fergenbaum et al., 2004).
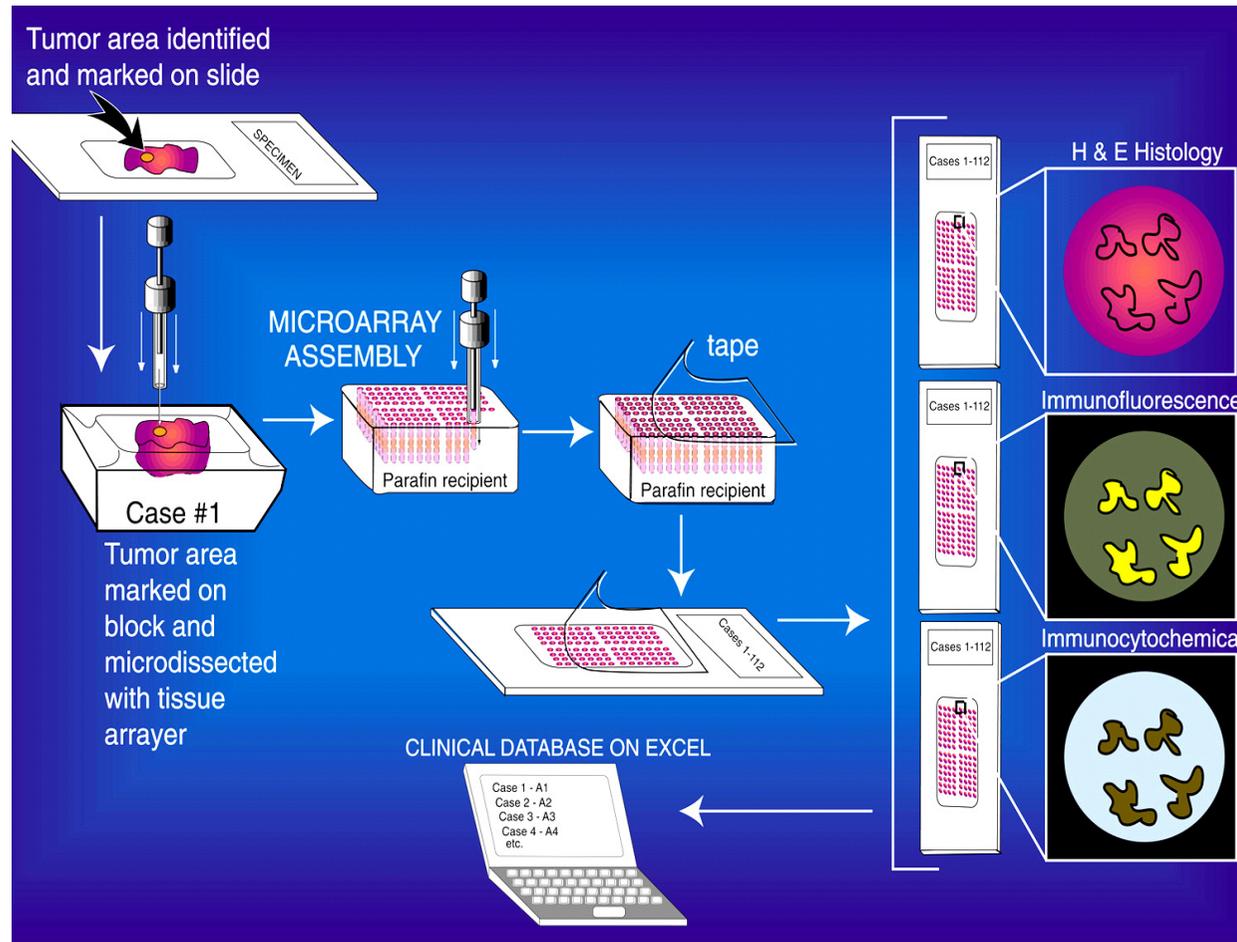
Figure 1.11-2: Tissue Microarray construction (image created by Mark A Rubin MD, Michigan University).

The size of the needles utilised in Tissue Microarray construction will affect the number of samples present within the Tissue Microarray. When designing a Tissue Microarray, it is essential that the ratio of needle size to the number of samples is carefully considered. In general, a Tissue Microarray recipient block is 18-22 mm in size, which can easily hold up to 500 0.6 mm diameter cores. As the diameter of the needles used increase, the distance between the cores within the block should also increase. General guidelines are available for designing a Tissue Microarray. Distances between the centres of cores must be uniform. Sub-arrays of samples within the Tissue Microarray block can be used in order to prevent reviewers loosing their position when performing analysis. Sub-arrays of 20-25 cores is considered a workable group (Braunschweig et al., 2004). An orientation core should be present within the block. Generally between 5-10% of all cores present should be normal or control tissue, especially where immunohistochemistry will be performed on the sides (Hewitt, 2004). The number of duplicate tissue depends on the tissue type and is greatly affected by the issue of tissue heterogeneity.

Another consideration is the amount of donor tissue available for sampling. There are guidelines for the amount of cores that can be sampled from a 1 cm^3 donor tissue block. For example, using 0.6 mm needles 256 cores can be sampled, using 1.0 mm needles 100 cores are possible. The general rule is to always fill the array; the more samples the better and in the absence of multiple samples, use larger cores (Braunschweig et al., 2004).

1.12 Sectioning of Tissue Microarray Block

Sectioning of Tissue Microarray blocks is performed utilising a microtome. It is possible to cut 200 sections from a single Tissue Microarray block (Kononen et al., 1998). However, once more then 70 sections are cut, missing tissue spots may become apparent. Sectioning a Tissue Microarray block at different times requires the block to be refaced, which results in loss of tissue (Milanes-Yearsley et al., 2002). Approximately 5% loss of the tissue occurs each time the block is re-inserted into the microtome (Wang et al., 2002a). Therefore, it is often preferable to section a block to its entirety in one sitting. It has been reported that between 17.6 - 20.0% of cases lacked a satisfactory core when the Tissue Microarray was refaced and sectioned a second time, this occurred 50 sections deep into a breast cancer block. However, only 3.2 - 4.8% of cases lacked a satisfactory core when Tissue Microarrays were sectioned at one time (Fergenbaum et al., 2004). To avoid unnecessary loss of tissue, microtome techniques must be perfected and performed by an experienced individual.

1.13 Creation of Tissue Microarray slides

Due to arraying, transferring of sections, vigorous staining procedures and the depth of the donor tissue samples are often missing within Tissue Microarrays. Loss of tissue can be minimised by including multiple samples from a specimen, inclusion of good quality tissue and maintaining uniform spacing between cores. It has been reported that between 10 - 30% of cores have been lost due to poor tissue transfer to the slides (Hoos and Cordon-Cardo, 2001). The challenge in transferring sections to slides is obtaining tissue that is not twisted, stretched or wrinkled and resistant to staining methods including harsh antigen retrieval procedures (Fergenbaum et al., 2004).

Currently two methods of transferring tissue to slides are utilised;

1.13.1 Adhesive Tape Transfer

When dealing with large Tissue Microarrays some experts recommend the use of tape transfer technique to transfer sections from the microtome to glass slides. An adhesive coated slides (Instrumedics, USA) system can be used to assist in tissue sectioning (Simon et al., 2003). With the tape transfer method, adhesive tape is placed directly onto the block within the microtome prior to sectioning. After sectioning the tissue remains on the tape, the tape and tissue is then transferred onto a slide coated in artificial resin. The slide is then placed under UV light in order to harden the resin and ensure the tissue transfers to the slide. The slides are then placed in solvent to remove the tape (Fergenbaum et al., 2004). Although this method is time-consuming, it has been shown to reduce tissue loss to less than 2% (Wang et al., 2002a).

1.13.2 Water Bath Transfer

Utilising water bath transfer achieves less tissue damage, less background staining and there are fewer missed cores on smaller Tissue Microarrays than when using adhesive tape transfer technique (Milanes-Yearsley et al., 2002). Using the water bath transfer technique, the section of tissue is taken from the microtome and placed on the surface of warm water; this evens the wrinkles which develop during sectioning on the microtome. A glass slide is then used to lift the section from the warm water. The slide is then allowed to dry, either in an oven or by air drying. The tissue is then adhered to the glass slide. In order to perform immunohistochemistry on the tissue, the adhesion of the tissue to the slide must be enhanced. Commercially available coated slides, which have been previously coated with organic polymers are used i.e. poly-L-lysine (Fergenbaum et al., 2004). Using water bath transfer, up to 70 sections from the Tissue Microarray have been removed before missing samples emerged (Hoos and Cordon-Cardo, 2001). However, large Tissue Microarrays are often twisted and lose their shape when using water bath transfer method.

1.14 Experimental usage of Tissue Microarray sections

Tissue Microarrays are suitable for all analyses currently performed on conventional whole biopsy sections (Kononen et al., 1998). Three assays are predominantly performed on Tissue Microarrays, they are as follows:

1.14.1 Fluorescent *in situ* hybridisation (FISH)

FISH is a molecular cytogenetic tool that permits detection of specific chromosomal aberrations in metaphase chromosomes and interphase nuclei. The process involves the samples DNA being denatured, the fluorescently-labelled probe of interest is added to the denatured sample mixture and it hybridises with the sample DNA at the target site as it reanneals back into a double helix. The probe signal can then be seen through a fluorescence microscope and the sample DNA scored for the presence or absence of the signal. FISH technology has been widely used in detection of chromosomal copy number changes, amplifications, deletions, and rearrangements both in research and clinical settings. The quality of FISH is influenced by several parameters including the quality of the probe and the accessibility of the hybridisation targets (Andersen et al., 2001).

The use of FISH technique on archival FFPE sections has been especially problematic, primarily because the formalin fixation greatly affects the accessibility of the hybridisation target. In a Tissue Microarray containing hundreds of different specimens the age of the individual samples as well as the type and length of the fixation used can vary considerably. Improved procedures for performing FISH on FFPE Tissue Microarrays have been developed and the new protocol results in greatly increased signal intensity and almost 30% increase in the number of tissue samples with invaluable hybridisation results (Andersen et al., 2001). There are challenges of FISH on Tissue Microarrays as the imaging methodologies require high magnification and counting many cells within each core to detect the specific targets and as the nature of tissue in histologic section does not result in every cell possessing an intact nucleus. However,

the significant savings in fluorescent probes continues to make Tissue Microarrays extremely inviting for investigators using FISH (Braunschweig et al., 2004).

1.14.2 *In situ* hybridisation

In situ hybridisation for RNA expression is used to characterise spatial and temporal response of specific gene targets in tissue sections. Probes were traditionally labelled with radioactive nucleotides and required placement of a photo emulsion on the slide after hybridisation. Development of the emulsion and visualisation by dark field microscopy required two weeks. Alternatively, the slide can be placed on a phosphorimager plate and the signal quantified. Chromogenic *in situ* hybridisation (CISH) using digoxigenin-labelled riboprobes, has been used for detection of amplified RNA and DNA viral sequences for some time but new protocols incorporating tyramide signal amplification (TSA) have recently come to the market. Others have used fluorescence probes (Braunschweig et al., 2004). The basic technique of ISH for RNA expression has always been challenging and requires meticulous technique for reproducibility. Due to manual time-consuming and tedious protocols, the application of ISH for RNA expression is limited.

1.14.3 Immunohistochemistry

The most commonly applied analysis to Tissue Microarrays is immunohistochemistry, approximately 80% of all Tissue Microarrays are analysed in this way (Braunschweig et al., 2004, Shergill et al., 2004). Immunohistochemistry has a well established role in routine pathology. Immunohistochemistry is a technique that allows the identification and localisation of a cell-bound antigen of interest; for example, a specific protein, lipid or carbohydrate molecule in a cell of tissue preparation detected by means of a specific antigen-antibody reaction, which is identified by a visible label that can be evaluated using microscopy. Immunohistochemistry is an extremely versatile technique as there is no limit to the range of molecules that can be localised in this way in cells and tissue, as long as suitable antibodies that recognise them are available (Brooks, 2005). It is recommended that, when possible, immunohistochemistry be the assay of choice for

biomarker validation (Hewitt, 2006). The main application of immunohistochemistry is the detection of differentiation antigens for classifying undifferentiated tumours, lymphomas, neuroendocrine and soft tissue tumours. In addition to antibody quality, tissue fixation, unmasking of epitopes and sensitivity of the detection system have a major affect on immunohistochemistry quality (Werner et al., 2000).

Detection methods for immunohistochemistry are broken down into two primary groups, chromogenic, so called "brown stains" and fluorescent detection methods. Detection occurs by the attachment of the oxidising colour reagent or fluorescent tag to the antibody-antigen complex (Braunschweig et al., 2004). Fluorescent detection methods offer increased dynamic range of staining and the capacity to perform multiple stains at one time; however, they are not as popular as chromogenic stains within FFPE tissue. Fluorescence in immunohistochemistry of Tissue Microarrays it is not widely used as it requires specialised equipment, expertise in analysis beyond that of the typical histopathologist and can be troubled by autofluorescence. Chromogenic stains are the most common approach in immunohistochemical analysis of Tissue Microarrays within FFPE tissue. Interpretation requires a standard light microscope with results that are relatively permanent. Contrast agents imparted by histochemical stains allow easy recognition of the staining pattern at both cytologic and histologic level (Hewitt, 2006).

The repetitive nature of immunohistochemistry is clearly amenable to automation. Automated staining minimise variations in the colour spectra of the chromogen produced by differences in incubation time and reagent amounts (Johansson et al., 2001). Instruments are available that regulate the temperature of slide incubation and allow the automation of complete immunohistochemical procedure including de-paraffinisation and antigen retrieval (Ventana Medical Systems; Vision BioSystems). Antigen retrieval is often the most critical factor to the success of an immunohistochemical procedure; therefore, these instruments provide a significant advantage over manual methods in terms of standardisation (Warford et al., 2004). Several instruments that automate immunohistochemical staining are available (Biogenex; DakoCytomation; Lab Vision). Properly used, they increase reproducibility

and can produce audit trails that are particularly useful in the controlled environment of the high-throughput laboratory. They also offer the possibility of continuous working beyond normal working hours. The purchase and running costs of automated immunostainers are high. However, in comparison with manual immunohistochemistry, this cost needs to be set against not only reproducibility, throughput and quality assurance issues, but also the freeing of laboratory staff from often long and very repetitive methods (Warford et al., 2004).

Immunohistochemically stained Tissue Microarrays provide a degree of consistency and standardisation that is not possible when staining multiple whole sections and, therefore, its has been suggested that due to technical difficulties of immunohistochemistry, data obtained from Tissue Microarrays can be superior to large section data, where immunohistochemistry is involved (Torhorst et al., 2001). However, all the challenges that are associated with immunohistochemistry on full face sections are also found with Tissue Microarrays and are often magnified by the small sample size of the tissue, and the diversity of fixation and processing conditions of tissue originating from different sources. Immunohistochemistry is not more challenging on Tissue Microarrays compared to whole sections; however, due to tissue originating from different sources, Tissue Microarray immunohistochemistry is more likely to unmask deficient protocols. Also, as previously mentioned, Tissue Microarrays are susceptible to loss of tissue during antigen retrieval and "edge effect" staining, as is the case with any small fragment of tissue present on a glass slide during immunohistochemistry (Braunschweig et al., 2004).

1.15 Microscope-based assessment of immunohistochemically stained Tissue Microarrays

Histopathology remains the gold standard for most diagnosis and therapeutic decisions in surgical and autopsy pathology. The interpretation of histologic sections; however, is an inherently subjective process based primarily on morphologic features (Cregger et al., 2006). The bulk of cases usually lie between where the research scientists can interpret the data; however, the quality of interpretation would improve with consultation by a pathologist (Hewitt, 2006). The general parameters of data recorded during assessment of immunohistochemically stained tissues using traditional microscopes are, intensity, the localisation, and the proportion of cells of interest that meet the first two criteria.

Human assessment of immunohistochemically stained Tissue Microarrays is considered the "gold standard" method of evaluation. Analysis of immunohistochemistry is considerably easier on Tissue Microarrays compared to whole sections, due to the fact that it is possible to compare staining intensities from different specimens on the same Tissue Microarray. More importantly, interpretation is limited to within a small predefined area. Therefore, the area under investigation is constant for all reviewers, unlike whole sections where different reviewers will select different areas of importance.

1.15.1 Observer variability in Immunohistochemistry

Observer variability can exist in three instances, Inter-observer variability (variability between observers), Intra-observer variability (variability by the same observer) and Inter-laboratory variability (variability between observers in different institutions). Poor inter-laboratory agreement is usually attributed to variability in tissue fixation, tissue processing, immunohistochemical protocols, antibodies and scoring systems used in different laboratories (Lacroix-Triki et al., 2006). Intra-observer variability has been reported as being less frequent than inter-observer variability. It has been suggested that each pathologist adhere to their own internal standards which in some cases, appear to be consistently reproducible (Kay et al., 1994). Inter-observer variability in relation to

microscope-based reviews of immunohistochemically stained tissue has been well documented in literature, especially for the assessment of HER-2.

Tsuada et al. (2002) examined the inter-observer variability of 6 pathologists for the assessment of HER-2 expression. One hundred and six whole sections breast carcinomas were immunohistochemically stained. The HercepTest™ scoring system, which traditionally quantifies staining using 4 categories, was further subdivided into two categories. Results from the HercepTest™ illustrated that when comparing 3+ to 0/1+/2+ a higher percentage of agreement was achieved than when comparing 0/1+ to 2+/3+. When cases were recorded as 3+, complete inter-observer agreement was almost achieved. Conversely, when cases were recorded as 2+ low level of inter-observer agreement was achieved (Tsuda et al., 2002).

As previously mentioned, numerous factors are attributed to influencing human interpretation of immunohistochemically stained tissue, and therefore introducing inter- and intra-observer variability. Human assessments can accurately and consistently identify the presence or absence of disease and low or high staining intensity. However, human assessment is not as capable when utilising intermediate categories and huge amount of variation is introduced as a result of over-using the intermediate category available during reviews (Kay et al., 1994).

Inter-observer variability, when performing tumour identification, is hugely dependant on the type of tumour assessed, the antibody under assessment and the standard criteria available to identify the tumour in question. With the use of standardised criteria, inter-observer agreement when assessing tumour presence can be achieved (Schnitt et al., 1992, Wei et al., 2004). It has also been suggested that Tissue Microarrays reduce intra- and inter-observer variability, due to the fact that a cohort of samples are analysed in a single review seating whereas traditionally this would have involved multiple review seating's (Tzankov et al., 2005, Zu et al., 2005).

1.15.2 Inherent Problems Associated with Human Vision

Observer variability is the greatest problem associated with human-based microscope assessment. The reason this occurs is that the accuracy of human vision is highly variable from person to person. Medical diagnostic reasoning is the process of assembling evidence to support the identification of diseases. It involves diverse cognitive activities including, information gathering, pattern recognition, problem solving, decision making, and judgment under uncertainty (Johnston, 2005a). Vision allows humans to perceive and understand the world surrounding them. If an image is to be analysed by a human, the information should be expressed using variables which are easy to perceive. These include psycho-physical parameters such as contrast, brightness, shape, texture and colour. Humans will find objects in images only if they may be distinguished effortlessly from the background. The replication of human vision would be relatively easy if the human visual system has a linear response to composite input stimuli; for example, a simple sum of individual stimuli. In fact, the sensitivity of human senses is roughly logarithmically proportional to the intensity of an input signal. Human perception of images provokes many illusions (Sonka, 1998).

Human vision is a complex process. Light enters the eye through the pupil; it passes through the lens and is projected on the retina at the back of the eye. Extraocular muscles move the eyeball in orbits and allow the image to be focused on the central retinal or fovea. The retina is a mosaic of two basic types of photoreceptors: rods, and cones. Rods are sensitive to blue-green light with peak sensitivity at a wavelength of 498 nm, and are used for vision under dark or dim conditions. There are three types of cones that give basic colour vision: L-cones (red) with a peak sensitivity of 564 nm, M-cones (green) with a peak sensitivity of 534 nm, and S-cones (blue) with a peak sensitivity of 420 nm. Cones are highly concentrated in a region near the centre of the retina called the fovea region. The maximum concentration of cones is roughly 180,000 per square mm in the fovea region and this density decreases rapidly outside of the fovea to a value of less than 5,000 per square mm. The standard definition of normal visual acuity (20/20 vision) is the ability to resolve a spatial pattern separated by a visual angle of one minute of arc. Since one degree contains sixty minutes, a visual angle of one minute of arc is

1/60 of a degree. The spatial resolution limit is derived from the fact that one degree of a scene is projected across 288 µm of the retina by the eye's lens. In this 288 µm, there are 120 colour-sensing cone cells. Thus, if more than 120 alternating white and black lines are crowded side-by-side in a single degree of viewing space, they will appear as a single grey mass to the human eye.

1.15.2.1 Acuity

Visual acuity is the ability of the eye to see fine detail. The human eye is less sensitive to slow and fast changes in brightness in the image plane, but is more sensitive to intermediate changes (Sonka, 1998). The eye has a visual acuity threshold below which an object will go undetected. This threshold varies from person to person. A number of factors affect visual acuity such as refractive error, size of the pupil, time of exposure of the target, eye movement, illumination and contrast (John Moran Eye Centre University of Utah, 2005).

1.15.2.2 Contrast

Contrast is the local change in brightness and is defined as the ratio between average brightness of an object and the background brightness. The human eye is logarithmically sensitive to brightness, implying that, for the same perception, higher brightness requires higher contrast. Apparent brightness depends very much on the brightness of the local background; this effect is called "conditional contrast". Figure 1.15-1 illustrates the fallibility of human perception of contrast. In Figure 1.15-1 A & B the internal squares are the same brightness; however, each internal square is surrounding borders of different brightness. As a result, the brightness of the internal boxes appear different (Sonka, 1998).
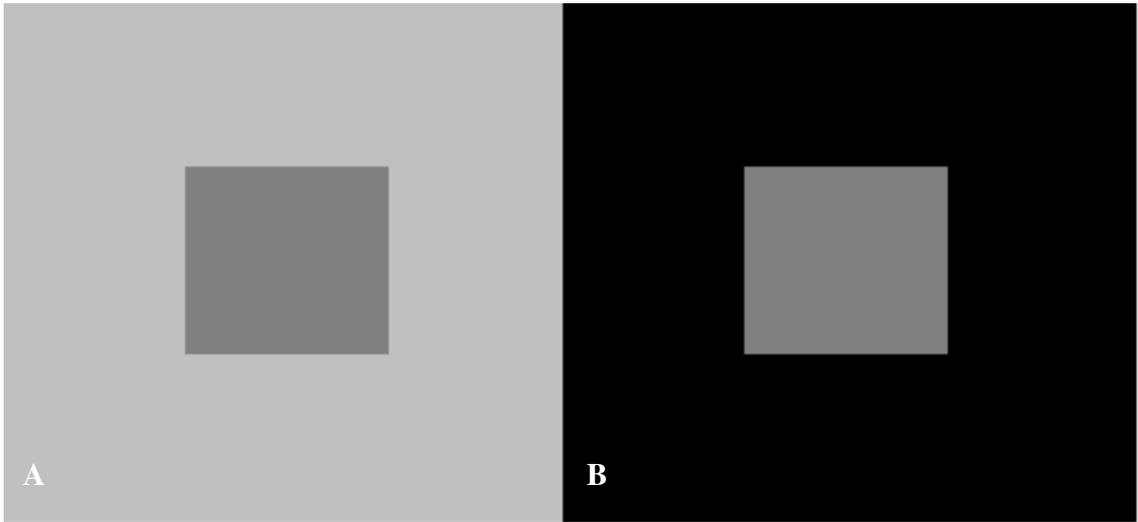
Figure 1.15-1: Internal squares are an identical shade; however, when surrounded by different colours the internal squares appear to be different shades.

Figure 1.15-2 illustrates a vertical bar with a single colour throughout. When viewed with the contrast of a white background it is quite obvious the vertical bar is a single colour.

Figure 1.15-2: Single coloured bar against white background.

Figure 1.15-3 illustrates the identical vertical bar as in Figure 1.15-2. However, in Figure 1.15-3 the vertical bar is surrounded by a background with a changing gradient of colour. As a result, the vertical bar no longer looks the same colour throughput.
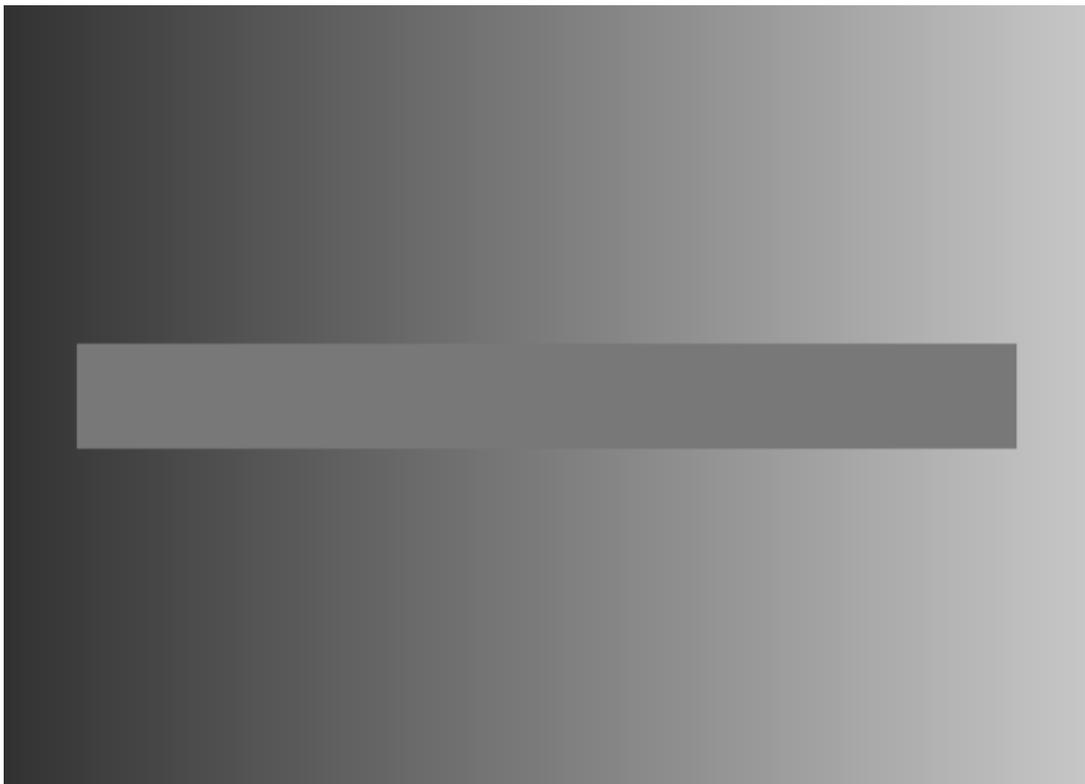
Figure 1.15-3: Single coloured bar against colour gradient background.

Contrast is extremely applicable in the assessment of membrane-bound immunohistochemical staining. The contrast between membrane and cytoplasmic staining may be hugely variable and can affect human perception. Figure 1.15-4 illustrates two images of bladder tissue probed with the antibody for E-cadherin. Both A and B within the image have strong membrane-bound immunohistochemical staining. Image A; however, has no cytoplasmic staining and image B is flooded with cytoplasmic staining. Within image B it is exceptionally hard to identify the membrane regions and to decipher the intensity of membrane staining only, without taking into account the cytoplasmic staining intensity.
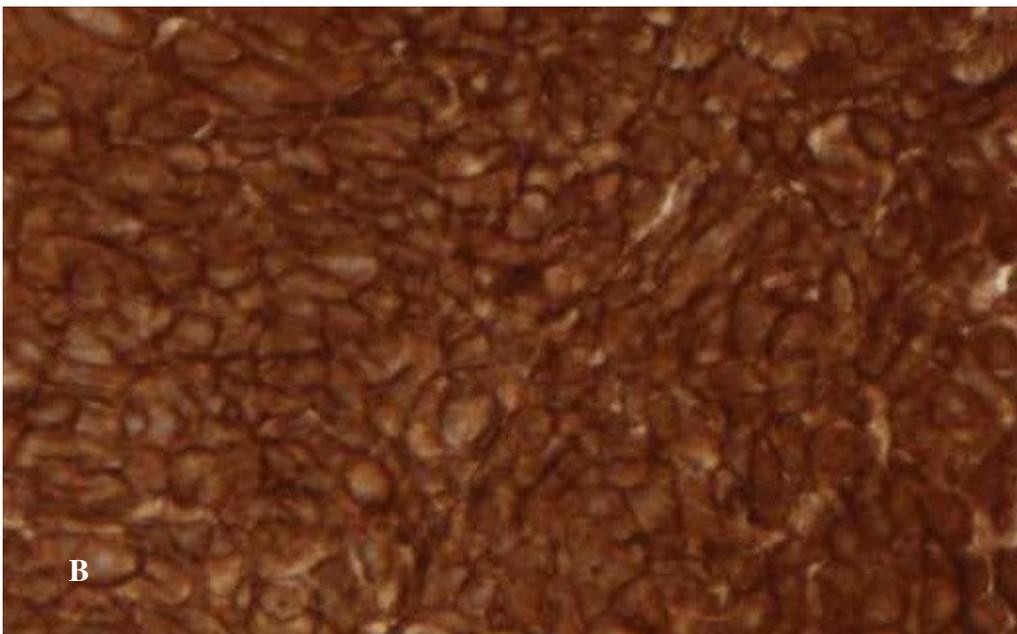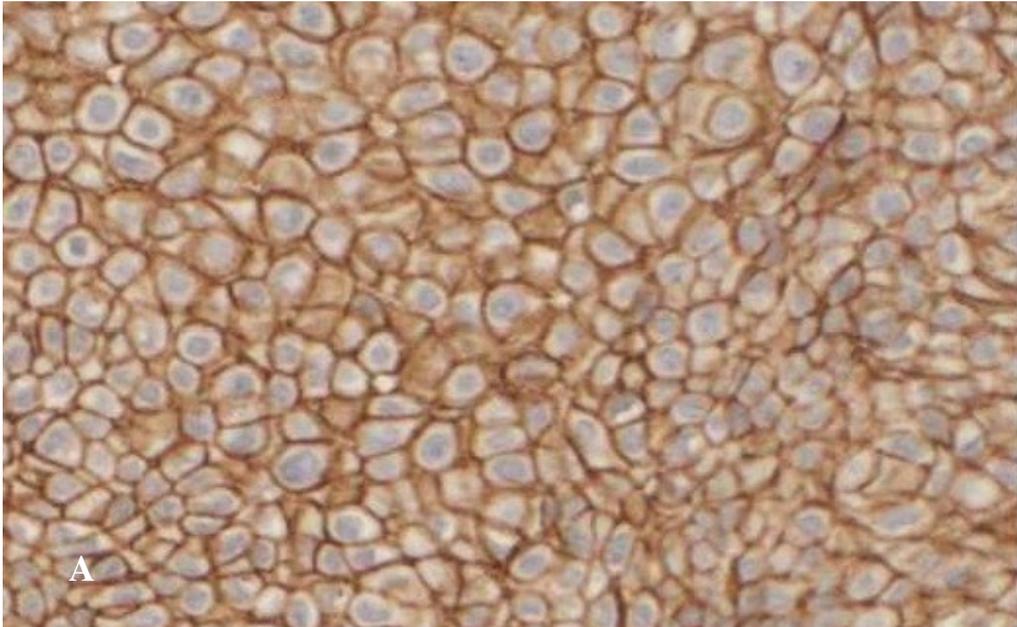
Figure 1.15-4: Image A and B have strong membrane-bound immunohistochemical staining; however, Image A has no cytoplasmic staining, and Image B is flooded with cytoplasmic staining, therefore making subjective membrane assessment very difficult.

1.15.2.3 Object border

Object borders carry a lot of information. Boundaries of objects and simple patterns such as blobs or lines enable adaptation effects similar to "conditional contrast". The Ebbinghaus illusion illustrates how humans can misinterpret size of particles when displayed in relative comparisons (Plodowski and Jackson, 2001). Figure 1.15-5 (A & B) displays two circles of the same diameters; however, as they are surrounded by circles of different diameters they appear to have different diameters (Sonka, 1998).

(A)　　　　　　　**(B)**
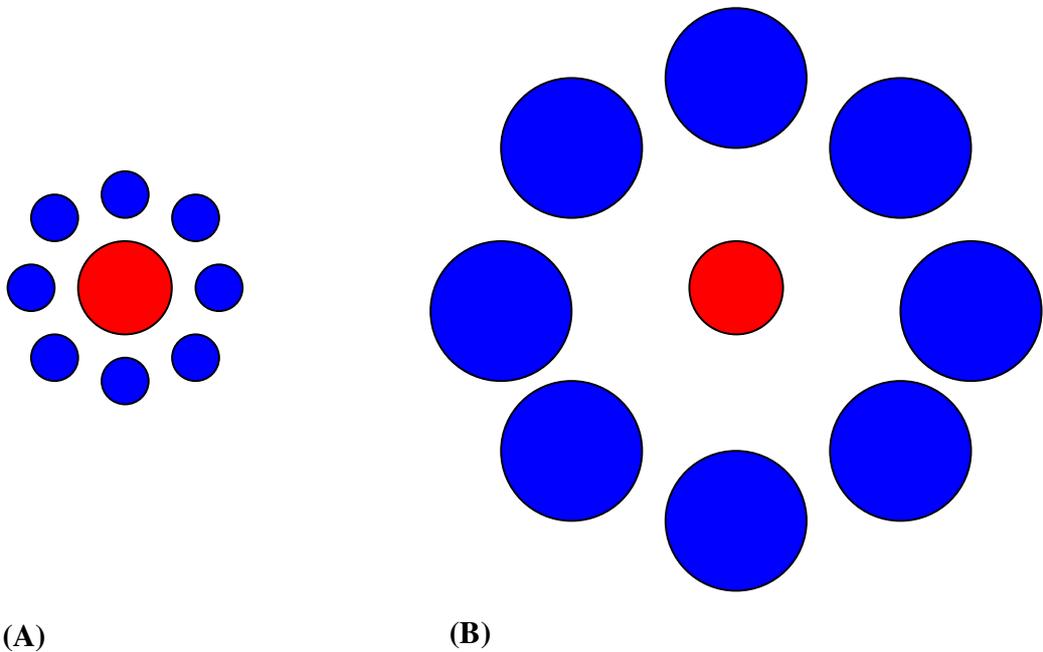
Figure 1.15-5: Ebbinghaus illusion illustrates how the interpretations of the size of objects are relative to their surroundings. The red circle within image (A) and (B) are identical; however, perception of the size of the red circle is altered by the blue circles surrounding them.

1.15.2.4 Colour

Colour is connected with the ability of objects to reflect electromagnetic waves of different wavelengths; the chromatic spectrum spans the electromagnetic spectrum from approx 400 to 700 nm. Humans detect colours as combinations of the primary colours red, green, and blue. Colour is very important for perception, since under normal illumination conditions the human eye is more sensitive to colour than to brightness. Colour quantification and representation can be expressed as a combination of the three basic components, red, green and blue (RGB), but colour perception is better expressed in the alternative Hue Saturation and luminosity (HSL) co-ordinate system. The Bezold Effect is used to describe how colours appear differently depending on their relationship to other colours. Figure 1.15-6 illustrates that the colour red appears lighter when it is surrounded by a white border, and darker when surrounded by a black border.
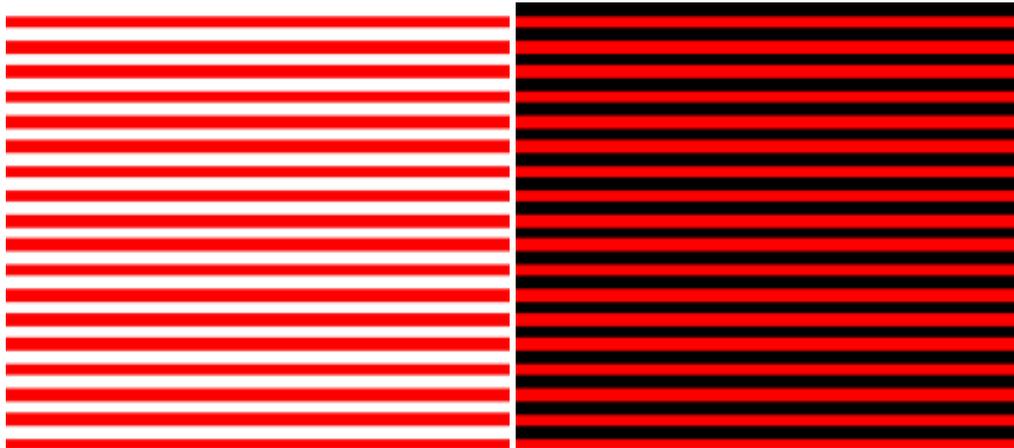
Figure 1.15-6: Bezold Effect illustrates how the appearance of colour is altered by the colours that surround them. In this case, the colour red appears lighter when surrounded by white, and darker when surrounded by black (Waranabe, 2007).

During the assessment of immunohistochemically stained Tissue Microarrays the comparison of colour is paramount. Reviewers often rely on previously reviewed Tissue Microarray cores to form their opinion of subsequent cores. For example a moderately stained core could be categorised as weak, if the core was reviewed following a series of

strongly stained cores, as human assessment is not a true value, rather a form of "comparison" of colours.

1.15.3 Reviewers workload

The aforementioned issues are the physiological factors of vision that introduce obstacles in human-based assessment of staining intensity. Those issues are inherent and are a constant disadvantage when assessment is performed by humans. In addition, there are difficulties in human-based reviews that are highly variable and are not solely restricted to vision. Pathologists must cope with three main areas of difficulty in their field, productivity, accuracy, and objectivity. Pathologists are under increasing pressure to improve productivity and, therefore, generating more data and process more slides. Tackling the mountain of work manually places a constant strain on resources, staff, and finances. For those analysing Tissue Microarrays, this burden is likely to worsen rapidly due to the volume of samples under analysis. In any field of science dependent on observation, accuracy is essential. However, it is well documented that, after prolonged visual study, eye and specifically cone-fatigue can significantly affect a person's ability to discern colour changes and identify unusual objects. The Tissue Microarray slide format has compounded this effect hundreds of times over. With densities exceeding 600 samples per slide, fatigue may quickly become an issue. Finally, there is the question of objectivity. The nature of the human eye is such that every person sees an object slightly differently from the way others see that same object. Subjectivity in this regard is therefore innate. Also, tints and shades can appear to change from one setting or context to another. Different observers may report seeing different features on the same object, as may a single observer at different times.

Visual inspection can also be confounded by the inherently subjective nature of human observation, which is affected by context, for example the amount of tumour present, background staining, stromal staining and even the order in which the spots are observed (Camp, 2005). Productivity, accuracy, and objectivity are further compounded by the following issues:

1.15.4 Scoring systems

Manual scoring systems are qualitative or semi-quantitative in nature, none reaching the level of true continuous quantitative data. Quantitative scales are either binary (+ or -) or normative (0, 1, 2, 3) typically providing categorization of a feature such as intensity. Simple qualitative scores are limited, such as percentage of positive cells. Qualitative scales have limitations in resolution which can be detected by eye, thus many researchers build a simple scale, as 0, 1, 2, according to negative, weak, strong. Manual reviews requires interpretative skills of well-trained investigators and frequently the efforts of a specialist primarily pathologist. Staining patterns that are anticipated to be used in clinical practise are usually scorable as positive or negative, whenever possible. Any added complexity in a scoring system introduces inter- and intra-observer reproducibility as is seen with challenges in scoring ER, PR, HER-2 (Braunschweig et al., 2004).

In comparison with human-based assessment there are numerous advantages of automated analysis in that it provides increased throughput, precise, reproducible, continuous, objective assessment of protein expression (Divito et al., 2004, Cregger et al., 2006). However, the greatest advantage is that it is truly quantitative. In contrast, the human eye, even that of a trained pathologist, has a difficult time accurately distinguishing subtle differences in staining intensity using a continuous scale. Consequently, scoring systems for pathologists tend to be nominal (for example, 0, 1+, 2+, 3+). In some cases, these can preclude the discovery of subtle sub-populations that cannot be identified using manual analysis (Camp et al., 2003, Camp et al., 2002, Camp, 2005).

1.15.5 Background lighting variation

Apart from the microscope itself, the next most important item in the reviewing process is the means by which the slides are illuminated; for example the microscope lamp (Micrographia, 2003). If too much light is exposed, information about the intensity of staining is lost. Adjustments in lighting settings can introduce huge variability in the

reviewing process. Consistency of light between slides and reviews is extremely difficult during microscope-based assessments.

The advent of virtual slides permitted the review of whole tissue slides across the Internet (Johnston et al., 2005). Virtual slides provide users with all the functionality of a microscope, but with numerous additional benefits, including concurrent access for multiple users, tracking of review movements and image annotation. Virtual slides are reminiscent of microscope use, they are favoured by pathologists over static digital images, due to the ability to change magnification and scroll laterally while reviewing the image. There are a growing number of interactive pathology sites available via the Internet. The diversity in their principle of operation, their application in telepathology, and their degree of sophistication promises an encouraging future in telepathology (Costello et al., 2003). Uniform lighting conditions can be applied across many slides, using digital imaging systems and therefore, eliminating the possibility of variance due to background lighting. In addition, when imaging analysis is utilised it is possible to normalise results across numerous slides (see section 2.6.1)

1.15.6 Reviewers Disorientation

Analysing Tissue Microarrays under a microscope can lead to misplaced spot orientation. It has previously been reported that it is difficult to accurately track the location of individual cores within complex Tissue Microarrays. In addition, microscope-based assessment typically relies on the manual entry of results first onto a worksheet and then subsequently into a spreadsheet or database system. Accurate manual tracking of the Tissue Microarray cores is challenging, prone to error and often leads to frustration (Tubbs et al., 2007).

1.16 Computer-assessed quantification of immunohistochemically stained Tissue Microarrays

The analysis of immunohistochemical staining patterns usually measures specific single targets rather than the relatively complex and intricate disease patterns, for example those seen on H & E staining. Immunohistochemical studies are, therefore, inherently amenable to automated image analysis (Joshi et al., 2007). As described in section 1.14.3, sources of variability in immunohistochemistry include fixation conditions, specimen pre-treatment, reagents, detection methods, and interpretation of results. Although it is not possible to standardise all the potential variables in immunohistochemistry, the interpretation of immunohistochemical results may be standardised through quantitative methods.

It has been proposed that computer-based analysis can quantify staining intensity more accurately and with greater reproducibility than human-based assessment (Weaver et al., 2003, Johansson et al., 2001). There are numerous commercially available, computer-based systems designed for the quantification of immunohistochemical staining, including BLISS and IHCscore of Bacus laboratories, Inc (Lombard, IL, USA); Ariol SL-50 Applied Imaging Corporation (USA); AQUA, HistoRx Inc (USA); ACIS developed by Chromovision (ChromaVision Medical Systems, Inc., USA). However, automated quantification systems can be expensive and time-consuming. Automated immunohistochemical protocols, in combination with a device that provides quantitative and objective output, could dramatically improve the quality of the data obtained from immunohistochemical studies (Cregger et al., 2006). For computer-based systems there are two key factors, stain quality and image quality. If image capture does not contain the information desired at adequate image resolution, the extraction of this information will be inaccurate (Hewitt, 2006).

Automated quantitative analysis (AQUA) is the most widely cited commercially available co-localisation application. Developed by Camp et al., 2003 it was specifically designed for quantification of immunocytochemistry on Tissue Microarrays (Camp et

al., 2002, Camp et al., 2003). Once the images are captured, the AQUA analysis is run using a set of algorithms. Rapid exponential subtraction algorithm (RESA), based on the principles of deconvolution theory, attempts to reduce the "out-of-focus" noise from parts of the specimen that are above and below the plane of focus. By subtracting an out-of-focus image from an in-focus image, RESA improves the resolution of the image, reducing the signal-to-noise ratio. Pixel-based locale assignment for compartmentalisation of expression utilises DNA binding fluorescent tags to distinguish tumour cells from stroma, as well as to define subcellular compartments (Dolled-Filhart et al., 2006). This algorithm also measures expression of a marker of interest within cellular and subcellular compartments. The output variable is a pixel intensity to pixel area value ranging from 0 to 255, based on the average intensity for all pixels evaluated (Cregger et al., 2006).

Automated Cellular Imaging Systems (ACIS) is produced by Clarient, formally known as Chromavision. ACIS is a digital microscope system with the ability to detect, count and classify cells based on colour shape and size. ACIS uses a combination of co-localisation and pattern recognition algorithms. ACIS is capable of detecting levels of hue, saturation and luminosity. Through the use of a digital camera, a voltage signal is produced which is proportional to the transmitted light intensity. The signal is then converted into a numerical density measurement. The system also has algorithms built into the software that allow it to analyse data and produce numeric scores for various parameters defined by the user. This device was originally designed for HER-2 analysis (Wang et al., 2002b) and was subsequently improved to allow analysis of Tissue Microarrays (Faith et al., 2004). It has been demonstrated in the scientific and pathology literature with many publications citing the use of ACIS in the analysis of immunohistochemistry (Tawfik et al., 2006, McKay et al., 2006). This system has enjoyed the greatest commercial success of all image analysis applications in pathology, although market penetration has been estimated only around the 5% level (Cregger et al., 2006).

TMAx is novel pattern recognition software application created by Beecher instruments, the leading provider for Tissue Microarray arrayers. TMAx utilises automatic pattern recognition software, with rule based operations and multiple iterative segmentations processes, together with fuzzy logic, to automatically identify cells and immunohistochemical deposits. As a result, colour channel values of immunohistochemistry can be calculated. TMAx also has a cell line application, which auto recognises cells, measures several key features describing the immunohistochemical staining intensity and counts the fraction of stained cells.

The validity of image analysis systems to be accurate, reproducible and to be at least as good as traditional methods of analysis has been extensively tested. Some of the major comparisons performed on image analysis systems are as follows:

1.16.1 Computer-assisted quantification of immunohistochemistry compared with Human Analysis

Comparisons of computer-assisted quantification and microscope-based assessment have shown some promising results, a number of case studies follow:

Initial validation of AQUA compared its accuracy, intra-observer variability, and predictive power to traditional human-based analysis. Tissue Microarrays derived from 340 node-positive breast carcinoma patients were immunohistochemical stained for the presence of ER. The correlation between AQUA and results from a human-based evaluation was high ($r^2 = 0.884$). Results illustrated AQUA has slightly better reproducibility than human analysis ($r^2 = 0.824$ versus $r^2 = 0.732$) (Camp et al., 2002).

Camp et al. (2003) compared traditional manual scoring with the AQUA system for the measurement of HER-2. Correlation of both reviewing methods with prognosis was performed. Utilising Tissue Microarrays, a cohort of 300 breast cancers was assessed. Manual examination of the immunohistochemical stained Tissue Microarray was performed by a single pathologist. As expected both methods identified a population of

high HER-2 protein overexpressing tumours with poor 30 years disease related survival. However, manual assessment failed to identify a population of tumours with low HER-2 protein expression, which when using AQUA proved to also be associated with poor prognosis.

Regression analysis demonstrated good correlation between the two methods (r^2 = 0.704). The percentage of HER-2-amplified cases in each manual category were 4.0% (0), 13.7% (1+), 71.4% (2+), and 75.0% (3+), and in each AQUA category were 17.5% (normal), 71.3% (intermediate), and 11.2% (high). There was a clear division between scoring 0/1+ and 2+/3+, there was no distinction between tumours scoring 0 and 1+. This result shows the difficulty in manually translating a continuous marker into a nominal four-point scale. Even for the trained eye of a pathologist, accurate distinction between nominal categories is difficult and often arbitrary. Examination of manual and automated techniques revealed that both were equally able to define a population of tumours expressing high levels of HER-2 and associated with poor outcome (p = 0.0007 and 0.0013, respectively). However, unlike manual analysis, automated analysis revealed that tumours expressing normal levels of HER-2 also showed a significantly worse outcome (p = 0.0091). Given the amount of overlap in the 0 and 1+ category from manual scoring, it is not surprising that manual assessment of stained slides has not previously identified the HER-2 normal population. Automated analysis identified a patient population that was otherwise detectable only by established biochemical assays (Camp et al., 2003).

Stromberg et al. (2007) compared human-based online assessments with the automated analysis performed by TMAx. A single Tissue Microarray which was constructed from 46 cell lines from different tissue types along with 12 cases of human tissue was examined by human and image analysis. Cell lines results from TMAx were compared to reviews performed by 7 pathologists who reviewed online intensity and fraction of stained cells. Intensity was recorded using a four point scale and fraction as a percentage of positive cells. The average result from the 7 reviewers was compared to the TMAx for 200 images. The scores from the TMAx were within the human-based review mean

of +/- 2 SD for both intensity and percentage of positive cells for a great majority of cases. Only 22 of 200 cases scored by TMAx were outside +/- 2 SD of the human analysis for both intensity and percentage staining cells. In addition, for each image, a combined output score for staining was calculated by multiplying the intensity mean value with the fraction mean value for the 7 pathologists. When comparing this score with the corresponding combined score from the TMAx, a high degree of correlation was observed ($r^2 = 0.94$, $p<0.0001$) (Stromberg et al., 2007).

Messerssmith et al. (2005) compared the ability of ACIS automated quantification to correlate with human-based assessment across a number of antibodies. A blinded pathologist scored a Tissue Microarray constructed from 18 human samples of colorectal cancers and matched normal mucosa. The pathologist scored the intensity of staining using a four graded scale (0 - 3+). Plotting the ACIS scores (normalised to the mean of the samples which were scored as 0) against the pathologists scores showed a variable, but definite relationship between ACIS and pathologist scoring. One way ANOVA analysis showed a significant correlation between ACIS and the pathologist scoring ($p< 0.001$) for all six proteins which were examined, including EGFR, Akt, pAkt, MAPK and Pmark (Messersmith et al., 2005).

Clearly automated analysis compares favourably with pathologist based interpretation of microarrays; however, the true criterion standard is outcome prediction (Camp et al., 2002).

1.16.2 Computer-assisted quantification of immunohistochemistry compared with Prognosis

Correlation between immunohistochemical staining and prognosis has been performed by computer-assisted analysis. A number of examples of these studies are as follows:

The role of B cell lymphoma 2 (Bcl-2) expression in melanoma and it's associated with patient outcome has previously been reviewed. However, inconsistency in association with clinical outcome has been reported. Some studies report that up-regulation of Bcl-2 correlates with advanced stage and poor prognosis, whereas others report that down-regulation of Bcl-2 is associated with disease progression, advanced stage, and poor prognosis. Divito et al. (2004) assessed the role of Bcl-2 in melanoma utilising AQUA. Automated quantitative analysis was performed on 2 replica Tissue Microarrays constructed from 402 cases of melanoma. Pixels were defined as melanoma and positive intensity of Bcl-2 was measured. A continuous index score was generated which was subsequently divided into quartiles and correlated with clinical variables. Kaplan-Meier survival curves illustrated that increased membranous/cytoplasmic Bcl-2 expression was correlated with better outcome (log-rank $p = 0.004$). Bcl-2 expression was significantly higher in the primary specimens than in the secondary specimens ($p < 0.0001$). Results were reproducible when using a second Tissue Microarray with different cores from the same patients (Divito et al., 2004).

Harigopal et al. (2005) assessed the prognostic value of E-cadherin expression in node-positive breast cancer, utilising AQUA. Automated quantitative analysis was performed on one immunohistochemically stained Tissue Microarray constructed from 280 invasive ductal and 61 invasive lobular breast cases. Results illustrated there was no significant difference in mean staining intensity between the primary and nodal specimens ($p = 0.80$). A scatterplot was generated which identified a subset of patients (25%) with high E-cadherin expression in nodal metastases, and this top quartile had improved survival ($p = 0.028$). On univariate analysis, increased E-cadherin expression in nodal metastases was strongly associated with extended survival ($p = 0.007$), whereas

expression in primary tumours was not (p = 0.130). Strong E-cadherin expression in lymph node metastases was highly predictive of improved survival. This suggests that expression of adhesion molecules at metastatic sites represents less aggressive tumour behaviour. This had not previously been quantitatively assessed (Harigopal et al., 2005).

Zerowski et al. (2007) determined the prognostic value of Cox-2 expression in a large cohort of breast cancer Tissue Microarrays. Cox-2 has been associated with the development, growth, and spread of cancers by stimulation of angiogenesis, inhibition of apoptosis, and enhancement of invasiveness. Previously, the prognostic value of Cox-2 was not well established. The authors constructed a Tissue Microarray composed of 669 primary breast cancer biopsies, and Cox-2 protein concentration was quantitatively assessed using AQUA. Cox-2 was localised on the membrane / cytoplasm. X-tile, a novel software application was used to determine optimum cut points of the continuous data. In order to rigorously validate the statistical significance of a cut point, X-tile divides the entire cohort into two populations matched for survival, one designated as the training set and the other as the validation set. Within the training set an optimal cut point was determined. When this cut point was tested in the validation set, the P-value was 0.0189 and 0.0055 for disease-free and overall survival respectively. This is highly significant in light of the fact that the validation set was now half the size of the original cohort. Using AQUA and X-tile analysis, the authors illustrated that Cox-2 expression was likely to be upregulated in invasive breast tumours (Zerkowski et al., 2007).

The role of ß-catenin in breast cancer and its prognostic value is controversial. The prognostic value has been assessed previously in a series of non-quantitative immunohistochemical studies, resulting in conflicting results. Using quantitative analysis methods, Dolled-Filhart et al. (2006) clarified that low-level membranous ß-catenin expression is associated with significantly worse outcome for breast cancer patients (38% versus 76%, 10-year survival, validation set log-rank p = 0.0016). The authors assessed ß-catenin protein expression by evaluating a 600 case cohort of breast cancer tumours which were constructed into a single Tissue Microarray. The Tissue Microarray was assessed by AQUA with a series of array-embedded cell lines for which

the ß-catenin concentration was standardised by an ELISA assay. Total ß-catenin protein levels (as ng/mg protein lysate) were detected by ELISA and by AQUA analysis of the same cell lines embedded into a Tissue Microarray. Comparison of the AQUA scores of the embedded cell lines and the ELISA concentrations showed a correlation of $r^2 = 0.853$. The calculation of these cell line concentrations provided a standard curve for the conversion of ß-catenin AQUA scores to a specific concentration for each of the tumour samples. X-tile software was also used to select optimal protein concentration cut points and to evaluate the outcome using a training set and a validation set (Dolled-Filhart et al., 2006).

1.16.3 Computer-assisted quantification of immunohistochemistry compared with Biochemical techniques

Computer-assisted quantification of immunohistochemistry is capable of determining the same level of protein expression as those obtained utilising quantitative bioassays. A number of studies that illustrate computer-assisted quantification of immunohistochemically stained tissue are described below:

Tawfik et al. (2006) compared the results of HER-2 protein overexpression and gene amplification in the same patient specimens by immunohistochemistry (HercepTest™) using the ACIS system and by FISH. Whole section analysis was performed on 247 primary breast cancers, by both ACIS and FISH. ACIS quantified at least 5 / 6 areas with the highest staining intensity. An average score for all selected areas was then calculated. The manufacturer of ACIS recommends that cases with an average score of 2.2 or higher be considered to have HER-2 protein overexpression, while cases with average scores lower than 2.2 should be considered not overexpressed. FISH results were reported as the ratio of the mean copy number of HER-2 signals to the mean copy number of CEP 17 signals. Specimens with a ratio of ≥ 2.0 were designated as having HER-2 gene amplification and those with a ratio < 2.0 as having no gene amplification. A pathologist verified the signal pattern in each case. The concordance rate between immunohistochemical-ACIS and FISH HER-2 tests was 94% for the 247 cases. This

result was markedly improved from an overall 72% concordance rate between immunohistochemistry by manual microscopy using the same HercepTest™ and FISH as previously performed in the authors' institution on 250 cases (unpublished data). The overall HER-2 2+ cases were 18.6% and 19% by immunohistochemistry-ACIS and FISH testing, respectively, which is significantly less than those reported in manual assessment. Despite the high levels of agreement between the tests there were still cases that fall into a grey zone in which there is disparity between the two tests. Accordingly, neither test can definitively guide clinicians in determining appropriate treatment for all patients (Tawfik et al., 2006).

Ciampa et al. (2006) compared HER-2 expression by immunohistochemistry quantified by ACIS with that of the gene amplification by FISH. One hundred and eight whole sections of primary infiltrating ductal carcinomas were immunostained using the HercepTest™ and the cases were divided into four groups: group 1 with HER-2 expression ACIS score ≤1.5; group 2 with a score 1.6-1.9; group 3 with a score 2.0-2.5; and group 4 with a score ≥2.6. FISH was performed on all of the 108 cases. All cases were also manually reviewed. Within the ACIS scoring system gene amplification was verified by FISH as follows: group 1: 4%, in group 2: 12%, in group 3: 11%, and in group 4: 86%. Furthermore, in group 4, 100% cases with an ACIS score of ≥3 were FISH positive. Correlation with manual immunohistochemical score and FISH showed that 9% immunohistochemical negative (0 and 1+) cases and 29% immunohistochemical positive (2+ and 3+) cases showed gene amplification by FISH. This study shows that the amplification of the HER-2 gene correlates better with overexpression of the HER-2 protein by immunohistochemistry when the score is <1.5 or > 2.6 by ACIS. Therefore, FISH may be useful to better evaluate HER-2/neu status in breast cancer in cases where the ACIS score by immunohistochemistry is 1.6-2.5. Since the correlation is so good, FISH may not be needed for HER-2 evaluation in cases with ACIS scores < 1.5 and > 2.6.(Ciampa et al., 2006)

Messersmith et al. (2005) compared quantitative results from ACIS, ELISA and western blots for EGFR expression in one CMA composed of eight head and neck tumours.

Experiments indicated that ACIS obtained similar results as ELISA and western blots. Scatter plots of ELISA results compared to ACIS results illustrated good correlation (r^2 = 0.697). It was suspected that one cell lines with low level expression may have negatively affected agreement between ACIS and ELISA (Messersmith et al., 2005)

1.17   Defining the need for this project

Despite the huge advances made by commercially available complete digital imaging devices, there are also prohibitive limitations to these systems. Currently, the cost of purchasing high-throughput scanners are extremely prohibitive, usually running at between 60,000 to 180,000 Euros (Rojo et al., 2006). As a result, commercial systems are not always viable in research or small laboratories (Camp, 2005, Camp et al., 2003, Camp et al., 2002). Collaborations can be created between institutions, where a scanner is purchased by a consortium of institutes and slides are posted for scanning. However, without access to a low cost web-based system for review, scoring and subsequent storage of images and data associated with virtual Tissue Microarrays, pathologists are reduced to recording results on worksheets and subsequently entering data into computer software applications. This system is completely counter-productive and negates the advantages and possibilities of virtual slide technology.

There are numerous publications listing the merits of image analysis as a means to quantifying protein expression on Tissue Microarrays. However, the influx of commercially available image analysis applications utilised to quantitate protein expression are extremely complex and difficult to use. Virtual Slide technology is currently being introduced in pathology training; however, it will be a number of years before this is the norm. The objective of automated quantification of protein expression has always been to free pathologists from the mundane task of calculating the level of "brown" on a slide and, in doing so, to produce results at least as accurate as the pathologist can. The commercially available image analysis systems have succeeded in freeing the pathologist from their traditional role at the bench. However, pathology is a complex and highly skilled field and, with the introduction of some commercially available image analysis systems, it has forced the onus on pathologists to also become experts in computer image analysis.

The majority of commercial image analysis systems are targeted at fluorescently-labelled tissue, due to the relative ease of compartmentalisation by this means

identification compared with chromogen-based methods. However, chromogen-based immunohistochemistry is the most common application of Tissue Microarrays. The use of fluorescently-labelled tissue introduces a number of strains on analysis: slides have to be digitised as soon as possible before signal is lost, therefore eliminating the possibility of laboratory consortiums, techniques are difficult and cost of antibodies and scanning equipment is substantial compared to immunohistochemical applications. Therefore, the development of novel chromogen-specific image analysis methodology is crucial in the ongoing investigation of tumour tissue via Tissue Microarrays.

CHAPTER 2: MATERIALS AND METHODS

SECTION A: LABORATORY METHODS

2.1 Cell Culture

2.1.1 Cell Lines

All cell culture work described within this body of work was carried out in a class II laminar airflow cabinet (Holten LaminAir, Denmark). The laminar airflow cabinet was swabbed with 70% industrial methylated spirits (IMS) before and after use, as were all items used within the cabinet. Only one cell line was used in the laminar airflow cabinet at any one time. Each cell line was assigned specific waste and medium bottles. After cleaning, and when working with different cell lines, the laminar airflow cabinet was allowed to clear for at least 15 minutes before and after use. The cabinets were cleaned weekly with the industrial disinfectant Virkon (Antech International, UK) as were the incubators. All of the cell lines included are anchorage dependent. Cell lines were maintained in 25, 75 and 175 cm^3 flasks (Corning Life Sciences, USA) at 37°C and in an atmosphere of 5% CO_2, where required. All cell lines were fed every two to three days. A complete list of the cell lines utilised, their source and the medium used is illustrated in Table 2.1-1. Cell lines were sourced from the European Collection of Cell Cultures (ECACC), National Cell and Tissue Culture Centre / National Institute for Cellular Biotechnology (NCTCC/NICB), and American Type Culture Collection (ATCC). Within Table 2.1-1 the following acronyms are used; NEAA - Non essential amino acids, FCS - Fetal calf serum and NA - Sodium.

Table 2.1-1: Cell lines and basal medium used within this body of work.

Cell Lines	Details	Source	Medium
A549 Parent	Lung Adenocarcinoma	ECACC	ATCC 5% FCS
A549 Taxol	Taxol-selected variant of A549	NCTCC/NI	ATCC 5% FCS
BT20	Invasive Ductal Carcinoma (Breast)	ATCC	RPMI 1640 10% FCS
BT474	Invasive Ductal Carcinoma (Breast)	NCTCC/NI	RPMI 1640 10% FCS
DLKP Parent	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLKP JC	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLKP Txt RL	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLKP Txt YL	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLKP VCR	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLKP YL	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLRP Parent	Poorly Differential Squamous Cell Lung	NCTCC/NI	ATCC 5% FCS
DLRP Carboplatin	Carboplatin-selected variant of DLRP Parent	NCTCC/NI	ATCC 5% FCS
DMS 53 Parent	Small Cell Lung Carcinoma	ECACC	RPMI 1640 10% FCS
DMS 53 Taxol	Taxol-selected variant of DMS 53	NCTCC/NI	RPMI 1640 10% FCS
DMS 53 Taxotere	Taxotere-selected variant of DMS 53	NCTCC/NI	RPMI 1640 10% FCS
DMS 53	Carboplatin-selected variant of DMS 53	NCTCC/NI	RPMI 1640 10% FCS
H460 Parent	Large Cell Lung Carcinoma	ATCC	RPMI 1640 10% FCS
H460 Taxol	Taxol-selected variant of H460	NCTCC/NI	RPMI 1640 10% FCS
HCC 1419	Breast Primary Ductal Carcinoma	ATCC	RPMI 1640 10% FCS
HCC 1954	Breast Primary Ductal Carcinoma	ATCC	RPMI 1640 10% FCS
MCF-7	Invasive Ductal Carcinoma (Breast)	ATCC	RMPI 10% FCS
MDA-MB-231	Breast Adenocarcinoma	ATCC	RPMI 1640 10% FCS
MDA-MB-453	Breast Metastatic Mammary Gland	ATCC	RPMI 1640 10% FCS
MDA-MB-468	Breast Adenocarcinoma	ATCC	RPMI 1640 10% FCS
NCI-H1299 Parent	Large Cell Lung Carcinoma	ATCC	RPMI 1640 1mM Na Pyruvate, 5% FCS
NCI H1299 Taxol	Taxol-selected variant of H1299	NCTCC/NI	RPMI 1640 1mM Na Pyruvate, 5% FCS
SK-BR3	Breast Adenocarcinoma	ATCC	RPMI 1640 10% FCS

SKLU-1 Parent	Lung Adenocarcinoma	ATCC	MEM 1% NEAA, 1mM Na Pyruvate 2mM L-glutamine
SKLU-1 Taxol	Taxol-selected variant of SKLU-1	NCTCC/NI	MEM 1% NEAA, 1mM Na Pyruvate 2mM L-glutamine
SK-MES-1 Parent	Squamous Cell Lung Carcinoma	ATCC	MEM 1% NEAA, 1mM Na Pyruvate 2mM L-glutamine
SK-MES-1	Carboplatin-selected variant of SK-MES-1	NCTCC/NI	MEM 1% NEAA, 1mM Na Pyruvate 2mM L-glutamine

2.1.2 Sub-culturing of adherent cell lines

Waste cell culture medium was removed from the tissue culture flasks into sterile waste bottles, and the cells were rinsed with pre-heated trypsin / EDTA (TV) solution (0.25% trypsin (GIBCO®, Invitrogen Corporation, UK), 0.01% EDTA (Sigma-Aldrich, Ireland) solution in PBS (phosphate buffered saline) (Oxoid, UK). The cells were rinsed with TV to ensure any residual medium was removed, as the serum content could act as a trypsin inhibitor. Pre-warmed TV was added to the flask, which was incubated at 37°C for approximately 5 minutes, in order to detach the cells from the flask. Approximately 1, 2, or 5 ml of pre-warmed TV was added to 25, 75, or 175 cm^2 flasks, respectively. The trypsin was deactivated by adding an equal volume of complete medium to the flask. The contents of the flask was then transferred into a sterile universal tube (30 or 50ml Sterilin, Barloworld Scientific, UK) and centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded and the cell pellet was re-suspended in a pre-warmed fresh complete growth medium, and cells were used to re-seed a flask at the required density.

2.1.3 Cell counting

An aliquot of trypan blue (GIBCO®, Invitrogen Corporation, UK) was added to an aliquot of cell suspension at a ratio of 1:5. The trypan blue and cell suspension was incubated for 3 minutes at room temperature, after which a drop of this mixture was added to a chamber of a haemocytometer which was enclosed with a glass cover-slip. Viable cells remained unstained by the trypan blue, non-viable cells stained blue. Viable cells in the 16 squares of the four outer corners of the haemocytometer were counted microscopically. The average cell count was multiplied by the dilution factor and 10^4 to determine the number of viable cells per ml in the original cell suspension.

2.1.4 Embedding Technique

The detachment of cells from the culture flasks using TV (as described in Section 2.1.2) was only utilised when cells needed to be passaged. When two 175 cm^2 flasks were 80% confluent, cells were collected in order to create a cell line pellet for Cell Microarray construction. It was not possible to use TV for detachment of cells from the flask at this

stage as trypsin potentially destroys cell surface protein markers. As a result, non-enzymatic cell dissociation solution (Sigma-Aldrich, UK) was used, as it causes minimal damage to the surface proteins of the cell.

To achieve this, waste cell culture medium was removed from the tissue culture flasks into sterile waste bottles, and the cells were rinsed with pre-heated PBS solution (Oxoid, UK). The cells were rinsed with PBS to ensure any residual medium was removed, as the medium could act as a cell dissociation solution inhibitor. Pre-warmed cell dissociation solution was then added to the flask, which was incubated at 37°C for approximately 15 minutes, in order to detach the cells from the flask. Approximately 5 ml of pre-warmed cell dissociation buffer was added to each 175 cm^2 flasks. The cell dissociation buffer was deactivated by adding an equal volume of medium to the flask. The contents of the flask was then transferred into a sterile universal tube (50 ml Sterilin, Barloworld Scientific, UK) and centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded and the cell pellet was re-suspended in pre-warmed PBS (20 ml). The content of the universal was centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded, and 20 mls of formalin (Serosep Ltd, Ireland) was added very slowly to the universal, ensuring not to disturb the cell pellet. Cells were fixed overnight at 4°C.

2.1.4.1 Agarose mold technique

Cells were centrifuged at 1000 rpm for 5 minutes, formalin was removed and the cell pellet was then packed into agarose molds as follows. Conventional agarose (2% Agarose Type I, Low EEO, Sigma-Aldrich, Ireland) was prepared in TE buffer. The agarose was poured into a Petri dish and allowed to set firmly, while placed on an even surface. After setting, cork borers (size 6 Sigma-Aldrich, Ireland) were used to cut dish-shaped molds out of the agarose. Using a smaller diameter cork borer (size 2), a central core was removed from each dish of agarose. The result was a donut-shaped agarose mold. Using a spatula, cell pellets were packed into the centre of the agarose mold and were wrapped in Bio-Wraps (Surgipath, UK), which had been moistened in 1% low

melting point agarose (Sigma-Aldrich, Ireland). The agarose molds where then placed into embedding cassette (Lennox, Ireland), submerged in 10% formalin (Serosep, Ireland), and sent for tissue processing (2.1.4.2).

2.1.4.2 Processing Cell Pellets

Cell Line samples were processed using a Citadel™ 2000 tissue processor (Thermo Fisher Scientific Inc, USA). The purpose of processing the cells is to ensure fixation, full dehydration and paraffin wax infiltration. To achieve this, processing cassettes containing the specimens were transferred to the Citadel™ 2000 processing baskets. Baskets were then inserted into the Citadel™ 2000, and processed for 16 hours. The stages involved in processing the cell lines are illustrated in Table 2.1-2.

Table 2.1-2: Stages of tissue processing performed to all tissue and cell lines samples assessed within this body of work. Tissue Processing was performed using Citadel™ 2000 automated processor.

Station	Reagent	Time (hrs)
1	10% Formalin	1
2	10% Formalin	2
3	70% Spirit	1
4	95% Spirit	1
5	Spirit (99% IMS)	1
6	Spirit	1
7	Spirit	1
8	Spirit	1
9	Xylene	1.5
10	Xylene	1.5
11	Wax	2
12	Wax	2

2.1.4.3 Paraffin wax embedding of cell line samples

Dehydrated cell lines housed in agarose molds were embedded in paraffin (Tissue Embedding Medium, Kendall Tyco Healthcare / Paraplast, USA) using a HistoEmbedder TBS88 (Medite, Germany).

2.2   Construction of Microarrays

2.2.1.1 Cell Microarrays

Cell Microarrays were designed and constructed in Tissue Array Research Program (TARP), National Cancer Institute (NCI), National Institute of Health (NIH), Washington DC, USA in collaboration with Dr. Stephen Hewitt. Cell Microarrays were constructed using the MTA-I manual arrayer (Beecher Instruments, USA) using 0.6 mm needles, following the protocol supplied by the manufacturer and as described in literature (Kononen et al., 1998). The array design included 31 cell line specimens and one liver control spot which was utilised for orientation. Sections were cut (4 µm) using a microtome (RM2155, LEICA Microsystems, Germany) and placed on adhesive coated slides (PSA-1X Paraffin tape transfer system, Instrumedics Inc, USA). Slides were vacuum packed and stored at 4°C until immunohistochemical staining was performed.

2.2.1.2 Tissue Microarrays

Tissue Microarrays were designed and constructed by staff of the Histopathology Department, Beaumont Hospital, Royal College of Surgeons Ireland (R.C.S.I). Tissue Microarrays were constructed using the MTA-I manual arrayer (Beecher Instruments, USA) using 0.6 mm and 2.0 mm needles, following the protocol pioneered by Kononen et al., 1998. Formalin-fixed-paraffin-embedded tissue blocks containing tumour material were retrieved from the archives, along with their respective haematoxylin and eosin (H&E) stained sections. Using the H&E stained slide as a template, representative areas of carcinoma were identified and marked. Sections were cut (4 µm) using a microtome (RM2135, LEICA Microsystems, Germany) from each Tissue Microarray mounted on to charged adhesive slides (Plus slides, Vision Biosystems, UK) using the water bath transfer technique and were air dried overnight at 55°C.

2.3   Experimental analysis of Tissue / Cells

Manual and Automated immunohistochemical staining was performed as part of this body of work. Fluorescent *in situ* hybridisation (FISH), Western blots and Enzyme Linked Immunosorbent Assay (ELISA) were not performed by the author.

2.3.1   Immunohistochemical staining

Both automated and manual immunohistochemistry was performed. In total, 6 primary antibodies were applied to sections of the Cell Microarrays and 4 primary antibodies were applied to Tissue Microarrays. All antibodies targeted antigens on the cell membrane. In addition, ß-catenin targeted the cytoplasm and PhosphoMet targeted the cytoplasm and nucleus. A list of primary antibodies utilised in this study is illustrated in Table 2.3-1.

Table 2.3-1: Primary antibodies and the dilutions utilised within this body of work.

Primary Antibody	Dilution	Microarrays
EGFR (Novocastra/Vision Biosystems, UK, NCL-L-EGFR-384)	1:75	CMA
MDR-1 (National Institute for Cellular Biotechnology, 6/1C)	Neat	CMA
MDR-3 (National Institute for Cellular Biotechnology, 6/1G)	Neat	CMA
MRP-1 (National Institute for Cellular Biotechnology, P268)	Neat	CMA
E-cadherin (Novocastra/Vision Biosystems, UK, NCL-E-Cad)	1:50	CMA/TMA
HER-2 (DakoCytomation, Code K5204)	Neat	CMA/TMA
PhosphoMet (Cell Signalling Technology, Tyr1234/1235)	1:50	TMA
ß-catenin (Labvision Corp, RB-9035-P1)	1:50	TMA

2.3.1.1 Manual Immunohistochemical Staining

Immunohistochemical staining for MDR-1/Pgp, MDR-3 and MRP-1 with the NICB antibodies (anti-MDR-1, anti-MDR-3 and anti-MRP-1) was performed manually using the technique described by (Hsu et al., 1981), using an avidin-biotin horseradish peroxidase (HRP) conjugated kit (ABC) plus an appropriate secondary antibody.

Briefly the protocol utilised is as follows; Paraffin-embedded tissue samples were dewaxed in xylene (2 x 5 minutes) rehydrated in grading alcohols 100%, 90% and 70% (2 x 3 minutes) and placed in Tris Buffered Saline (TBS/0.05% Tween-20). Endogenous peroxidase activity was quenched by placing tissue sections in 3% (vol/vol) H_2O_2/distilled water for 5-7 minutes at room temperature. All slides were blocked for non-specific staining with 20% normal rabbit serum (DakoCytomation, UK)/TBS for 20 minutes at room temperature. Primary antibodies optimally diluted in TBS/0.05% Tween-20 where applicable (antibody MDR-1 clone 6/1C, MDR-3 clone 61/G and MRP-1 clone P268) were applied to each sample. Primary antibodies were incubated overnight at $4^{o}C$. Samples were then washed (3 x 5 minutes) with TBS/0.05% Tween-20. This was followed by a 30 minute incubation with biotinylated secondary antibody (rabbit anti-mouse IgG (1/300 dilution in TBS/0.05% Tween-20) (DakoCytomation, UK) or rabbit anti-rat (1/500 dilution in TBS/0.05% Tween-20). Finally, following another (3 x 5 minutes) wash step, Vectastain Elite ABC reagent (HRP conjugated) (Vector Laboratories, UK) was applied for 25 minutes at room temperature. The peroxidase substrate, 3',3-daminobenzidine tetrahydrochloride (DAB) containing 0.02% H_2O_2 (Vector Laboratories, UK) was added for 10 minutes at room temperature. All slides were then washed (3 x 5 minutes) TBS/0.05% Tween-20. Tissue sections were then lightly counter stained with haematoxylin (Vector Laboratories, UK).

Slides were subsequently dehydrated in grading alcohols 70%, 90% and 100% (2 x 3 minutes). Samples were then cleared in xylene and mounted in DPX (BDH, UK). Negative control samples in which primary antibody were replaced by 1x TBS/0.05% Tween-20 were included in all experiments. Positive controls (normal breast and lung tissue), using the same experimental conditions, were included in all experiments.

2.3.1.2 HercepTest™ Immunohistochemical Staining

The assessment of HER-2 protein overexpression was carried out using the HercepTest™ kit according to the manufacturer's instructions. The HercepTest™ protocol differs from that described in section 2.3.1.1 as follows; after dewaxing in xylene and rehydration through graded alcohols to distilled water, the section was subjected to heat induced epitope retrieval by immersion in preheated epitope retrieval solution (95 – 99°C) in a waterbath for 40 minutes. The sections were then allowed to cool down for 20 minutes at room temperature. Endogenous peroxidase activity was blocked by a five minute treatment with peroxidase blocking agent. Sections were then incubated with the anti-HER-2 polyclonal antibody for 30 minutes at room temperature, followed by incubation with a visualisation reagent (labelled streptavidin–biotin–immunoperoxidase). The antigen–antibody reaction was visualised using 3-3'-diaminobenzidine as chromogen and counterstaining was performed with haematoxylin. Suitable negative and positive control slides were treated in a similar manner to ensure appropriate staining (Gulmann et al., 2004).

2.3.1.3 Automated Immunohistochemical Staining

Immunohistochemical staining of Microarrays with E-cadherin, EGFR PhosphoMet and β-catenin was performed using the automated stainer BondMax™ (Vision Biosystems™, UK) and according to manufacturers instructions.

2.3.2   Western Blots

Western blotting was performed by fellow post-graduate / post-doctorate members of the NICB (Brigid Browne, Denis Collins, Brendan Corkery and Norma O'Donovan), and as described in literature (Moran et al., 1997a). Western blot analysis was performed on cell lines within 5 – 10 passages of those used to construct the Cell Microarray.

2.3.3   Enzyme-Linked Immunosorbent Assay

Protein expression, when quantified by ELISA. ELISA testing was performed by fellow post-graduate / post-doctorate members of the NICB (Brigid Browne, Denis Collins, Brendan Corkery and Norma O'Donovan) as part of a wider study and results were made available for comparison within this study. Commercially available ELISA systems were utilised for quantification of total HER-2 (Calbiochem®, UK) and EGFR (DuoSet® IC Human Total EGFR ELISA, R&D Systems Inc., USA) expression. ELISAs were performed on cell lines within 5 – 10 passages of those used to construct the Cell Microarray. ELISA results were classified as positive (> 20 ng/mg) or negative ($\leq$ 20 ng/mg).

2.3.4   Fluorescent *in situ* hybridisation

The PathVysion® HER-2 DNA probe kit and paraffin wax pre-treatment kit (Vysis Inc, UK) were used in accordance with the manufacturer's recommended protocol. Slides were stored at –20°C in the dark before signal enumeration with a Nikon Eclipse E600 Microscope (Nikon Inc, USA) using x 100 magnification (oil immersion) and appropriate multiband pass filters. Appropriate positive (HER-2 gene amplification) and negative (non-amplification) control slides were included in the staining run. Analysis of the FISH score was carried out by comparing the ratio of the average copy number of the HER-2 gene with that of the chromosome 17 centromere in 60 nuclei/case (Kay et al., 2004).

SECTION B: COMPUTER APPLICATIONS

2.4   Digitisation of Slides

Slides within this body of work were digitised using a number of techniques. The digitisation techniques utilised were updated with advances in technology. The most advanced technology was utilised toward the end of the thesis. The technologies utilised to digitised slides are listed below, and are arranged in chronological order.

2.4.1   Olympus BX-40 Microscope

An Olympus BX–40 microscope (Olympus, USA) incorporating a Prior H101 motorised stage was used to capture digital images of tissue within the initial creation of the Virtual Tissue Matrix in Chapter 3. Images were captured at 4 x using a Plan Achromatic lens, and at 20 x using a Plan Fluorite lens. The camera used to scan the immunostained Tissue Microarrays was a 3-chip JVC KY 55 B 3 CCD. The camera has a red, green and blue (RGB) digital signal output to an Imaging Technologies IC RGB frame grabber (Coreco Imaging Incorporated, USA), which was housed in a Silicon Graphics ZX10 imaging workstation. A software algorithm was constructed using the Optimas development environment (Media Cybernetics, USA) that facilitated the remote control of the stage and the construction of wide field-of-view images from a montage of smaller fields. Scanning time at 20 x was approximately 12 - 16 minutes for a 20 mm x 20 mm biopsy. All software was written in-house by previous members of the Medical Informatics Group, Dublin City University (Johnston, 2005b).

2.4.2   Aperio ScanScope T3 Scanner™

Immunohistochemically stained Tissue Microarrays were digitised using an Aperio ScanScope T3 Scanner™ (Aperio Technologies, USA). Images were captured at 20 x magnification, saved in SVS format utilising JPEG compression and were approximately 90 Mb per Tissue Microarray. Scanning time at 20 x was approximately 6 - 8 minutes for a 20 mm x 20 mm biopsy. The Aperio ScanScope T3 Scanner™ was used to digitise slides within Chapters 4 and 5.

2.4.3 NanoZoomer Digital Pathology System

Immunohistochemically stained Tissue Microarrays and whole sections within Chapter 6 were digitised using a NanoZoomer Digital Pathology (NDP) System (Hamamatsu, UK). The NDP system utilises CCD TDI technology to achieve scans with a spatial resolution of 0.46 µm / pixel. Scanning time at 20 x was approximately 3 minutes for a 20 mm x 20 mm biopsy. Images were approximately 90 Mb per Tissue Microarray and ranged from 55-487 Mb per whole section biopsy.

2.4.4 Nikon Eclipse E600 Microscope

Fluorescently labelled tissues described in Chapter 6 were digitised using Nikon Eclipse E600 Microscope (Nikon Inc, USA), by members of staff at Beaumont Hospital /Royal College of Surgeons. Slides were stored at –20°C in the dark before signal enumeration using x 100 magnification (oil immersion) and appropriate multi-band pass filters.

2.5 Development of an Algorithm

2.5.1 Image-Pro Plus[®]

Image analysis was performed in Image-Pro Plus[®] (Media Cybernetics, USA). The image analysis algorithm was developed to perform the following functions:

- To isolate positive membrane staining only.
- To create a histogram of intensity of pixel values and output the histogram values, mode, mean and standard deviation. The algorithm also recorded the background lighting staining intensity value, and the number of positive membrane and tissue pixels.

The algorithm was a combination of AutoPro™, Media Cybernetics own programming language and Visual Basic. Once created and extensively validated, the algorithm was saved as a macro. The macro facilitated automated analysis of the images. Human intervention was reduced to selecting the Tissue Microarrays to be analysed. On average, image analysis required four minutes per Tissue Microarray spot. The image analysis was performed on networked laptop computer with a 40 GB hard drive, 1.0 GB of RAM and an Intel® Pentium ® M processor 1.50GHz.

The key aspect of the algorithm is the detection of membrane staining only, which was accomplished using a combination of filters and a colour cube. The sequence of the algorithm is divided into three stages and is depicted in Figure 2.5-1. The algorithm was tested on numerous images to ensure it was equally effective on varying degrees of staining intensities.
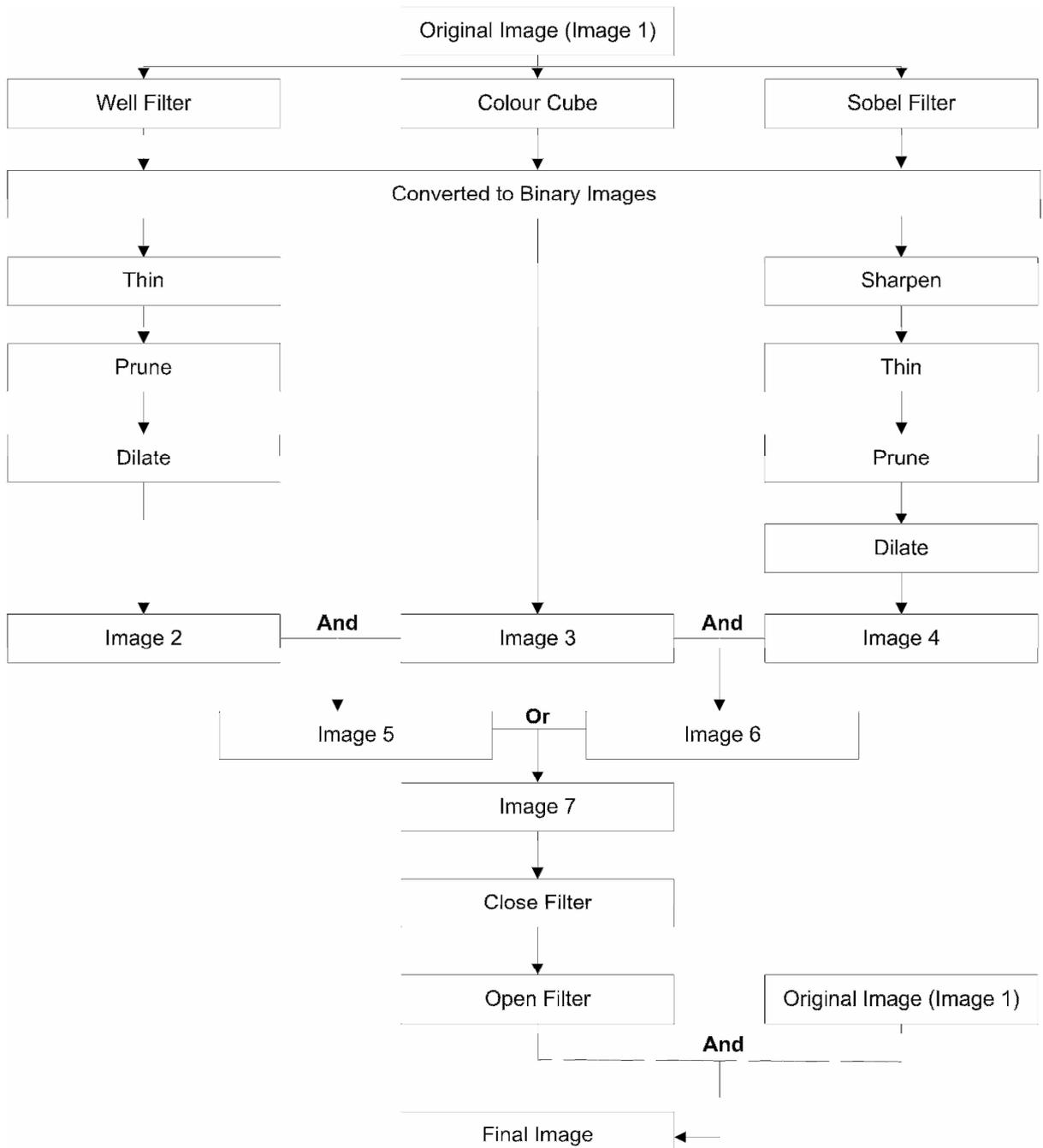
Figure 2.5-1: Sequence of the image analysis algorithm, which isolates membrane-bound immunohistochemical staining. Colour cube is used to identify brown staining. Filters are used to isolates areas of membrane-bound staining (Well, Sobel, Prune, Dilate, Thin, Sharpen, Open and Close).

Stage 1 involves the creation and application of a colour cube, which represents the range of colours present when tissue stained positive for immunohistochemical staining. Using the eyedropper tool within Image-Pro Plus®, colours were selected from numerous images to comprehensively represent all colour intensities present for each of the antibodies under assessment. The colour cubes were then applied to the original images and a mask was created, as illustrated in Figure 2.5-2. The mask converted pixels within the original image that were not a colour present within the colour cube to black, and those with pixels colours within the colour cube were converted to white.

Stage 2 involves applying a number of filters to the image in order to identify membrane staining only; the types of filters used are as follows:

- Well: detects and emphasises points that are darker than the background.
- Sobel: enhances the principal edges, this identifies the membrane.
- Thin: reduces a binary image to its skeleton.
- Prune: reduces a binary image to its skeleton and removes protrusions.
- Dilate: enlarges the edges of bright objects and erodes dark edges.
- Sharpen: enhances fine details.
- Close: Fills gaps and enlarges protrusions to connect objects that are close together.
- Open: Smoothes object contours, separates narrowly connected objects and removes small dark holes.

The filters "Sobel" and "Well" were of particular importance in identifying membrane staining.

Stage 3 involved a number of operations which were used to combine multiple images. The "And" function performs a logical operation between two images, in which only bit values that were on in both images will be present in the output image. The "Or"

function performs a logical operation between two images, in which bit values that are on in either image will be present in the output image. The operations resulted in a single image which encompassed all the filters performed on the original image, as in Figure 2.5-2. An example of the original and final images after image analysis quantification of E-cadherin and is illustrated in Figure 2.5-2.
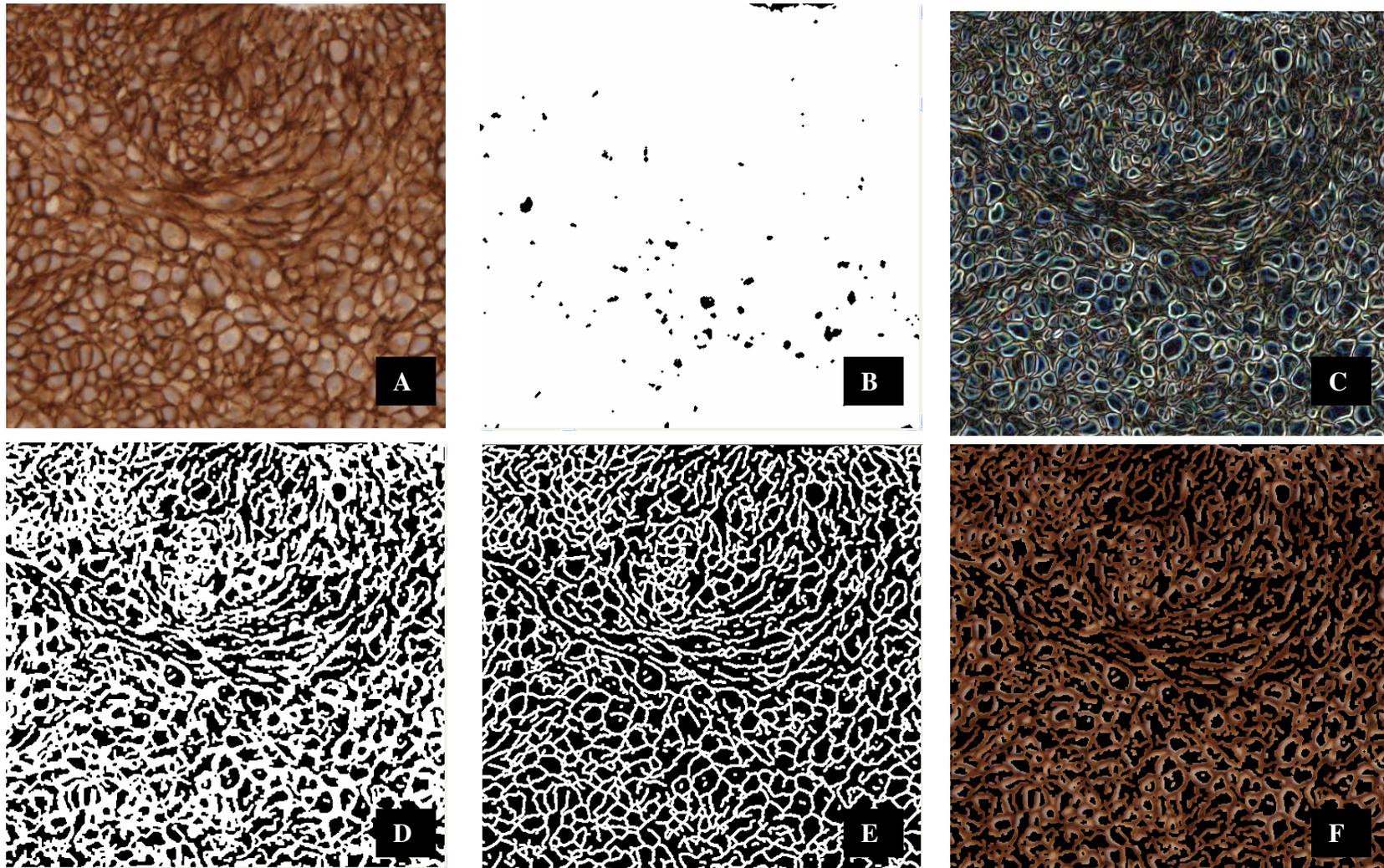
Figure 2.5-2: Various stages of the image analysis algorithm, which detects membrane-bound E-cadherin immunohistochemical staining (A) Original image, (B) Mask of the colour cube, (C) Sobel filter, (D) Well filter, (E) thinning and (F) final image of membrane staining. The tissue has been immunohistochemically stained for E-cadherin.

2.5.2 Distiller

Distiller (SlidePath, Ireland) is a web-enabled image informatics management system. It has been used as a replacement to the Virtual Tissue Matrix to facilitate virtual online reviews, store images and associated information. The Virtual Tissue Matrix is described in detail in Chapter 3. In addition, Distiller also provides the functionality of automated de-arraying. De-arraying is the process of fragmenting a Tissue Microarray image and locating each individual core, with minimal human intervention.

Due to file size restrictions it is not possible to perform high-throughput analysis across large images using Image-Pro Plus®. Therefore, the algorithm was recreated in Java and integrated into Distiller. Java code was created using ImageJ, an open source software library (Rasband, National Institutes of Health, USA). Netbeans was used as the development environment (Sun Microsystems Inc., USA). The creation of the algorithm in Java was performed by SlidePath engineers. Distiller provides a harness for manipulating virtual images. The harness divides images into tiles of 1000 X 1000 pixels, and performs image analysis on a tile by tile basis. From example, a 90 Mb whole section slide when divided into 1000 X 1000 pixels generates 721 tiles. Using the harness it is possible to output staining data per tile or per complete image. An example of the original and final images after image analysis quantification of ß-catenin is illustrated in Figure 2.5-3. A histogram of staining intensity values is generated from image analysis; an example of a histogram of staining intensity is illustrated in Figure 2.5-4.
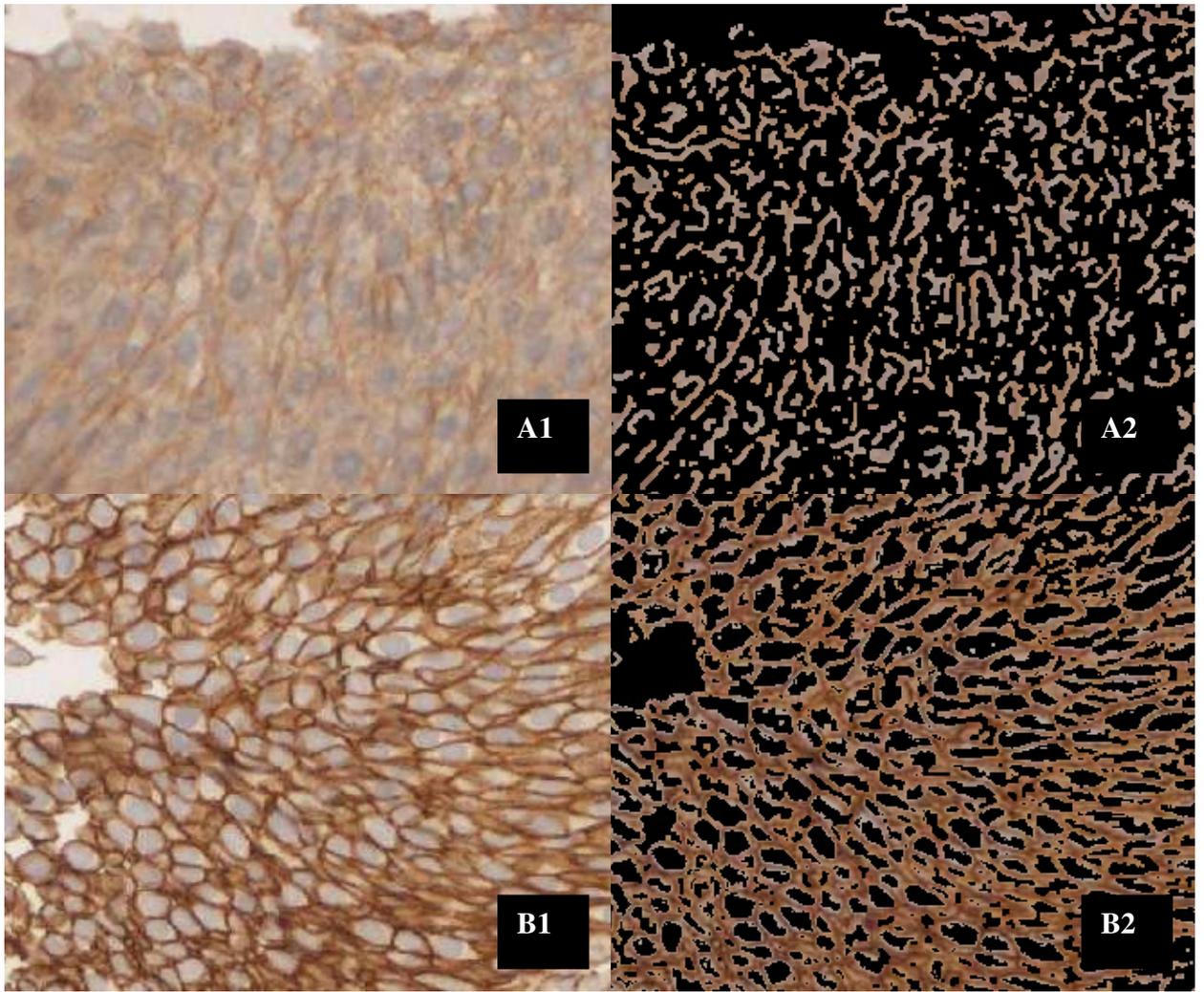
Figure 2.5-3: Original and Final images from the image analysis assessment of tissue immunohistochemically stained for ß-catenin. Images A and B, represent, Weak and Strong staining respectively. A1 and B1 are the original input images, and A2 and B2 are the final membrane stained only output images.
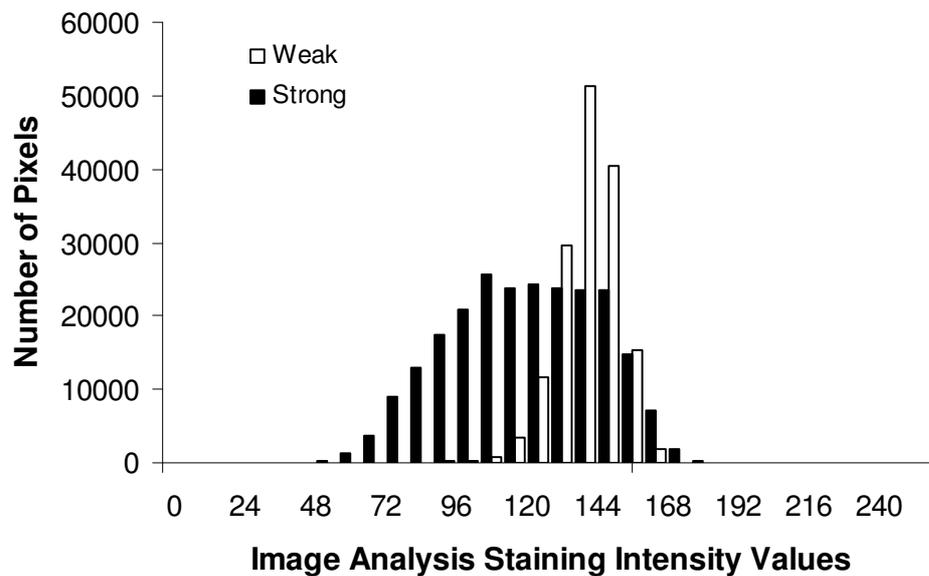
Figure 2.5-4: Illustrates the distribution of positive membrane-bound immunohistochemical staining observed within Figure 2.5-3, when assessed by image analysis.

After image analysis was performed, intensity values were uploaded into the Distiller database. Once in the database, intensity values were downloaded into Microsoft Excel™ and SPSS for statistical analysis. The fully automated image analysis workflow is depicted in Figure 2.5-5. DLL refers to dynamic link library; which is a collection of small programs that is called upon by a larger program running on the computer; CSV refers to comma-separated values, which is a format of data representation.
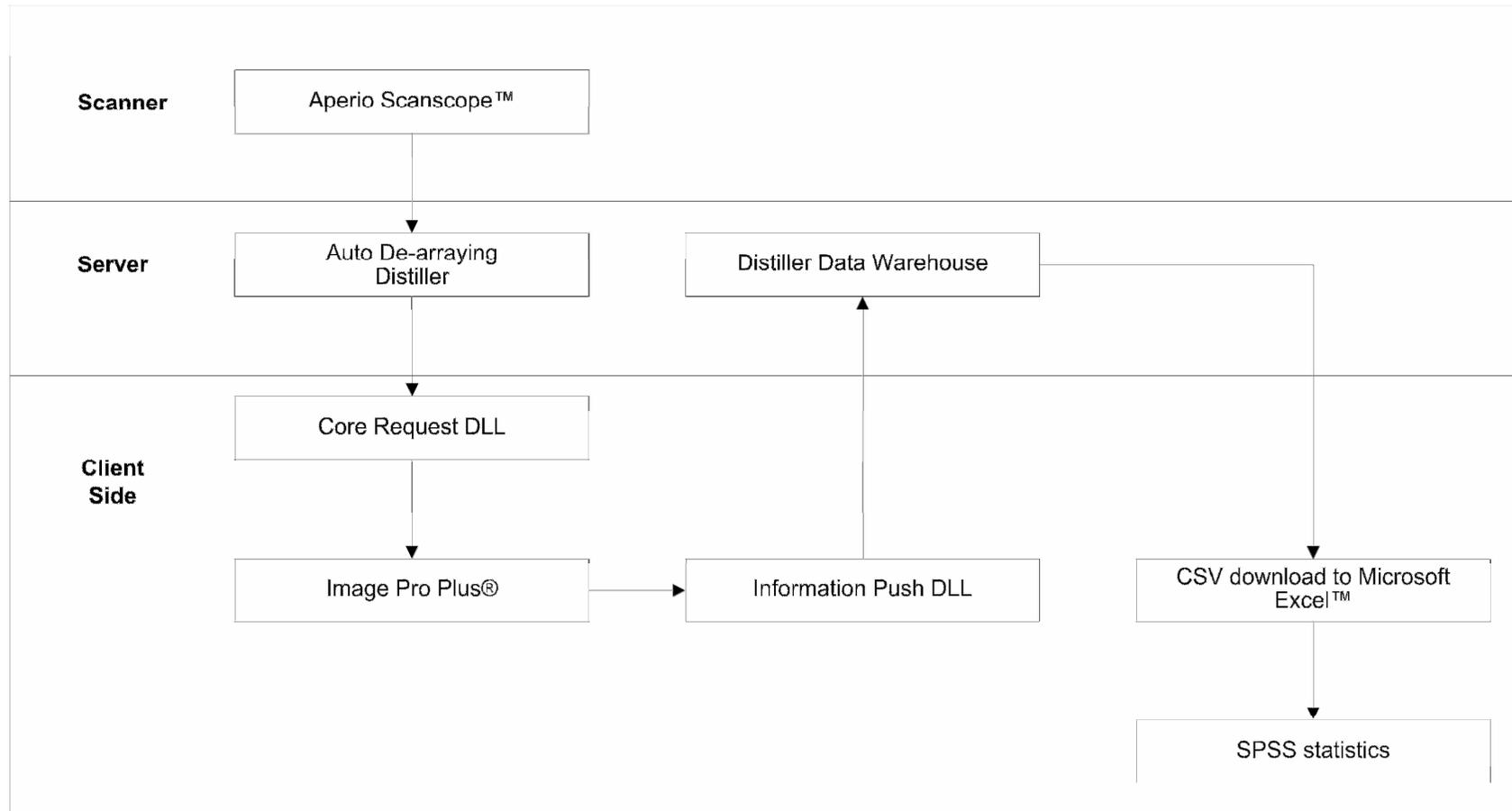
Figure 2.5-5: Fully-automated image analysis workflow facilitating digitisation, de-arraying and image analysis quantification of membrane-bound immunohistochemical staining intensity. In this example an Aperio Scanscope™ has been used to digitise the slides. However, the flow is independent from the scanning system used.

2.6   Interpretation of Image Analysis Results

Using image analysis the intensity of membrane staining was classified on a continuous scale, ranging from 0 to 255 staining intensity values, where 0 is strongest staining possible and 255 is the lightest staining intensity possible. The staining intensity value was calculated as the mode of greyscale intensity for all the "membrane" pixels. The mode was calculated by automatically generating a histogram of recorded membrane intensity values per Tissue Microarray spot, or whole section biopsy using an eight point intensity interval. The mode was recorded as the value that appeared most often in the histogram and, therefore, the most commonly occurring pixel intensity value. The average of the modal value from replica cores was used to calculate the membrane staining intensity value for each patient within the Tissue Microarrays. In addition, the number of positive membrane pixels, tissue pixels, and background lighting intensity was also recorded.

2.6.1   Normalisation

To account for variance in immunohistochemical staining across multiple slides, and variance from slide-to-slide background lighting intensity, all staining intensity mode values were normalised. Figure 2.6-1 illustrates the process of normalising image analysis data. Utilising image analysis, staining intensity values were quantified for all tissue cores (Tissue Core Value) and control cores (Control Core Value) present across each Tissue Microarray. In addition, using image analysis the background lighting intensity of each Tissue Microarray was also recorded (Slide Background Lighting Value). As a result of normalisation, staining intensity data was converted into natural logarithm values (Ln). During whole section analysis and where control cores were not present slides normalisation was performed by obtained the natural log of the Modal staining intensity value / background lighting intensity (Ln(I/Io).

$$Normalised = Ln\frac{I}{Io}$$
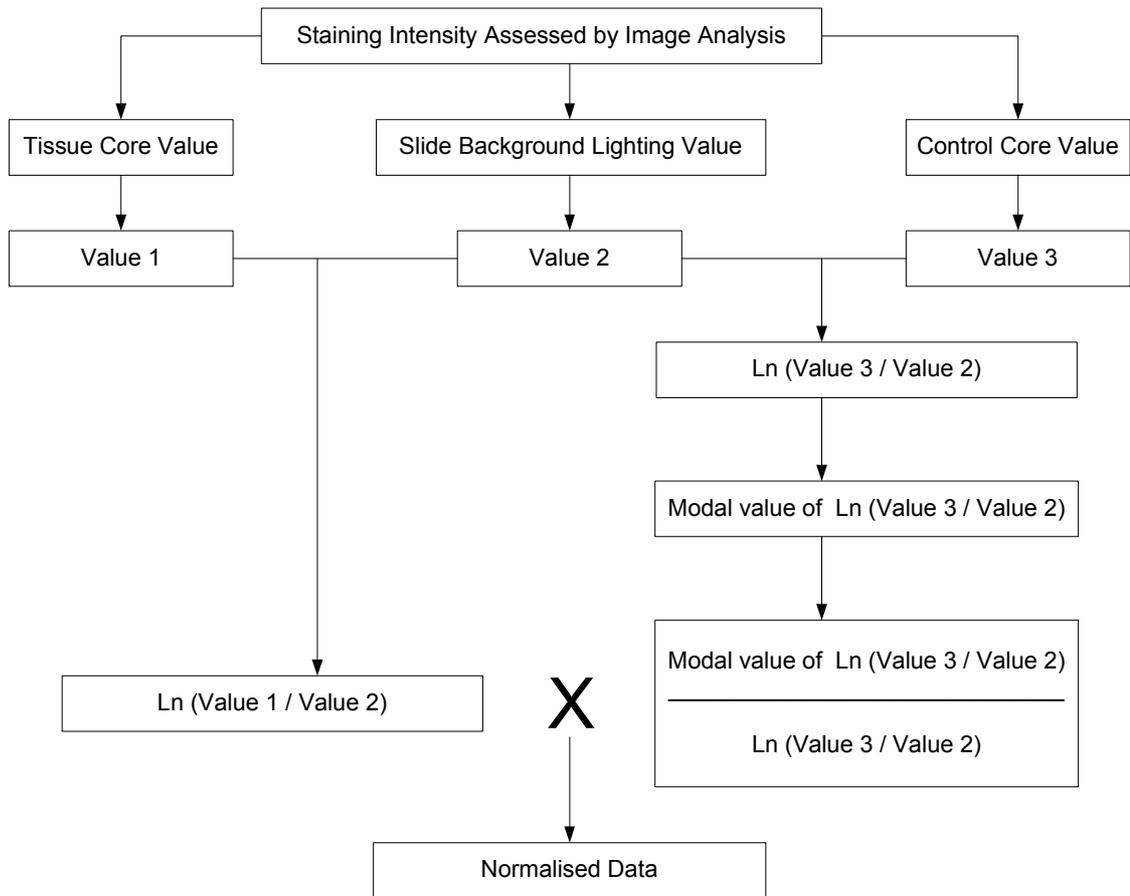
Figure 2.6-1: Process of normalising data generated from image analysis quantification of immunohistochemical staining. Results are converted into a natural log (Ln).

2.6.2   Generating Cut-Points

In order to compare continuous (image analysis results) and nominal datasets (human analysis), image analysis results where converted into categorical results. Using prognostic information and quantitative results from alternative assays it was possible to create optimal cut-points within image analysis datasets. The optimal cut-points were created as follows:

2.6.2.1 Prognostic Data

In theory, associations between tumour biomarker expression and patient outcome should reveal the existence of biologically meaningful tumour classifications; however, in practice there is no universal methods for discovering such associations (Camp et al., 2004). As a result, studies often group tumour into set divisions, for example, upper and lower quartiles, which fails to reflect the underlying biology of most biomarkers. Therefore, this study utilised X-tile which is a freeware software application that facilitates optimum cut-point selection of continuous data sets. X-tile plots provide a single, global assessment of every possible way of dividing a population into categories. As illustrated in Figure 2.6-2 , X-tile software allows the user to move a cursor over a histogram of continuous data (C) and provides on the fly subsets of populations with an associated Kaplan-Meier curve (D) and statistical analysis data (B), whiles displaying the distribution of cases (A) (Camp et al., 2004).

As it is statistically invalid to test multiple divisions and accept the best P-Value, within large cohorts of patients, X-tile provides the functionality of dividing the data set into "training set" and then validating results in a separate patient cohort, known as the "validation set". The application of X-tile in prognostic evaluations is expanding in literature (Camp et al., 2004, McCabe et al., 2005, Dolled-Filhart et al., 2006, Zerkowski et al., 2007, Rajput et al., 2007).

Statistics						Chi-Square	6.1563	
Pt No	% Total	Events	Rate	Rank	Range	Uncorrected P	0.01278	
34	23.45	12	35.29	0 to 25	0.56 thru 0.80			
111	76.55	19	17.12	26 to 103	0.80 thru 1.71			
145	100.00	31	21.38	0 to 103	0.56 thru 1.71			
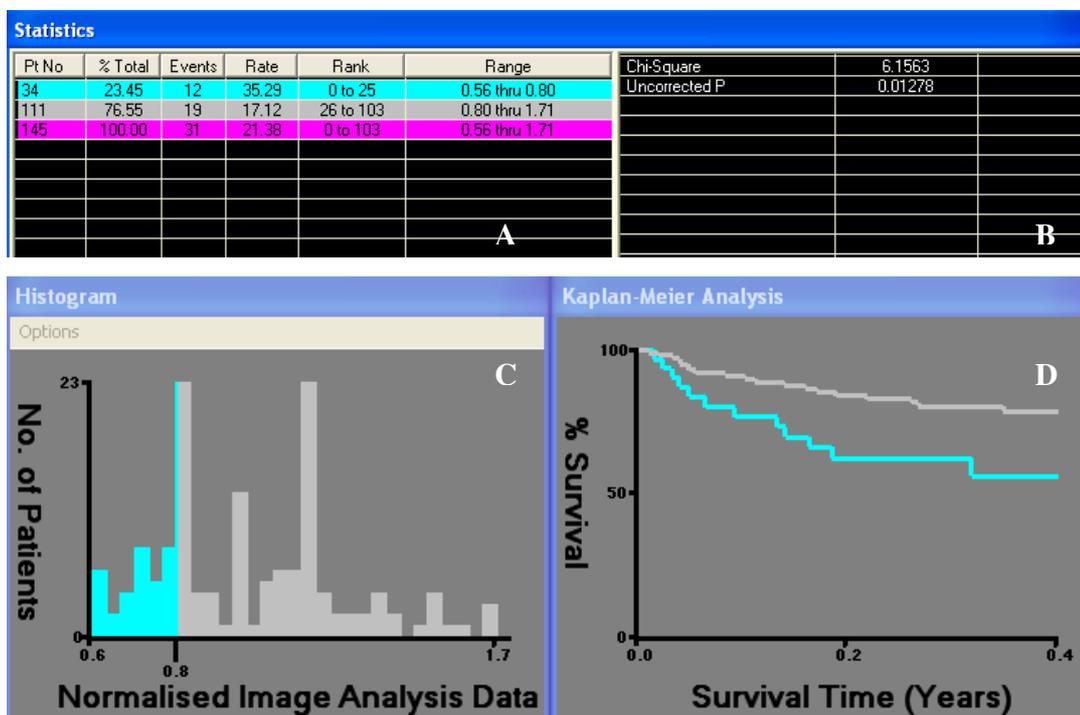
Figure 2.6-2: X-tile application selecting optimum cut-point of continuous data sets. X-tile provides a distribution of cases (A), statistical values (B), histogram of results (C) and Kaplan-Meier curves (D).

2.6.2.2 Quantitative / Qualitative Assays

Optimal cut-points can be generated by comparing image analysis results with data obtained from alternative quantitative assays performed on the same tissue set. Staining intensity values were quantified using image analysis, and the normalised image analysis results were compared with protein overexpression data quantified by Western blots/ELISA or gene amplification data quantified by FISH. The output from image analysis was categorised into positive or negative for protein overexpression, based on the modal staining intensity value and the number of positive membrane pixels.

SECTION C: STATISTICAL ANALYSIS METHODS

2.7 Statistical Analysis

Statistical analysis was performed using SPSS version 14.0 Windows (Chicago, IL, USA). ANOVA, Tukey HSD tests, T-tests and Pearsons Chi squared were used to compare human and image analysis assessments. Where applicable, confidence intervals (CI), degrees of freedom (df), and relative risk (RR) are included in the results. Statistical significance for all tests was set at $p < 0.05$. Statistical Analysis was predominantly used to illustrate levels of correlation between multiple reviewers, assays (see section 2.7.1), and measuring time related events (see section 2.7.2).

2.7.1 Cohen's Un-weighted Kappa

Cohen's un-weighted kappa values was one of two methods used to quantify the level of agreements achieved when comparing two datasets generated from human and image analysis reviews. Kappa was not the primary statistics used in the dataset comparisons involved in this body of work; however; Kappa has been included as it is widely used in comparisons of observer agreement of this type. Landis and Koch kappa interpretation scale was used to evaluate the level of kappa agreements; the interpretation scale is illustrated in Table 2.7-1 (Landis and Koch, 1977). However, within this body of work, the predominant method used to quantify the levels of agreement achieved when comparing two datasets generated from tissue / cell line reviews was the percentage of cases where the two datasets were in complete agreement.

Table 2.7-1: Interpretation of Landis and Koch 1977 kappa values. Ranges of Cohen's Kappa values are described within the table.

Kappa Value	Interpretation
0-0.2	Slight
0.2-0.4	Fair
0.4-0.6	Moderate
0.6-0.8	Substantial
0.8-1.00	Almost Perfect

2.7.2   Survival Statistics

Overall survival analysis was assessed by Kaplan-Meier curves with statistical significance assessed by Mantel-Cox log rank test. Kaplan-Meier curves were used as a univariate test to assess the rate of death due to Urothelial cell carcinomas of the Bladder. Kaplan-Meier curves facilitate the analysis of patients with varying lengths of follow-up information. Therefore, patients can be mathematically removed from the curve at the end of their follow-up period. Patients were deemed uncensored if they died of the disease under assessment during the follow-up period. Patients that were alive or dead due to other cancers or causes within the follow-up period were censored. The interpretation of Kaplan-Meier curved is illustrated in the following example.

Figure 2.7-1 illustrates a typical Kaplan-Meier curve. Within this curve, the effect of two treatments on a particular disease is assessed. The two lines represent treatment 1 and 2. Censored patients are signified within the curves by the following symbol (+). A decline within the curves represents a death due to the disease under assessment. The depth of decline within the curve signifies the total number of deaths at that time point. The y-axis represents the percent survival of patients, and the x-axis presents the time of survival under assessment.   Within this example, approximately 59% of all patients receiving treatment 2 were alive after the 5 year follow-up period, whereas, 80% of patients were alive after receiving treatment 1.

Kaplan-Meier Curve
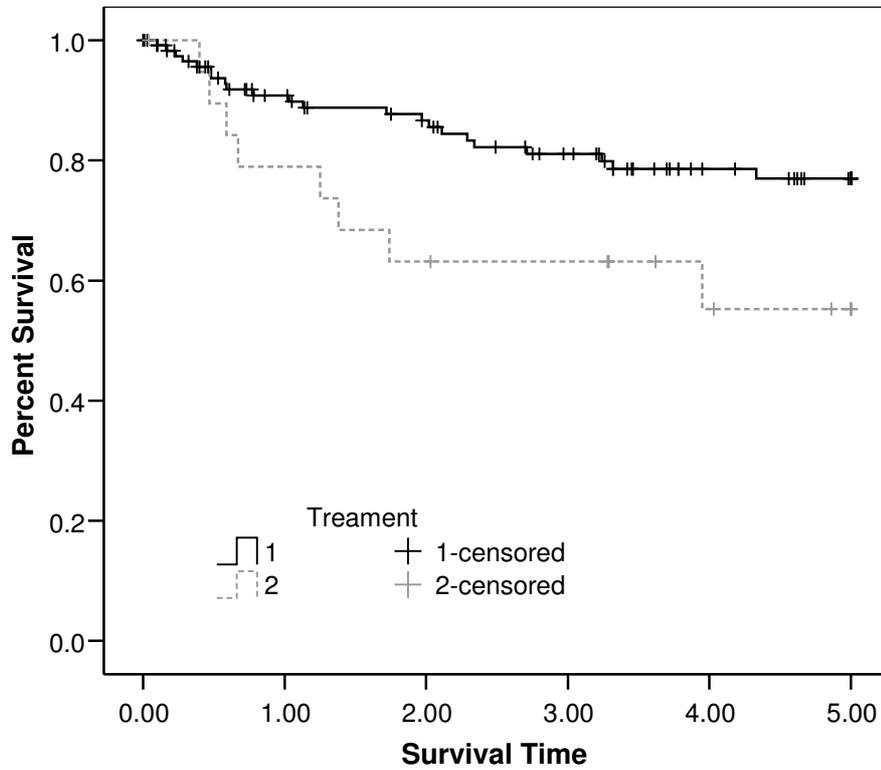
Figure 2.7-1: The effect of two different treatment types on disease is represented in this Kaplan-Meier curve. Survival time from disease under assessment is illustrated on the X-axis. Percent survival from the disease under assessment is illustrated on the Y-axis.

CHAPTER 3: DESIGN AND VALIDATION OF THE VIRTUAL TISSUE MATRIX; A SOFTWARE APPLICATION WHICH FACILITATES ONLINE REVIEW OF TISSUE MICROARRAYS

3.1 Introduction

Tissue Microarrays provide high-throughput analysis of tissue samples for *in situ* hybridisation and immunohistochemistry, by means of arranging multiple tissue samples in a uniform structure on the surface of a glass slide. Tissue Microarrays allow for large numbers of tissue samples to be analysed simultaneously at DNA, RNA or protein level.

Kononen et al. (1998) first illustrated the use of Tissue Microarrays in 1998 (Kononen et al., 1998). The technique involves the excision of cores of varying diameter (0.6 mm to 2.0 mm) from regions of histological importance on donor tissue blocks and the subsequent insertion of these excised cores into precise co-ordinates on a recipient block. This process is repeated until a two-dimensional matrix of cores is inserted into the recipient block. Once the block is complete, sections can be cut from the block, which are then available for any analysis currently performed on full-face tissue sections.

A large amount of data is associated with Tissue Microarrays, ranging from information on the tissue (patient information), to their construction, subsequent staining and assessment. It is becoming apparent that applications to assist in pathologist's reviews of Tissue Microarrays are required, as bottlenecks in the storage and manipulation of the data generated are beginning to emerge.

There have been previous attempts to create software applications that facilitate review of Tissue Microarrays (Sharma-Oates et al., 2005). The technology has varied from using Microsoft Excel™ spreadsheets, to the creation of complex databases. Manley et al. (2001) developed a relational database to store data and images, which focus on clinical outcome (Manley et al., 2001). This system consisted of several databases to store Tissue Microarray images, Tissue Microarray information and pathological and clinical information, in Microsoft Access™. All data was manually entered into a main online form and then transferred into the relevant database table. Each image was scanned using a grid structure that overlaid the image of the array. The images were

composed of six separate 10 x fields, stitched together to form a single image, which was saved as a JPEG image (200-300 kb). However, despite the advances made by this system, rapid file sharing over the internet was limited by large image sizes with slow internet connections. Also, as the software utilised were commercial applications, adaptations to the functionality of the programs were not possible.

Liu et al. (2002) utilised a combination of commercial and in-house applications to store data, digitise images and perform statistical analysis. Information was stored in Microsoft Excel™ spreadsheets and reformatted by a program called Tissue Microarray deconvoluter, into a structure that can be further manipulated to allow statistical analysis and hierarchical clustering. Although Microsoft Excel™ spreadsheets are traditionally used by scientists to store data, there is always a significant risk of human error, as large amounts of data entry are required and the object-oriented nature of the data does not lead to optimal data storage in spreadsheets (Liu et al., 2002).

The ability to interpret, review and grade histology in Tissue Microarray images across the Internet was assessed by Bova et al. (2001). This study evaluated the reviewer's ability to interpret images of Tissue Microarray cores, in order to assess the presence or absence of prostate cancer and to Gleason grade tumours. In 99% of cases, the images were deemed interpretable; this was done by visual inspection. However, it was suggested that, on occasion, lengthy downloading times would limit the system's practical use. The authors recommended that compressed files of less than 200 kb should be evaluated for viable image quality, as using files of this size would reduce downloading times. Inter- and intra-observer variability was found to be no greater, and in some cases less, than that reported when using traditional microscope-based Gleason grading. This was evaluated by comparing inter- and intra-observer variability observed in their study of online analysis with those previously reported in literature for inter- and intra-observer of Tissue Microarray microscope-based analysis. When evaluating levels of inter- and intra-observer agreement, the authors used percentage of complete agreement, and Cohen's un-weighted kappa coefficient of agreement. They concluded that web-based technology was an acceptable means to review Tissue Microarrays. The

authors believed that a limitation of their study was that web-based analysis was not directly compared with their microscope-based review and so recommended that this technology be tested using data resulting from immunohistochemical and *in situ* hybridisation reviews (Bova et al., 2001).

The advent of virtual slides permitted the review of whole tissue slides across the Internet (Johnston et al., 2005). Virtual slides provide users with all the functionality of a microscope, but with numerous additional benefits, including concurrent access for multiple users, tracking of review movements and image annotation. Virtual slides are reminiscent of microscope use and they are favoured by pathologists over static digital images, due to the ability to change magnification and scroll laterally while reviewing the image.

The aim of this chapter was to develop and validate a software application that would combine the benefits of virtual slides and online relational database technology, to facilitate Tissue Microarray reviews and scoring via the Internet. Unlike existing applications, the software system developed provides a relational database for data storage and smaller image tiles for rapid image access across the internet. To validate the software system developed, a study was created to ensure the system could achieve comparable results to those obtained from traditional glass slide analysis. The study examined users' ability to agree when performing virtual-based and microscope-based (glass) reviews of six immunohistochemically stained Tissue Microarrays, across eight parameters.

3.2   Study Design

3.2.1   Patient Cohort

Forty eight bladder tumours were used in this evaluation. Urothelial cell carcinomas of the Bladder (UCB) were selected from the files of the Mater Hospital histopathology department. A total of 6 Tissue Microarrays were constructed, with 48 cases and 196 2.0 mm cores, including 12 control spots composed of liver tissue applying the technique pioneered by Kononen et al. (1998). Tissue Microarrays were probed with three antibodies, E-cadherin (Novocastra/Vision Biosystems, UK), ß-catenin (Labvision Corp., USA) and PhosphoMet (Cell Signalling Technology, Inc., USA). The immunohistochemical staining procedure has been described in Chapter 2 section (2.3.1.3). The antibody to E-cadherin shows membrane staining; the antibody to ß-catenin shows membrane and cytoplasmic staining (Wijnhoven et al., 2000) and the antibody to PhosphoMet shows membrane staining, cytoplasmic staining and nuclear staining (Moran et al., 1997b). Tissue Microarrays were digitised using an Olympus BX–40 microscope (Olympus, USA) incorporating a Prior H101 motorised stage, the technology used was described in Chapter 2 section (2.4.1).

3.2.2   Validation of Image Quality

Tissue Microarrays were initially scanned at 4 x to create a tiled 'thumbnail' image of the entire array. This overview image was used to locate cores manually through a custom Graphical User Interface. The user clicked on the centre of each core on the overview image and the coordinate generated was used to seed an automatic scanning algorithm for all cores at 20 x. The array of captured images (6×8) were then tiled together to form a montage bitmap image of approximately 60.5 Mb (4607×4592 pixels).

The Macromedia Flash application, Zoomify™, was used to display images within the Virtual Tissue Matrix (VTM) framework. Zoomify™ Droplet, a software tool provided by Zoomify™, uses the original scanned BMP image as an input and converts it into a

set of JPEG image tiles. This tileset, once uploaded to a webserver, can be displayed via the Internet using the Zoomify™-embedded object within a conventional web page. Zoomify™ initially presents the user with a low power view of each Tissue Microarray spot. The users can then scroll around the image, and when required, zoom into a maximum magnification of 20 x.

When using a lossy compression algorithm, such as JPEG, image quality is reduced when compared with that of uncompressed images. To ensure that the compression rate used by Zoomify™ provided images of sufficient quality, a series of consultations with pathologists and scientists were performed where compressed and uncompressed images were compared. The outcome of this consultation was that the images generated by Zoomify™ were suitable for scoring. Figure 3.2-1 illustrates the quality of images available in the VTM.
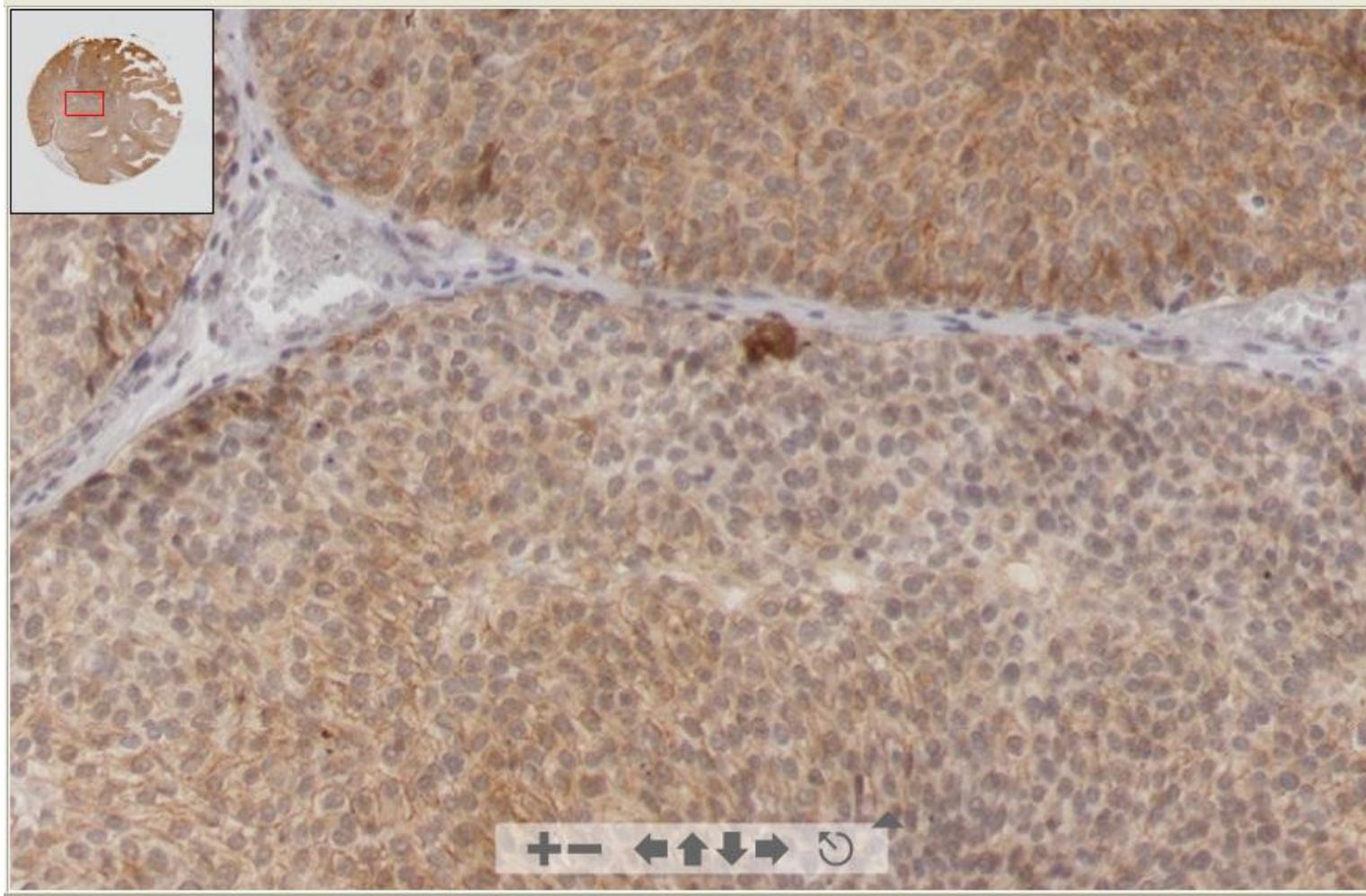
Figure 3.2-1: Digital image of Tissue Microarray spot presented in VTM using Zoomify™ application. On the top left corner of image is the thumbnail overview; the red box identifies selected location within the spot. The key at the bottom of the image allows the user to change position or magnification, which can also be controlled by the cursor.

3.2.3   Design Phase of the VTM

The design objectives of the VTM system were to provide Tissue Microarray images of sufficient quality to review over the World Wide Web, to present scoring forms to record Tissue Microarray results and to create a relational database that can store and, subsequently, retrieve data gathered during scoring.

PHP, Javascript, HTML and Oracle were used to create the VTM. PHP is a server side scripting language, which creates dynamic web pages, through embedding PHP code in HTML pages. Through the use of SQL queries, it can also extract data from many conventional databases (Oracle, MySQL, etc). Javascript adds interactive client side functionality to otherwise static HTML pages. An Oracle relational database was used to store all the information generated in this study.

3.2.4   Database Design

The construction of Tissue Microarrays facilitates the generation of hundreds of Tissue Microarray slides from a single Tissue Microarray block; therefore, every Tissue Microarray slide produced can potentially be stained with a unique immunostain. A major benefit of storing Tissue Microarray images and results within a relational database is the ability to extract all the results associated with a single core, which potentially may have been immunostained hundreds of times. This functionality is available within the VTM.

Based on analysis of conventional Tissue Microarray datasets, it was established that ten groupings of data would be sufficient to record all relevant information; therefore, ten tables where created in the VTM relational database. Each table contains information relating to a specific aspect of a Tissue Microarray review. For example, the USER table contains information relating to users only; SCORE table contains information relating to the Tissue Microarray analysis results only.   A unique identifier interlinks all tables and, by using SQL statements, information can be retrieved from multiple tables simultaneously. For example, results relating to a specific user can be obtained by

creating an SQL statement that requests information from the SCORES and USER tables. Table 3.2-1 lists all the tables that exist in the database and examples of the information they record. A complete schema of the database is illustrated in Figure 3.2-2.

Table 3.2-1: Tables within the VTM database and examples of their content.

Table Name	Information held	Examples
User	User details	E- mail address
TMA Manufacturer	Manufacturer details	Name
TMA	Tissue Microarrays	Diagram of cores
Patient	Patient information	Gender
Biopsy	Biopsy details	Biopsy notes
Core	Core details	Tissue type
Spot	Spots in a slide	Magnification scanned
Experiment Slide	Experiment procedure	Treatment name
Score	Results	% of Nuclear staining
Intensity Table	Staining intensity details	Staining type

Patient
Patient details
Patient ID
Date of birth
Gender
Diagnosis
Treatment
Outcome
Patients notes
Date added

Biopsy
Biopsy ID
Biopsy name
Biopsy notes
Patient ID
Date added

Core
TMA ID
Core ID
Core name
Biopsy ID
Tissue type
Core notes
X location
Y location
Data added

User
User ID
Fname
Sname
Email
Field of work
Username
Password
Ipaddress
Time of reg
Privileges

Score
Score ID
Spot ID
User ID
Ts notes
Time of exam
Nuc pos/neg
Mem pos/neg
Cyt pos/neg
Nuc percent pos
Mem percent pos
Cyt percent pos
Nuc total average
Nuc total range
Mem total average
Mem total range
Cyt total average
Cyt total range
Scoring scenario

Spot
TMA ID
Slide ID
Core ID
Spot ID
Image URL
Image Thumbnail
URL
Magnification

TMA
TMA ID
TMA name
Manufact. ID
X no. of cores
Y no. of cores
Diam. of cores
Diam. of cores units
TMA notes
Input times

Experiment Slides
TMA ID
Slide ID
Treatment notes
Scoring notes

Intensity Table
Score ID
Staining type
Category start score ID
Category finish score ID
Category name
Staining intensity average
Staining intensity range

TMA Manufacturer
Manufact ID
Name
Address
Manufact. info

Figure 3.2-2: Complete Schema of the VTM database.

The database has been designed to eliminate replication of data input. For example, a manufacturer of Tissue Microarrays details are only entered into the database once regardless of the number of Tissue Microarrays they have constructed. When their account is created they can be associated with multiple blocks/slides and studies. There is also the ability to add information into the database after the initial study has been created. For example, if patient or biopsy information is not known at the time of entering the Tissue Microarray review results into the database, they may be entered at a later date by simply entering the patient or biopsy information and then selecting the cores to which the information applies. However, users are unable to overwrite the Tissue Microarray review results already present in the database.

Depending on the internet connection used (times described here are for LAN settings of 10 Mbps), it takes approx 5 seconds to download the thumbnail overview of a Tissue Microarray slide; once a core has been selected it takes approx 2-3 seconds to view the individual core in the Zoomify™ window. There are no restrictions on the number of images that can be displayed within the VTM, currently there are 196 images. However, the number of images displayed is dependent on the capacity of the server available to the users.

3.2.5 Interface Design

The user interface had to be easily navigated, interpretable and provide images at sufficient speed and resolution for review. There are two types of user account within the VTM, namely administrator or user. A user has restricted access to the site; they can only review and score images. An administrator has additional privileges within the site; they can create new studies, add new images and scoring forms and they can view all data stored in the database.

The VTM was constructed so that the user is lead through the site, not having to concentrate on the sequence of events, freeing up analysis time for the reviewing process. Figure 3.2-3 illustrates the options available to both user and administrator within the VTM.
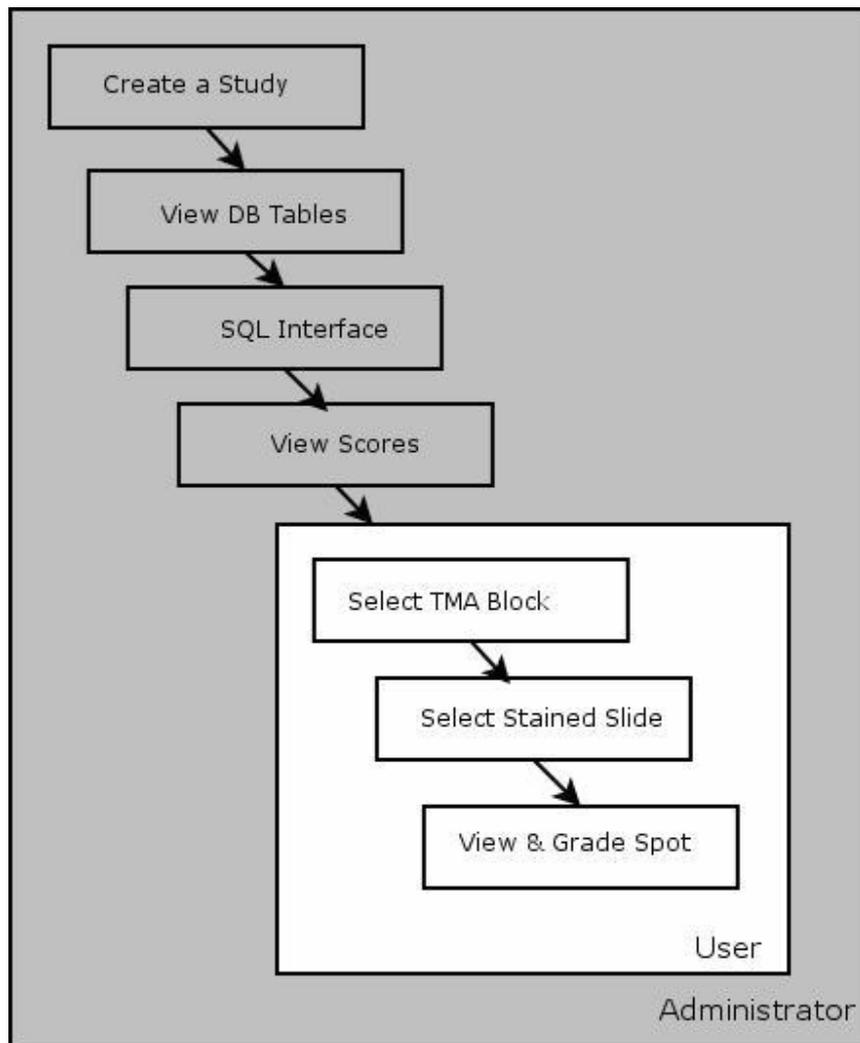
Figure 3.2-3: Options available through the VTM interface. Schematic diagram showing options available to administrators and users within the VTM interface. DB: Database, and SQL: Structure Query Language.

Figure 3.2-4 illustrates the user's view of a Tissue Microarray slide before a specific spot is selected. A spot must be clicked on to view a scaleable image. Once selected, the spot is reviewed in a pop-up window and the option to record results via a scoring form is presented. Figure 3.2-5 illustrates the spot with the associated scoring form. The administrator has the ability to create new scoring forms, depending on the user's specifications.
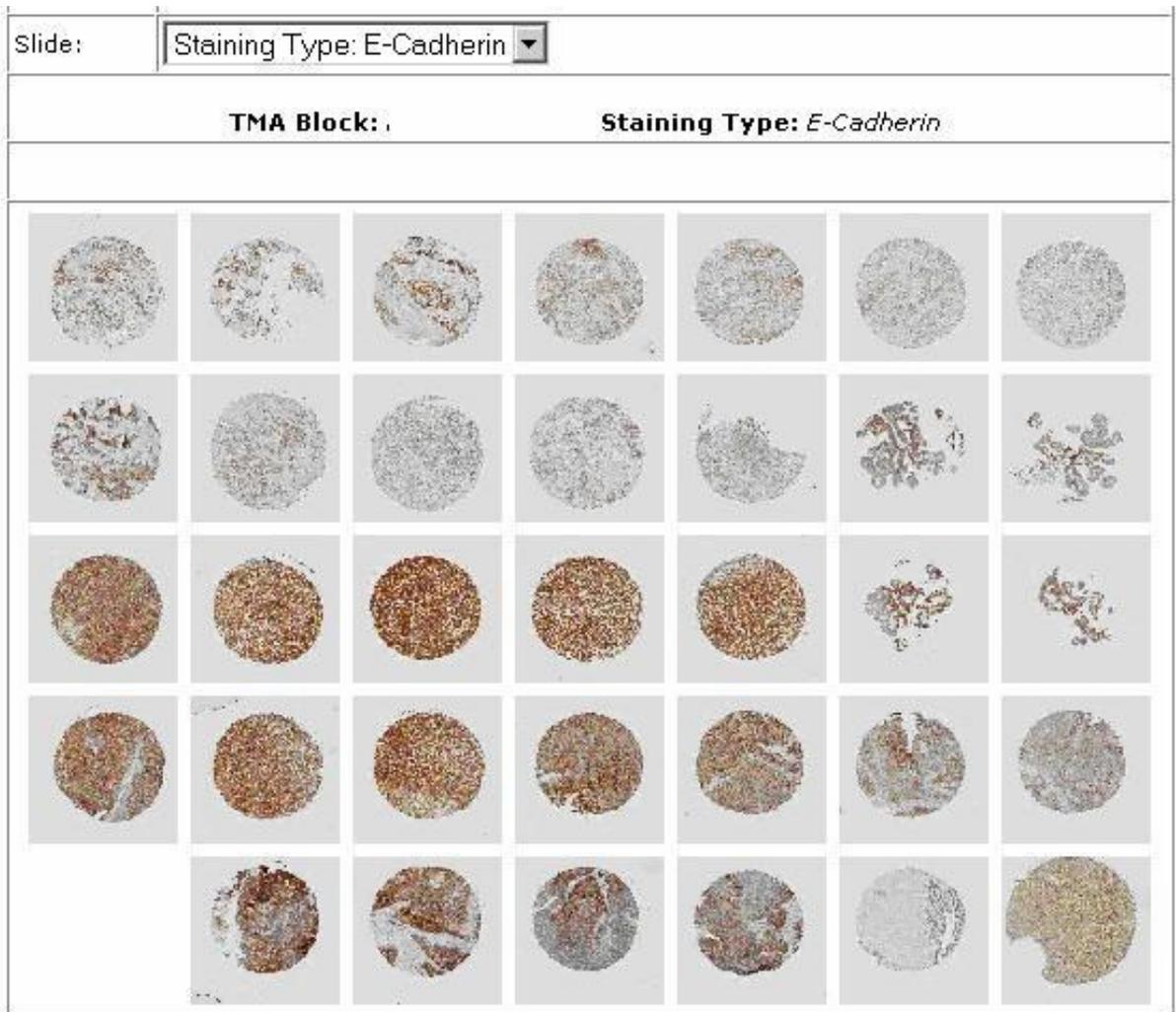
Figure 3.2-4: Overview of digital Tissue Microarray slide as presented in the VTM interface. Clicking on any spot will result in an enlarged version of the spot being provided, as in Figure 3.2-5.

Figure 3.2-5: Scoring form presented to users within the VTM interface. Results can be entered into the scoring form on the left, the image can be magnified and scrolling is possible via the controls provided.

3.2.6   Users, Scales and Parameters recorded

Two pathologists and one research scientist scored the immunohistochemically stained Tissue Microarray slides. One of the former was an external examiner, completing a virtual review of one slide from an international location. This review was conducted to ensure the VTM functioned in a remote location and to ensure that a user who had no contribution to the design phase could use the VTM software system. As a result of having only one data set from this user, their observations were not included in the results that follow.

Review methods included a microscope-based review (glass slide) and a virtual slide review. The glass slide review was a traditional microscope-based process; the virtual review was performed over the World Wide Web using the VTM. Eight parameters were examined for each spot. The amount and intensity of membrane, cytoplasmic and nuclear staining, as well as the amount of core and tumour present in each spot, were recorded. A five-point scale was used to record the staining intensity. Within the five-point scale, grade 0 represented no staining, grade 1 was weak staining, grade 2 was moderate staining and grade 3 was strong staining. Grade 4 was included to record variable staining; this option was only available when scoring membrane staining.

A four-point scale was used to record the amount (%) of staining present in tumour-containing regions of each spot. Grade 0 represented no staining, grade 1 represented 1-30% of relevant cells staining, grade 2 represented 31-50% staining and grade 3 represented greater than 50% staining. An eleven-point scale was used to record the amount of core and tumour present. Grade 0 was negative/no tissue, and grade 1 to 10 increased in 10% increments to 100%.

3.3   Results

Two reviewers (Users A and B) examined 183 Tissue Microarray spots (196 minus 12 control spots and one un-reviewed spot) stained with 3 immunostains using both review methods (microscope and VTM). Neither user reported any technical difficulties when performing virtual or glass analysis. The levels of intra- and inter-observer variability between virtual and microscope-based (glass) Tissue Microarray reviews were assessed for parameters examining the amount of core and tumour present, the amount and intensity of membrane, cytoplasmic and nuclear staining. Table 3.3-1 illustrates the intra-observer variability between virtual and glass Tissue Microarray reviews. Good levels of agreement (> 60%) between methods were observed when quantifying the amount of core present, the amount of membrane, nuclear and cytoplasmic staining and nuclear intensity. Low levels of agreement between methods were observed when quantifying the amount of tumour present and membrane intensity.

Virtual vs. Glass was the average of two users' agreements when comparing virtual with glass Tissue Microarray reviews. User A vs. B Virtual was User A virtual Tissue Microarray review compared with User B virtual Tissue Microarray review. User A vs. B Glass was User A glass Tissue Microarray review compared with User B glass Tissue Microarray review.

Table 3.3-1: Agreement levels (%) and un-weighted kappa values by measured parameter for each comparison of Tissue Microarray reviews. Agreement levels which are extremely low have been highlighted in grey.

Parameters	Intra-Observer		Inter-Observer			
	Virtual vs. Glass		User A vs. B Virtual		User A vs. B Glass	
	% Agreement	Kappa	% Agreement	Kappa	% Agreement	Kappa
% Core Present	71.3	0.507	67.1	0.414	67.4	0.448
% Tumour Present	47.3	0.407	37.9	0.283	33.0	0.244
% Membrane Staining	81.2	0.577	78.2	0.553	78.6	0.520
% Cytoplasmic Staining	64.0	0.373	42.8	0.198	65.9	0.280
% Nuclear Staining	85.0	0.226	84.9	0.524	95.8	0.239
Membrane Intensity	32.4	0.167	25.3	0.153	77.1	0.447
Cytoplasmic Intensity	58.2	0.306	50.8	0.263	65.2	0.234
Nuclear Intensity	82.1	0.109	83.3	0.474	89.2	0.234

Table 3.3-1 illustrates the inter-observer agreement achieved when performing both virtual and glass Tissue Microarray reviews. Inter-observer agreements achieved when performing virtual Tissue Microarray reviews were comparable with inter-observer agreements achieved when performing glass Tissue Microarray reviews, for four out of eight parameters. The virtual Tissue Microarray review of the amount of tumour present achieved greater levels of inter-observer agreement than the glass Tissue Microarray review, of this parameter; however, the level of agreement between users for this parameter was very low for both methods of assessment. The glass Tissue Microarray review of the amount of cytoplasmic staining and intensity, and membrane staining intensity achieved greater levels of inter-observer agreement than the virtual Tissue Microarray review, of these parameters.

The parameter, % Tumour present, was difficult to assess, with low agreement recorded for inter- and intra-observer agreements across virtual and glass Tissue Microarray reviews. As this parameter has equivalent levels of inter-observer variability for both virtual and glass Tissue Microarray reviews, it has to be assumed that poor performance is not based on the method, but more likely the size of the range used. A 10% interval was used to quantify the amount of tumour present, which appears to have been too restrictive for this parameter.

Reviewers found quantification of the parameter membrane intensity difficult to reproduce between virtual and glass Tissue Microarray reviews. When users assessed membrane intensity using a microscope, they predominantly used two categories of staining intensity, negative and moderate. With the virtual Tissue Microarray review, users appeared to use the classifier more extensively, as illustrated in Figure 3.3-1. The assessment of intensity of immunohistochemistry, particularly the intensity of membrane-bound immnunostains is inherently difficult. For example, problems with HER-2 assessment by immunohistochemistry are well documented (Hoang et al., 2000, Ellis et al., 2005).
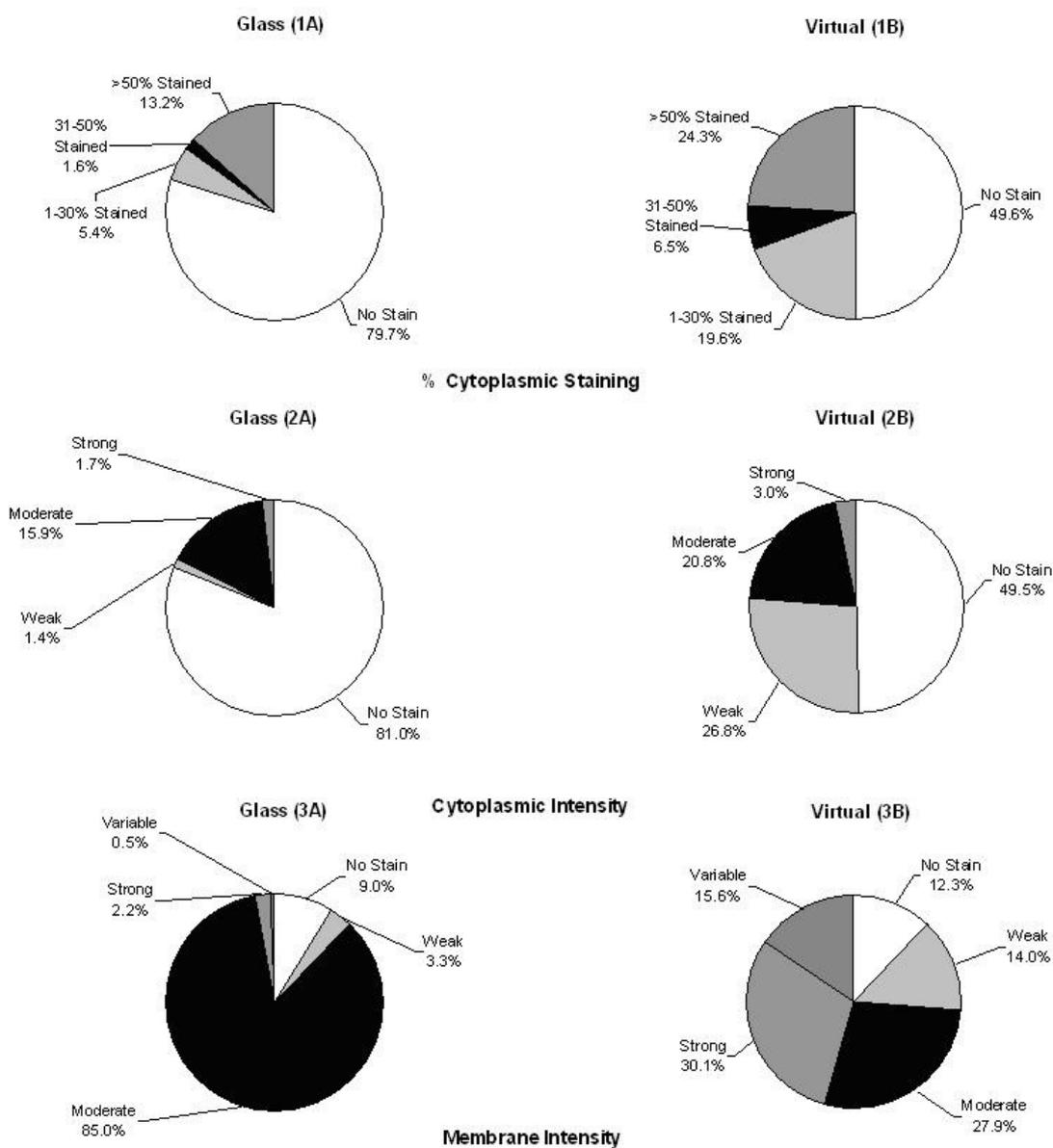
Figure 3.3-1: Distribution of the results for virtual and glass TMA reviews of cytoplasmic / membrane staining intensity and % cytoplasmic staining. Illustrates the distribution of the classifiers when using virtual and glass methods to (1A) Amount of cytoplasmic staining assessed by glass TMA review (1B) Amount of cytoplasmic staining assessed by virtual TMA review (2A) Cytoplasmic staining intensity assessed by glass TMA review (2B) Cytoplasmic staining intensity assessed by virtual TMA review (3A) Membrane staining intensity assessed by glass TMA review (3B) Membrane staining intensity assessed by virtual TMA review.

3.4 Discussion

The objective of this chapter was to design and validate an online software application that presents Tissue Microarray images and stores associated review and clinical data. The result was the Virtual Tissue Matrix (VTM), which consists of Tissue Microarray images available at multiple magnifications, scoring forms to gather Tissue Microarray review data and a relational database to store the generated results.

The VTM displays virtual Tissue Microarray images via a web site and facilitates the storage of Tissue Microarray data via a relational database. There are numerous advantages of using the VTM over other proposed software systems of its type. Downloading of the images is rapid. Only views that are requested by the user are returned at maximum resolution, thereby downloading the minimum required dataset. The VTM was designed in consultation with scientists and pathologists and, as a result, the reviewing process emulates the workflow involved in conventional Tissue Microarray reviews. The VTM interface is delivered in HTML, via a conventional web browser, allowing for intuitive user interaction. The VTM database is relational; a structure more suited to the storage of the object oriented dataset generated from Tissue Microarray experimentation than previous efforts incorporating flat files and spreadsheets for data storage.

The ability to interpret virtual images using the VTM was validated, via assessment of inter- and intra-observer variability on two users' evaluations of immunohistochemically stained Tissue Microarrays, using digital and microscope analysis. Eight parameters were evaluated, the amount of core and tumour present, the amount and intensity of membrane, cytoplasmic and nuclear staining. Comparisons evaluated in this study illustrated that intra- and inter-observer virtual Tissue Microarray reviews produced equivalent levels of agreement as intra- and inter-observer glass Tissue Microarray reviews, for four out of the eight parameters examined. Where discrepancies occurred it was dependent on the parameters and users involved. In all comparisons, low levels of agreement for the amount of tumour present were observed. This was not surprising, as

the application of classifiers to any data continuum (data that does not naturally fall into discrete clusters) results in scoring variability around the interfaces of the classifier. This variability is increased when the number of classes are increased creating more interfaces. Also, of the two reviewers used, one was a scientist and one a pathologist. The scientist's accurate interpretation of tumour/non tumour may potentially be questioned as a result of this work.

Of particular interest were a large number of observations that were considered positively stained by virtual Tissue Microarray reviews, which were considered negatively stained when reviewed using a microscope. This was particularly evident when quantifying the amount of cytoplasmic staining, where virtual Tissue Microarray reviews observed substantially more positively stained spots than glass Tissue Microarray reviews. The additional positively stained spots were largely considered to stain between 1-30% of the tumour area and/or to be weakly stained. This suggests that virtual Tissue Microarray reviews may be more successful in allowing the identification of small areas of staining and/or where staining intensity is low.

One proposed reason for the identification of staining when using digital images that was not observed with a microscope was the use of correcting adjustments to the image data during the digitising of Tissue Microarrays. Bulbs used in microscopes have a characteristic tint; in general this is yellow or straw coloured. However, this tint is removed when digitising slides using a corrective algorithm, potentially unmasking weak staining that would otherwise be attributable to background tint. Also, with microscope-based analysis, background light is adjusted to best suit each individual spot. When digitising the slides for this study, a constant background light intensity was used to digitise all slides for this study.

Excluding nuclear staining, where positive staining was infrequent, agreement levels were low when examining staining intensity. When using a conventional microscope, in general, users failed to utilise all grades within the classifier to characterise positive

staining intensity; the category of moderate staining was repeatedly used when positive staining was observed, particularly for membrane staining. However, with digital reviews, all grades within the classifier were utilised more extensively, which suggests that the review of digital images gives a user more confidence to discriminate between different intensities and that subtle differences in intensity may be easier to detect when evaluating digital slides, than when using a microscope. This may be due to the standardisation in lighting while preparing the images.

Exceptionally low levels of inter- and intra-observer agreement for both methods of review were observed when assessing membrane staining intensity. Users had difficulty on agreeing on the presence and degree of membrane staining intensity when using a four grade scale, and when agreement was observed membrane staining intensity was negative. The level of agreement achieved for membrane staining intensity was also influenced by the immunostain assessed. Intra-observer agreements levels differed by 46.8% between users when comparing virtual and glass slide assessments of membrane-bound E-cadherin staining intensity. Due to the level of inter- and intra-observer variability recorded for the quantification of membrane-bound immunohistochemical staining intensity, it is assumed that the accurate quantification of this parameter is inherently difficult. Therefore, the application of a computer-assisted image analysis algorithm for the quantification of membrane-bound immunohistochemical staining intensity may be beneficial. Image analysis would be particularly applicable to immunostains that prove difficult to assess; in this case E-cadherin staining.

Human observers, while excellent at object classification, are inherently poor at quantifying intensities and areas to any degree of accuracy. Studies have shown that image analysis produces more reproducible results than pathologists for quantifying the intensity of staining, in relation to ß-catenin expression in Tissue Microarrays for colon cancer (Camp et al., 2002). Image analysis systems may identify subtle differences in staining intensity, which are not quantifiable by a human reviewer, thus leading to the better correlation of expression data to prognostic indicators.

Since the creation of the VTM there have been numerous advances in the technologies used for image acquisition (Gomez et al., 2006, Stelldinger and Kothe, 2006, Yasuda et al., 2005) and image analysis techniques and applications have been well documented in literature (Warford et al., 2004, Ho et al., 2006, Hansen et al., 2004, Patton et al., 2006, McCullough et al., 2004, Della Mea et al., 2005). Integrated intuitive systems such as Aperio or Dmetrix are now available that rely on minimal human intervention when scanning slides (Weinstein et al., 2004). Numerous commercial image acquisition applications are now available (Aperio Technologies Inc., USA: D.Metrix, USA: Applied Imaging, USA); however, cost of purchase is often high for these integrated systems, putting them out of reach for many research laboratories.

The validation of the VTM illustrated the need for online databases which integrate Tissue Microarray images and an associated database for information storage. Since the VTM was created, the software has been developed considerably and released as a commercially available software application called Distiller by a Dublin City University based company (SlidePath, Ireland). Distiller is not constricted to Tissue Microarray images or data. It is a fully configurable and customisable web-enabled image informatics management system for life sciences data. Distiller also provides statistical analysis capabilities. Due to the extra functionalities available within Distiller, it has been used in subsequent chapters as an alternative to the original VTM technology.

3.5   Conclusions

The Virtual Tissue Matrix (VTM) was created to assist in Tissue Microarray analysis, by providing digital Tissue Microarray images at multiple magnifications online and submitting Tissue Microarray review data from an online form into an associated database. The VTM illustrated that digital Tissue Microarray analysis obtained equivalent levels of agreement as microscope-based analysis, for four out of eight parameters. The remaining four parameters achieved greater levels of agreement when performed using microscope analysis. However, on further investigation of the four parameters, it is proposed that the digital reviews may be providing the user with greater capability to accurately assess staining presence and intensity. Results illustrate that users were incapable of agreeing when comparing digital and microscope Tissue Microarray reviews. Particularly when classifying staining intensity. Results illustrated that staining was identified when using digital images that was not observed with microscope-based reviews. It is suggested this is due to the background correction step involved in digitising the slides.

Comparisons of digital with standard glass reviews of immunohistochemistry stained slides is well documented in literature; however, in order to validate the VTM, it was necessary to perform this study. As previously reported in literature, there was some degree of inter- and intra-observer variability. However, the ability of users to observe more positive staining when performing digital reviews, and the inability of users to utilise all categories within the classifiers provided when performing glass reviews are previously unreported in literature.

Membrane staining intensity proved to be inherently difficult to assess, and high levels of inter- and intra-observer variability were observed for this parameter. The type of immunostain under assessment also influenced the level of inter- and intra-observer agreement. Inter- and intra-observer agreement was particularly low for E-cadherin staining, compared to the two alternative antibodies (ß-catenin and PhosphoMet). It is suggested that the introduction of a computer-assisted image analysis algorithm for the

quantification of membrane-bound immunohistochemical staining intensity may be beneficial, especially for immunostains that prove difficult to assess.

CHAPTER 4: CREATION OF CELL MICROARRAYS; PROTEIN PROFILING PERFORMED BY AUTOMATED IMAGE ANALYSIS

4.1 Introduction

Worldwide, lung cancer is the most common cancer in terms of both incidence and mortality with 1.04 million new cases diagnosed per year and 921,000 deaths, with the highest rates currently observed in Europe and North America. Once diagnosed, the 5-year survival rate for lung cancer in Europe fluctuates between 8-12% (Guessous et al., 2007). Breast cancer is a leading cause of cancer deaths in women all over the world, with approximately 12% of women directly affected by this disease (O'Driscoll and Clynes, 2006). Between 40-50% of patients diagnosed with breast cancer will eventually die of the disease, despite advance in treatment (Jemal et al., 2002). In order to decrease disease-related mortality, it is vital to determine the optimal treatment method for each individual patient and to identify subgroups of patients who might benefit from individualised treatment strategies. Identification of prognostic and predictive biological markers will facilitate this (Larkin et al., 2004).

Cell culture is an important research tool. It plays a vital role in assessing the characteristics of tumour cells, it helps us to determine how their growth can be manipulated by different culture conditions and also aids in the identification of potential therapeutic agents (Waterworth et al., 2005). The most commonly applied technique for the study of proteins in cultured cells are immunoprecipitation and Western blotting, which use protein-antibody interactions to qualify protein levels. These techniques are qualitative but do not provide information on the subcellular localisation of proteins, unless cell fractionation is performed first (Li et al., 2005). Immunocytochemical study of cultured cells has the advantage of directly visualising intact cells and enables localisation of the target proteins to subcellular compartments, but it is associated with many technical difficulties (Li et al., 2005). However, due to the heterogeneity of most tumours it is important to use an investigational technique such as immunohistochemistry, that can provide direct morphological confirmation of the presence of protein rather than "bulk" methods such as Western blotting (Moran et al., 1997b).

Tissue Microarrays facilitate gene expression and copy number surveys of large cohorts of tumours simultaneously (Kononen et al., 1998). Utilising Tissue Microarrays, disease-specific expression of a gene can be identified and, therefore, a possible diagnostic test and treatment against a specific target can be developed (Simon et al., 2003). In general, current knowledge regarding mechanisms underlying malignant transformation, differentiation, and other cellular processes originates from experiments performed on cell lines (Stromberg et al., 2007). Microarrays constructed from cell lines are an adaptation of the traditional Tissue Microarray format. Such Cell Microarrays (CMAs) are useful for rapid characterisation of the expression profiles of multiple cell lines relevant for cancer research. Cell Microarrays are powerful tools for the screening of new reagents, including hybridisation probes and antibodies. The standardisation of staining procedures and reduction of intra-assay variability can also be significantly improved with this technique (Hoos and Cordon-Cardo, 2001).

Image analysis algorithms, which quantify immunohistochemical staining intensity, can be developed using cell lines, rather than using complex heterogeneous tissue (Stromberg et al., 2007). The application of image analysis algorithms to quantify immunohistochemically stained cell lines has been described in literature (Pick et al., 2007, Brown et al., 2004, Stromberg et al., 2007, Messersmith et al., 2005, Dolled-Filhart et al., 2006). Automated Quantitative Analysis (AQUA) was used to compare protein expression levels of HSP90 across 10 cell lines and 655 human breast carcinomas. HSP90 protein expression is reported as being high in breast cancer cell lines; however, expression levels had not been previously reported in human tumours. Results illustrated that greater variability in HSP90 expression was seen in tumours than in cell lines. Subsequently, AQUA was used to evaluate the correlation of HSP90, ER, PR and HER-2 protein expression with breast carcinoma-specific survival across the entire patient cohort. Results indicated that HSP90 expression retained its prognostic significance in a multivariate setting. Within this study, utilising AQUA the authors identified a potential biomarker in breast carcinoma (Pick et al., 2007).

The prognostic value of ß-catenin in breast cancer has been assessed using AQUA (Dolled-Filhart et al., 2006). ß-catenin protein expression across 8 cell lines created into a Cell Microarray was assessed using ELISA and AQUA. A standard curve of protein expression was generated from ELISA and AQUA results. Comparison of AQUA and ELISA results showed a high level of correlation ($r^2 = 0.728$). The standard curve was subsequently used to quantitatively evaluate ß-catenin protein concentration in breast tumour Tissue Microarrays from AQUA results. The creation of the standard curve for ß-catenin expression was possible as the 8 cell lines selected showed a range of protein expression levels by ELISA. Using the standard curve model and AQUA, the authors concluded that ß-catenin is an independent predictor of patient outcome in breast cancer (Dolled-Filhart et al., 2006).

Quantitative assessment of protein expression across Tissue and Cell Microarrays composed of colorectal tumour and normal colon tissue was performed by Automated Cellular Imaging System (ACIS), and compared against cell line Western blotting, ELISA and human scoring by a pathologist (Messersmith et al., 2005). Cell Microarrays were immunohistochemically stained with 6 antibodies for 6 antigens (EGFR, pEGFR, Akt, MAPK and pMAPK). Similar protein expression levels were observed when assessed by ACIS and Western blots, for all 6 targets. However, protein expression was only compared for a single cell line. ELISA results from 8 cell lines indicated variable expression of EGFR was compared with ACIS analysis. A scatter plot of ACIS scores and ELISA results demonstrated good correlation between the two methods of assessment ($r^2 = 0.697$). A total of 99 normal and cancerous colon tissue samples were examined for total protein expression by ACIS and human analysis using a semi-quantitative scale (0-3+). A significant correlation was observed when comparing ACIS and human analysis scores for all six antibodies (ANOVA, $p < 0.001$) (Messersmith et al., 2005).

This chapter describes the creation of a Cell Microarray from 31 cell lines (lung and breast). Sections from the Cell Microarray were immunohistochemically stained with 6 antibodies. Slides were digitised and protein expression was quantified by image

analysis, human analysis and Western blot/ELISA, in order to validate the application of the low cost high-throughout algorithm for Cell Line analysis. To date, this is the largest cohort of cell lines which has been analysed by automated image analysis. Using image analysis, protein expression which was previously unpublished has been quantified. A brief description of the antigens under assessment within this Chapter follows:

A major problem encountered during chemotherapy treatment of many tumour types is the development, by the target tumour cells, of resistance to a broad spectrum of drugs, a phenomenon known as Multidrug Resistance (MDR) (Moran et al., 1997b). MDR occurs in approximately 84% of patients who receive chemotherapy (Larkin et al., 1999). Multidrug resistance is characterised by resistance to a broad range of structurally and functionally unrelated chemotherapeutic drugs including antracyclines, vinca, alkaloids, epipodophyllotoxins, and taxanes. Having prior knowledge of the presence of such resistance would decrease morbidity from unsuccessful therapy and allow for the selection of individuals who may benefit from the co-administration of MDR-inhibiting drugs (Kurdziel et al., 2007).

Several mechanisms of resistance have been identified including; disruption in cell signalling, alterations in drug targets or activation of de-toxifying systems. One of the mechanisms that may be clinically active in lung and breast carcinomas is the prevention of intracellular drug accumulation by the expression of transporter proteins that pump drugs out of cells before they can reach their sites of action. Several of these proteins belong to the ATP (adenosine triphosphate)-binding cassette (ABC) transmembrane protein superfamily that utilise energy from ATP hydrolysis to translocate substrates across cell membranes (O'Driscoll and Clynes, 2006, Leonessa and Clarke, 2003).

Multidrug resistance is mediated by P-glycoproteins (P-gp). P-gp belongs to the ATP-binging cassette (ABC) transmembrane proteins. P-gp is an energy-dependent pump, capable of transporting hydrophobic compounds out of the cell (Ambudkar et al., 1999). P-gp is frequently found to be overexpressed in post-chemotherapy tumours and its

expression is associated with decreased cellular uptake of chemotherapy drugs. Therefore, allowing cells to survive in the presence of higher drug concentrations. At least three P-gp genes have been identified in humans, MDR-1/P-gp, MDR-3 and MRP-1.

The Multiple drug resistance 1 (MDR-1) gene encodes P-glycoprotein a 170 kDa plasma membrane protein consisting of 12 transmembrane domains and 2 ATP-binding domains. MDR-1/P-gp is expressed in a large number of normal tissues including, gastrointestinal tract, brain, testes, ovaries, kidney, adrenal glands, large intestine and liver. The physiological role of MDR-1/P-gp has not been clearly defined; however, it may include detoxification and excretion of xenobiotics, as well as hormone transport (O'Driscoll and Clynes, 2006). Tumours arising from tissues that normally express MDR-1/P-gp may be intrinsically resistant to chemotherapeutic agents or, alternatively, tumours that were initially responsive to chemotherapy may develop multidrug resistance during the treatment regimen and subsequently not respond to therapy (Moran et al., 1997b). MDR-1/P-gp gene and protein expression in untreated breast carcinomas have been reported to range from 0-100% (Larkin et al., 2004, Leonessa and Clarke, 2003, Kroger et al., 1999). The role of MDR-1/pgp in lung cancer has not been widely reported in literature.

MRP-1 is a plasma membrane protein involved in the active efflux of cytotoxic materials from the cell (Moran et al., 1997b), the protein may play a role in the prediction of response to chemotherapy of breast cancer. MRP-1 is expressed by a considerable number of untreated breast tumours (range 10-100% by immunohistochemistry) (Leonessa and Clarke, 2003, Trock et al., 1997, O'Driscoll and Clynes, 2006). MRP-1 expression is an important predictor of poor prognosis in patients with recurrent breast cancer (Nooter et al., 1997). Both MDR-1/P-gp and MRP-1 were found to be frequently expressed in non small cell lung carcinomas (38 cases); however, MDR-1/P-gp was predominant over MRP-1 protein (Roy et al., 2007).

MDR-3 is thought to be the most highly conserved P-gp gene amongst various mammalian species (Larkin et al., 1999, Matthews et al., 2006, Morita et al., 2007, Russell et al., 1994, Swales et al., 2006). MDR-3 is highly expressed in liver and is thought to function as a hepatic transporter of phospholipids into bile. There is evidence to suggest that the overexpression of MDR-3 may be a potential prognostic indicator in chronic lymphocytic leukaemia (Matthews et al., 2006). However, MDR-3's role in breast or lung carcinoma, if any, remains undefined.

The epidermal growth factor receptor (EGFR), a member of the ErbB / HER family of receptor proteins, initiates a complex signalling cascade which influences many facets of tumour cell biology including growth, survival, and responsiveness to chemotherapy (Messersmith et al., 2005, Mendelsohn and Baselga, 2003, Baselga and Arteaga, 2005). EGFR protein is overexpressed in 16 - 48% of human breast carcinomas and up to 80% in non small cell lung carcinomas (Klijn et al., 1992, Tsutsui et al., 2002, Grenier and Soria, 2007). The validity of EGFR as a useful prognostic factor for breast cancer is still uncertain. The lack of standardized assessment methods and interpretation criteria for EGFR expression may contribute to the conflicting results in the prognostic role of EGFR expression in breast cancer (Park et al., 2007). However, in non small cell lung carcinoma, EGFR mutational analysis is an excellent predictor of responsiveness to treatment with tyrosine kinase inhibitors, such as gefitinib (Daniele et al., 2007). At least three anticancer agents targeting EGFR are already FDA-approved and several others are in development (Messersmith et al., 2005).

HER-2/neu is a member of the human epidermal growth factor (HER-2) receptor family. The overexpression of this gene has been found to contribute to oncogenic transformation, tumourgenesis and metastatic potential and it is reported to be a negative prognostic factor (Tawfik et al., 2006). HER-2 is amplified and overexpressed in 10–30% of women with breast cancer (Slamon et al., 1987, Slamon et al., 1989, Borresen et al., 1990). HER-2 as a prognostic indicator in breast carcinomas has been extensively investigated in literature. A humanised monoclonal antibody (Herceptin[TM]) targeting the oncogene shows therapeutic benefits especially when combined with conventional

chemotherapeutic agents in treatment of patients with metastatic breast carcinoma (Camp et al., 2003, Slamon et al., 1987, Seidman et al., 2001, Winer and Burstein, 2001, Baselga et al., 1996, Carter et al., 1992). HER-2 amplification is observed in 2-23% of non small cell lung carcinomas (Swanton et al., 2006). HER-2 gene amplification and protein overexpression are a indication of poor prognosis in non small cell lung carcinoma (Szelachowska et al., 2006, Zeng et al., 2006).

E-cadherin is a transmembrane glycoprotein involved in intercellular adhesion (Mahnken et al., 2005, Bukholm et al., 1998). E-cadherin mediates the selective adhesion of epithelial cells that is required for the interaction and maintenance of normal integrity (Takeichi, 1991). Reduced expression of the E-cadherin adhesion complex in lung cancer cells is associated with dedifferentiation, increased invasiveness, metastasis, and poorer survival (Bremnes et al., 2002b, Shibanuma et al., 1998, Sulzer et al., 1998). Loss of E-cadherin expression also correlates with poor prognosis in breast carcinoma (Jeschke et al., 2007). Reduced expression levels of E-cadherin has been observed in invasive lobular carcinoma, but not in invasive ductal carcinoma of the breast (Turashvili et al., 2007).

4.2   Study Design

Cell Microarrays were created from 31 cell lines (lung and breast). The creation and immunohistochemical staining of the Cell Microarrays is described in detail in Chapter 2 section 2.3.1. Protein expression was quantified by human/image analysis and compared with Western/ELISA data. See Chapter 2 section 2.3.2 and 2.3.3, respectively. Cell Microarrays were digitised at 20 x magnification using an Aperio ScanScope T3 Scanner™ (Aperio Technologies, USA). Images were approximately 30 Mb per Cell Microarray.

4.2.1.1 Human Analysis

Human analysis of the immunohistochemically stained cells was performed by one reviewer, utilising Distiller (SlidePath, Ireland). Protein expression was categorised as positive or negative. Cases that were deemed positive by human analysis were further assessed by image analysis.

4.2.1.2 Image Analysis

Image analysis categorised staining intensity on a continuous scale between 0 and 255, were 0 is the strongest staining possible and 255 is the weakest staining possible. Subsequently, cut-points were created to classify staining intensity as Weak (> 140), Moderate (84 - 140) and Strong (< 84). Cut-point selection was based on human interpretation of image analysis data.   Normalisation was not performed on image analysis data, as inter-slide comparisons were not performed. Therefore, the variance of background light or staining procedure did not affect results.

4.2.2   Sample Cohort

In total, this study consisted of 186 tissue spots immunohistochemically stained for 6 antibodies (31 cell lines x 6 antibodies). However, a number of spots were lost or damaged due to tissue processing (n = 14). Therefore, 172 cases were available for analysis. Protein expression, as quantified by Western blot (n = 39) or ELISA (n = 23),

was available for 62 cases in total. The distribution of samples available for analysis per antibody is illustrated in Table 4.2-1.

Table 4.2-1: Distribution of cell lines available for immunohistochemical analysis and protein expression results available when assessed by Western blot /ELISA.

Antibody	Cell Lines	Western blot	ELISA	Total Assays
E-cadherin	22	2	0	2
EGFR	31	2	10	12
HER-2	30	4	13	17
MDR-1	29	15	0	15
MDR-3	30	4	0	4
MRP-1	30	12	0	12
Totals	172	39	23	62

4.3 Results

Cell Microarrays composed of 31 cell lines were immunohistochemically stained with six antibodies targeting MRP-1, MDR-3, MDR-1/P-gp, HER-2, EGFR and E-cadherin proteins. Protein expression was qualified by human analysis of immunohistochemistry, Western blots, ELISA and image analysis. The level of agreement between human analyses of immunohistochemically stained cell lines and Western blot/ELISA was assessed. Across this cohort of cell lines, previously unreported protein expression was determined utilising high-throughput image analysis.

4.3.1 Comparison of Protein Expression

The expression levels of six proteins across 31 cell lines were assessed utilising immunohistochemistry, Western blots and ELISA. Immunohistochemistry was assessed visually by a single reviewer (the author) using Distiller. Protein expression recorded by the three classification methods were categorised as positive or negative. The level of agreement between immunohistochemistry and Western blots/ELISA assessments is illustrated in Table 4.3-1. Overall agreement between immunohistochemical reviews and Western blot/ELISA was 70.96%. The highest levels of agreements were observed when HER-2 and EGFR protein expression was assessed. There was no agreement observed when E-cadherin protein expression was assessed. However, only two Western blots were available for comparison. Generally, the majority of disagreement arose when protein expression was considered positive by Western blots/ELISA methods and negative when assessed by immunohistochemistry. In addition, disagreement between Western blot and ELISA results was observed for one case out of a total 6 cases where results were available for both assays (83.33% agreement). In this one case of disagreement between Western blot and ELISA, image analysis results correlated with ELISA results.

Table 4.3-1: Complete agreement between protein expression levels when quantified by immunohistochemistry and Western blots/ELISA. Disagreement between review methods occurred when protein expression was observed by immunohistochemical analysis and not Western blot/ELISA (IHC Pos+) or protein expression was observed by Western blot /ELISA and not immunohistochemical analysis (Western blot /ELISA Pos+).

Antibody	% Agreement	Cases Compared	IHC Pos+	Western blot /ELISA Pos+
E-cadherin	0.00	2	-	2
EGFR	91.66	12	1	-
HER-2	88.23	17	1	1
MDR-1	60.00	15	1	5
MDR-3	75.00	4	-	1
MRP-1	50.00	12	-	6
Totals	70.96 (Overall)	62	3	15

Figure 4.3-1 and Figure 4.3-2 illustrate cell lines immunohistochemically stained. For all cases illustrated in Figure 4.3-1 and Figure 4.3-2 the level of protein expression differed when assessed by immunohistochemistry and Western blot/ELISA. E-cadherin protein expression was not observed for DLKP Parent and DLKP VCR when assessed by immunohistochemistry (Figure 4.3-1). However, both of the cell lines apparently expressed E-cadherin protein when assessed by Western blots. Figure 4.3-2 illustrates the immunohistochemical staining of BT20 and DLKP Txt YL with HER-2 and EGFR, respectively. BT20 and DLKP Txt YL expressed HER-2 and EGFR protein, respectively, when assessed by immunohistochemistry. However, no expression of these proteins was observed when BT20 and DLKP Txt YL were assessed by Western blot and ELISA, respectively.

DLKP VCR – E-cadherin

DLKP Parent – E-cadherin
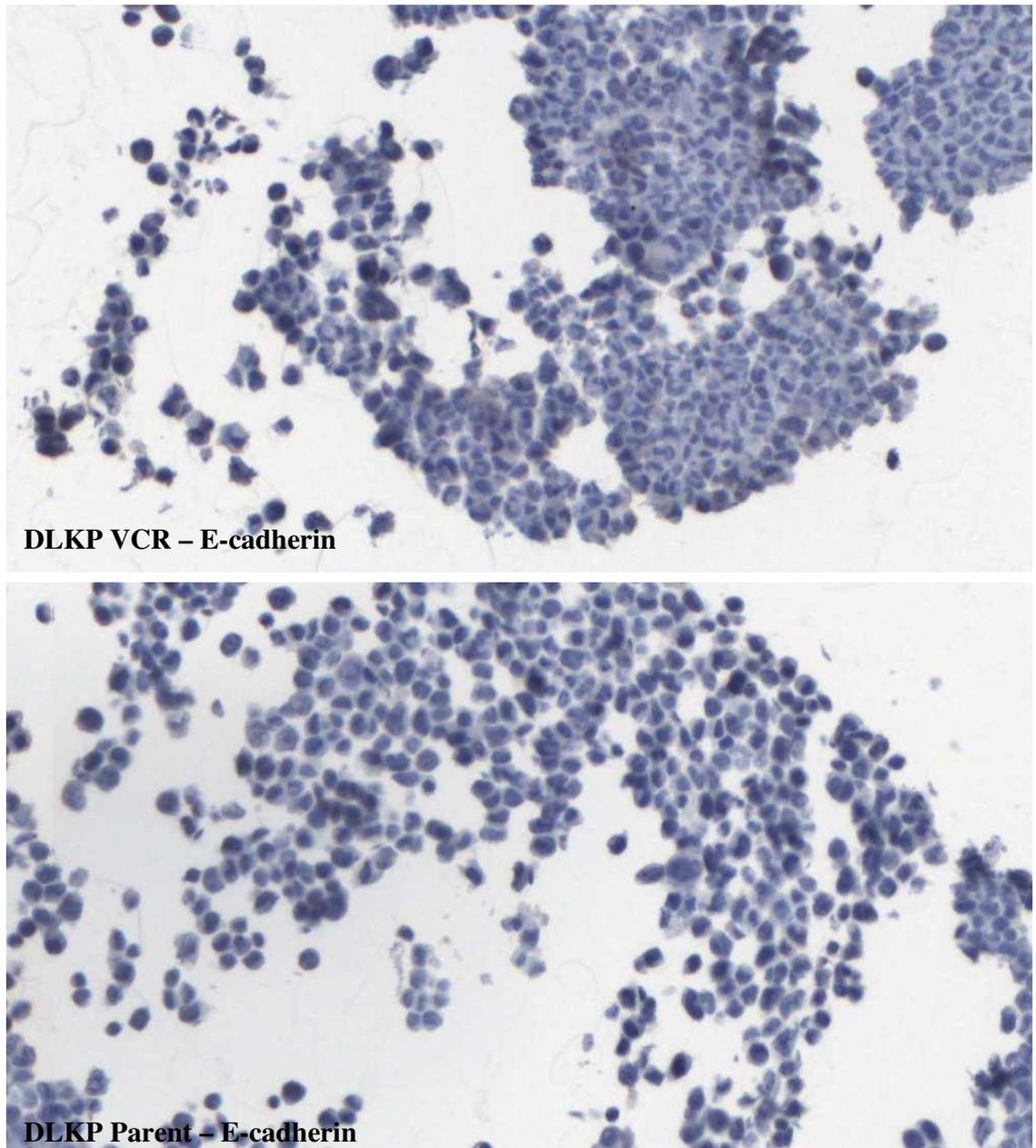
Figure 4.3-1: DLKP VCR and DLKP Parent do not express membrane E-cadherin when assessed by immunohistochemistry. However, when assessed by Western blot/ELISA protein expression was observed.

DLKP Txt YL - EGFR

BT20 - HER-2
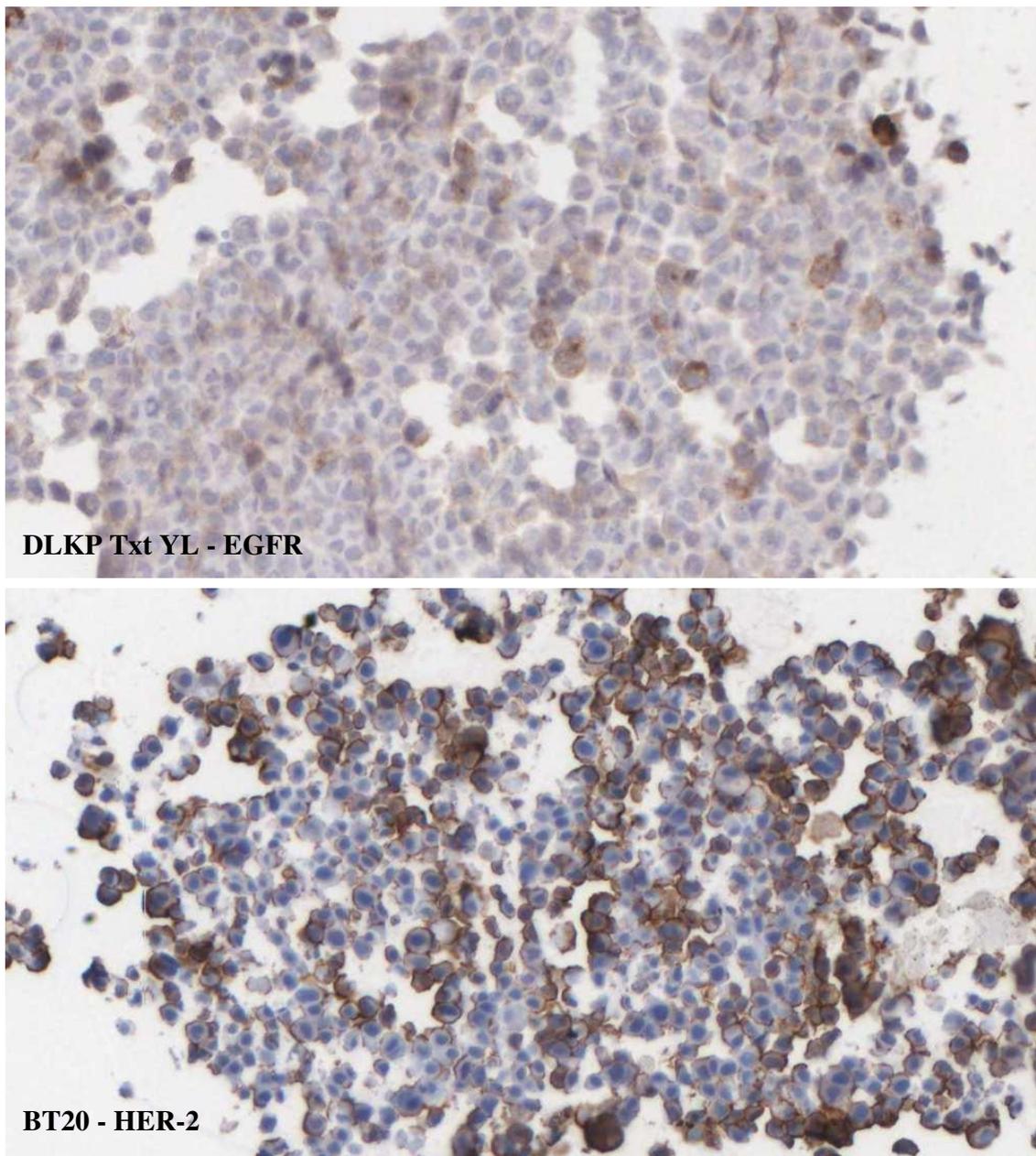
Figure 4.3-2: DLKP Txt YL and BT20 expressing membrane EGFR and HER-2, respectively, when assessed by immunohistochemistry. However, neither cell lines express EGFR or HER-2 proteins when assessed by Western blot or ELISA.

4.3.2 Protein Expression assessed by Image Analysis

Table 4.3-2 illustrates membrane expression levels for 6 membrane proteins across the 31 cell lines, when assessed by image analysis. Cases where immunohistochemical assessment and Western blot/ELISA differed have been noted within the Table. Staining intensity was categorised as Weak (> 140), Moderate (84-140) or Strong (< 84), based on cut-offs generated by comparisons with human analysis of immunohistochemical staining. The majority of positive staining was classified as weakly stained. There were a number of cell line cores which were lost due to processing and are noted in the Table as (-). The majority of protein expression illustrated in Table 4.3-2 is previously unpublished. Cases where Western blot or ELISA data was available for comparisons are shaded within Table 4.3-2.

Table 4.3-2: Protein expression of 6 membrane proteins across 31 immunohistochemically-stained cell lines. Cases where immunohistochemistry disagrees with Western blot/ELISA have been noted within the Table. Cases were protein expression was observed when assessed by Western blot/ELISA methods, and not during immunohistochemical analysis are denoted as [1] within the table. Cases were protein expression was observed when assessed by immunohistochemistry and not during Western blot/ELISA analysis are denoted as [2] within the table. Cases were Western blot/ELISA results were known are shaded within the table.

Cell Lines	MRP-1	MDR-3	MDR-1/P-gp	HER-2	EGFR	E-cadherin
A549 Parent	Negative [1]	Weak	Negative	Negative	Strong	Negative
A549 Taxol	Negative [1]	Weak	Negative	Negative	Moderate	Negative
BT20	Weak	Weak	Weak	Moderate [2]	Weak	-
BT474	Negative	Weak	Negative	Moderate	Negative	Moderate
DLKP Parent	Negative [1]	Weak	Negative	Negative	Negative	Negative [1]
DLKP JC	Negative	Negative	-	Negative	Weak	-
DLKP Txt RL	-	-	-	-	Weak	-
DLKP Txt YL	Weak	Weak	Negative [1]	Negative	Weak [2]	Negative
DLKP VCR	Negative [1]	Weak	Negative	Negative	Weak	Negative [1]
DLKP YL	Negative	Weak	Negative	Negative	Weak	-
DLRP Parent	Negative	Negative	Negative	Negative	Moderate	-
DLRP Carboplatin	Negative	Weak	Negative	Negative	Strong	-
DMS 53 Parent	Negative	Negative [1]	Negative	Negative	-	Negative
DMS 53 Taxol	Negative	Negative	Negative [1]	Negative	Negative	Negative
DMS 53 Taxotere	Negative	Negative	Negative [1]	Negative	-	Negative
DMS 53 Carboplatin	Weak	Negative	Weak	Negative	Negative	Negative
H460 Parent	Negative [1]	Weak	Negative	Negative	Weak	-
H460 Taxol	Negative [1]	Weak	Negative	Negative	Moderate	-
HCC 1419	Negative	Negative	Negative	Weak	Negative	Moderate

HCC 1954	Negative	Weak	Negative	Strong	Weak	Moderate
MCF-7	Weak	Weak	Weak	Negative [1]	Negative	Moderate
MDA-MB-231	Negative	Negative	Negative	Negative	Moderate	Negative
MDA-MB-453	Negative	Weak	Negative	Weak	Negative	Negative
MDA-MB-468	Negative	Negative	Negative	Weak	Negative	Negative
NCI-H1299 Parent	Negative	Weak	Negative [1]	Negative	Weak	Negative
NCI H1299 Taxol	Negative	Weak	Negative [1]	Negative	Weak	Negative
SK-BR3	Negative	Weak	Negative	Moderate	Weak	Negative
SKLU-1 Parent	Negative	Weak	Negative	Negative	Strong	Negative
SKLU-1 Taxol	Weak	Weak	Weak [2]	Negative	Weak	Negative
SK-MES-1 Parent	-	-	-	Negative	Weak	-
SK-MES-1 Carboplatin	Weak	Weak	Negative	Negative	Weak	Negative

4.4   Discussion

Thirty-one lung and breast cell lines were cultured and created into a Cell Microarray. Sixteen cell lines were uniquely developed within the NICB. Sections were cut and immunohistochemically stained with 6 antibodies targeting cellular membrane proteins. Immunohistochemical staining was qualified by human analysis and compared to expression data obtained from Western blots and ELISAs. Positive membrane staining across all Cell Microarrays was classified using high-throughput automated image analysis.

When membrane staining intensity was qualified by human analysis, agreement levels between Western blot/ELISA and immunohistochemistry varied greatly depending on the antibody assessed. Agreement levels between methods were lowest when assessing the P-glycoproteins (MRP-1, MDR-1/P-gp and MDR-3). No agreement was observed between methods when assessing E-cadherin expression. However, only two cell lines were compared for E-cadherin expression (DLKP Parent and DLKP VCR). No E-cadherin protein expression was observed for either of the cell lines when assessed by immunohistochemistry (Figure 4.3-1); however, protein expression was observed for both cell lines when assessed by Western blots. In this instance, literature would suggest that immunohistochemical analysis was correct when qualifying E-cadherin expression in DLKP Parent (Liang et al., 2004); however, protein expression for DLKP VCR is unreported in literature. On average, agreements between immunohistochemistry and Western blot/ELISA were high when assessing protein expression with the 6 antibodies.

The majority of disagreements between Western blot/ELISA and immunohistochemistry occurred when protein expression was observed by Western blot/ELISA and not observed when assessed by immunohistochemistry. For a number of cases, the positivity observed when assessed by Western blot/ELISA was also reported in literature (DLKP Parent - *MRP-1*, DLKP VCR - *MRP-1*, DLKP VCR - *MDR-1*, DLKP Parent - *MDR-1*, DLKP Txt YL - *MDR-1* (Liang et al., 2004) and H460 Parent - *MRP-1* (O'Connor et al., 2004)). There were cases where both Western blot/ELISA and immunohistochemistry suggested the protein was not

expressed, which contradicted reports in literature (MDA-MB-468 - *EGFR* (Konecny et al., 2006) and DLKP VCR- *MDR-1/P-gp* (Liang et al., 2004)).

One proposed reason for the identification of staining when using Western blot/ELISA that was not observed with immunohistochemical analysis was the threshold created to identify positive membrane staining was set too high. The presence of positive staining was identified by a single reviewer's assessment of the immunohistochemically stained cell lines. Human analysis, as previously described, is highly flawed when quantifying staining intensity, and in this study may have failed to identify expression which was identified by Western blots or ELISA. Direct comparison between quantitative ELISA and image analysis data within this study was not possible, due to the narrow range of protein expression observed within the 23 cases assessed by ELISA. In future cell line assessments, it may be beneficial to have a validation set of samples with a complete set of quantitative data available in order to generate optimal cut-points. In addition, using the same antibody for ELISA and immunohistochemistry may prove beneficial, as differing affinities of antibodies for an antigen may contribute greatly to such discrepancies. However, the application of a single antibody for immunohistochemistry, Western blots and ELISAs analysis may prove difficult to achieve.

Digital assessment of immunohistochemically stained cell lines, rather than tissue, adds an additional layer of complexity for human analysis. Consistent and high image quality is a prerequisite for reliable and accurate scoring of digital images. To ensure high-throughput and high quality images, all immunohistochemically stained Cell Microarrays were scanned using an automated slide scanning system. Certain cell types within the cell lines are compact cells with little cytoplasm, which are features complicating the process of obtaining well-focused images (Stromberg et al., 2007). During the digitisation of all spots in the Cell Microarrays, it is likely that a subset of images with suboptimal focus may have been included and, therefore, possibly adversely affecting the analysis. In addition, the density of the cells was extremely high. Cell Microarrays were constructed from cell pellets created by packing cell line pellets into agarose molds (as described in Chapter 2, section (2.1.4.1)). This resulted in a very dense cell pellet, which may have been too high

density of cells for immunohistochemical analysis. Control cell lines present within the Cell Microarrays were constructed within the Tissue Array Research Program (USA), using cell suspension techniques. This technique appeared to be better suited to Cell Microarray creation as cells were evenly distributed within the cores.

High-throughput image analysis was performed in order to quantify protein expression across all cell lines which were immunohistochemically stained with 6 antibodies. For the majority of these, protein expression was previously unreported in literature. Utilising image analysis, protein expression was quantified on a continuous scale, which was subsequently divided into a semi-quantitative scale (Weak, Medium and Strong); thus, allowing comparison with previous and subsequently published expression levels, the majority of which are classified semi-quantitatively. Utilising the harness provided by Distiller, image analysis was extremely high-throughput. Analysis required approximately 5-8 minutes per Cell Microarray. Construction, immunostaining and image analysis on low density Cell Microarrays, such as the one described in this study, greatly increases the speed, standardisation and uniformity of analysis compared to Western blots/ELISA.

Cell culture plays a critical role in tumour characterisation and the identification of novel therapeutic targets. Once the Cell Microarray block is created, the cell lines are formalin-fixed paraffin-embedded and are, therefore, preserved for future analysis. The Cell Microarray generated in this study provides a repository of cells available for subsequent protein expression profiling, optimisation of newly-developed antibodies, and identification of positive controls. To the best of our knowledge this is the largest repository of immunohistochemically-stained cell lines which have undergone protein expression profiling via automated image analysis.

4.5 Conclusion

Cell Microarrays composed of 31 lung and breast cancer cell lines were immunohistochemically stained with 6 antibodies. Protein expression was qualified by human analysis, Western blots, ELISA and image analysis. The level of agreement between human analyses of immunohistochemically stained cell lines and alternative assays (Western blots and ELISA) was assessed.

Overall, results illustrated a high level of agreement between immunohistochemistry and Western blots/ELISAs testing when assessing protein expression. Agreement was hugely influenced by the antigens under assessment. When disagreement between methods was observed, it was largely when immunohistochemistry failed to identify protein expression which was present when assessed by Western blots. Results illustrate the importance of creating optimal cuts-points when validating image analysis algorithms.

In order to optimally validate the image analysis algorithm, a complete set of quantitative data would ideally be available. In addition, the application of a single antibody across Western blot/ELISA and immunohistochemistry if possible may introduce a high standard of uniformity into the analysis of protein expression.

A Cell Microarray has been created which provides a repository of cells available for subsequent protein expression profiling. Utilising automated image analysis, a high-throughput method for protein profiling of immunohistochemically-stained cell lines in a Tissue Microarray format has been achieved.

CHAPTER 5: APPLICATION OF AUTOMATED IMAGE ANALYSIS IN THE QUANTIFICATION OF IMMUNOHISTOCHEMICALLY STAINED BLADDER CANCER BIOPSIES

5.1 Introduction

Immunohistochemistry is a well-established and versatile technique, which is routinely used in molecular and surgical pathology (Kononen et al., 1998, Cregger et al., 2006). Immunohistochemistry allows for the identification and localisation of cell-bound antigens and can be performed on numerous cells and tissue preparations (Fejzo and Slamon, 2001). Immunohistochemistry is widely used due to its relatively low cost, availability of materials in routine pathology laboratories and relatively rapid turn around. However, at best, immunohistochemistry is a semi-quantitative technique.

Traditionally, human analysis has been considered the optimal method for qualifying immunohistochemical staining. However, the ability to quantify staining intensity by human analysis has produced varied results and is inherently flawed (Conway et al., 2006). In addition, results from previous chapters indicate that the quantification of immunohistochemical staining is greatly influenced by the complexity of the immunostain under assessment. Human analysis generally quantifies staining intensities into broad categories, rather than assigning exact staining intensity values. At present, alternative methodologies can accurately quantify protein signal when performed in conjunction with computer-assisted analysis. However, in some instances immunohistochemistry remains the primary technique utilised (Bartlett et al., 2003, Ellis et al., 2004, Hsi and Tubbs, 2004, Ellis et al., 2000, Hicks and Tubbs, 2005, Kay et al., 2004).

It has been proposed that computer-based analysis can quantify staining intensities more accurately and with greater reproducibility, than human-based analysis (Weaver et al., 2003, Johansson et al., 2001). There are numerous commercially available, computer-based systems designed for the quantification of immunohistochemical staining, including BLISS and IHCscore of Bacus laboratories, Inc (Lombard, USA); Ariol SL-50 Applied Imaging Corporation (San Jose, USA); AQUA, HistoRx Inc (New Haven, USA) (Cregger et al., 2006). However, automated quantification systems can be expensive, time-consuming and extremely complex.

Tissue Microarrays provide high-throughput analysis of immunohistochemically stained tissue (Kononen et al., 1998). Tissue Microarrays have numerous benefits over whole section analysis including uniform experimental conditions, conservation of scarce tissue and a reduction in the volume of reagents used (Simon and Sauter, 2002, Al Kuraya et al., 2004, Milanes-Yearsley et al., 2002, Hoos and Cordon-Cardo, 2001). Image analysis of Tissue Microarrays provides a high-throughput, reproducible and quantitative means of analysing immunohistochemically stained tissue.

Urothelial cell carcinoma of the Bladder (UCB) is the 7th most common cancer worldwide in men (10.1 new cases per 100,000 person per year) and the 17th in women (2.5 new cases per 100,000 person per year) (Murta-Nascimento et al., 2007). There is quite a difference between incidence and mortality rates for patients with UCB; this would suggest that UCB has a long progression period. In the USA, the 5 year relative survival rate ranges from 97% for those diagnosed with stage 1 UCB, to 22% for those with stage 4 UCB, according to the Tumour Node and Metastasis (TNM) classification (Gloeckler Ries et al., 2003, Murta-Nascimento et al., 2007). The tumour stage and grade are already recognised as significant indicators of prognosis; more reliable and more useful indicators are required for the prognosis of patients with invasive bladder cancer (Kashibuchi et al., 2007).

As described in Chapter 4; E-cadherin is a transmembrane glycoprotein involved in intercellular adhesion (Mahnken et al., 2005, Bukholm et al., 1998) which mediates the selective adhesion of epithelial cells for the interaction and maintenance of normal integrity (Takeichi, 1991). The correlation between E-cadherin expression and survival data continues to be debated (Campos et al., 2006, Kase et al., 2000, Zhou et al., 2002, Gamboa-Dominguez et al., 2005, Nakopoulou et al., 2000, Tamura et al., 2005, Chow et al., 2001, Brecelj et al., 2005, Bremnes et al., 2002a, Bringuier et al., 1993, Byrne et al., 2001, Shariat et al., 2001). However, as evident from the previous chapter, high levels of inter- and intra-observer variability are present when E-cadherin staining is assessed by human analysis. Therefore, E-cadherin assessment performed by image analysis maybe of benefit.

E-cadherin is known to play a role in cell-cell adhesion through an intercellular Ca^{++} dependent hemophilic junction. E-cadherin associated molecules, including α-, ß- and γ-catenin, link the E-cadherin to the actin cytoskeleton. The dysfunction of cell-cell adhesion molecules is related to the initial step of the invasion of the surrounding tissues and metastasis to distant organs. It has been recognised that the reduced expression of E-cadherin is closely related to the cell differentiation status and invasiveness of a tumour (Kashibuchi et al., 2007, Frixen et al., 1991, Vleminckx et al., 1991). The association of the catenins with E-cadherin is an essential requirement for the correct function and maintenance of cell adhesion within a tissue (Kashibuchi et al., 2007). α–catenin mediates the interaction of the cadherin and ß-catenin complexes with the cytoskeleton (Nathke et al., 1993), while ß-catenin directly binds to the cytoplasmic domain of E-cadherin and regulates cadherin-mediated cell recognition and adhesion (Peifer, 1993). Evaluation of effect of loss of E-cadherin and ß-catenin expression on prognosis in bladder cancer has previously been investigated (Moyano Calvo et al., 2006, Clairotte et al., 2006, Romanenko et al., 2006, Nakopoulou et al., 2000). Both E-cadherin and ß-catenin loss of expression correlates with poor prognostic outcome; however, in a multivariate setting, poorer prognostic outcome is associated with E-cadherin expression only (Nakopoulou et al., 2000). Also, tumours which lack E-cadherin or ß-catenin have lower recurrence-free time (Moyano Calvo et al., 2006). To date, no fully automated quantitative analysis of urothelial cell carcinoma of the bladder stained for E-cadherin and ß-catenin has been correlated with survival.

The validation of an image analysis algorithm which quantifies membrane-bound immunohistochemical staining intensity is described in detail in this chapter. The validation of the algorithm involved performing quantification of membrane-bound E-cadherin protein expression across 26 Tissue Microarrays composed of urothelial cell carcinoma of the bladder. In order to access the optimal method of review and to verify the prognostic value of E-cadherin expression, human and image analysis were correlated with survival data. In addition, the correlation between E-cadherin and ß-catenin expression with prognostic outcome has been assessed across 25 of theses 26 Tissue Microarrays. As previously mentioned, the VTM has been replaced

with a commercially available software application know as Distiller, which was derived from the VTM prototype.

5.2   Study Design

5.2.1   Patient Cohort

A total of 26 Tissue Microarrays were constructed, comprising of 418, 2.0 mm cores from 145 consecutive cases of urothelial cell carcinomas of the bladder. The Tissue Microarrays were constructed and immunohistochemically stained for E-cadherin (Novocastra/Vision Biosystems, USA) and ß-catenin (Labvision Corp., UK) as described in Methods Chapter 2 section 2.3.1.3. A single slide was not available for the assessment of ß-catenin; therefore, the sample size was decreased to 141 for ß-catenin assessment. A liver control spot derived from one biopsy was present on all slides. The liver control spot stained positive for E-cadherin and ß-catenin, and was used as a positive control to verify uniform staining across all Tissue Microarrays. Tissue Microarrays were digitised using an Aperio ScanScope T3 Scanner™ (Aperio Technologies, USA) at 20 x magnification, as described in Chapter 2 (section 2.4.2).

5.2.2   Tumour Heterogeneity

Multiple cores were sampled from different tumour areas to account for heterogeneity within any one tumour. The number of replicate cores from a single biopsy ranged from 1-5. Table 5.2-1 illustrates the number and distribution of replica cores taken from a single biopsy. The majority of biopsies were sampled 4 times (n = 72 biopsies). All cores could not be used for image analysis assessment of E-cadherin and ß-catenin staining intensity. Cores were excluded from image analysis based on amount of core present. In order to eliminate mis-representation of the staining intensity present, all cases with less than 45% of the core present were excluded from this study. Cases available for image analysis assessment of E-cadherin are within *(brackets)* in Table 5.2-1. An equal number of biopsies were assessed by human and image analysis. However, the number of cores assessed by human and image analysis differed slightly (418 and 386 cores respectively).

Table 5.2-1: Distribution of replica cores taken from a single biopsy. Cases within the *(brackets)* are the number of cases available for image analysis assessment of E-cadherin staining intensity. Cases without brackets are the number of cases assessed by human analysis.

Replication	Number of biopsies	Number of Cores
1x	44 (48)	44 (48)
2 x	9 (16)	18 (32)
3 x	16 (21)	48 (63)
4 x	72 (57)	288 (228)
5 x	4 (3)	20 (15)
Total	145	418 (386)

5.2.3   Normalisation of Image Analysis Results

A liver control spot was present on all Tissue Microarrays. The control spot was used to verify uniform staining across all Tissue Microarrays. When assessed by human analysis, the level of E-cadherin staining present on the control spots was deemed constant across all Tissue Microarrays.   Image analysis was performed across all control spots in order to quantitatively assess staining intensity variations.

Figure 5.2-1 illustrates the modal intensity of E-cadherin staining present on control tissue present across 26 Tissue Microarrays, when assessed by image analysis. It was expected that staining intensity would have been constant across all control spots, as immunohistochemical staining was performed by an automated staining. The majority of slides (n = 18) obtained identical staining intensity values. However, there was a range of staining intensity observed for the remainder of slides (n = 8).

Image analysis was performed across all control spots immunohistochemically stained with ß-catenin. The level of staining variation was comparable to that observed across Tissue Microarrays immunohistochemically stained for E-cadherin. As a result, it was deemed necessary to perform normalisation of data obtained from the quantification of both E-cadherin and ß-catenin staining intensity. Image analysis data was normalised as described in Chapter 2 section (2.6.1).
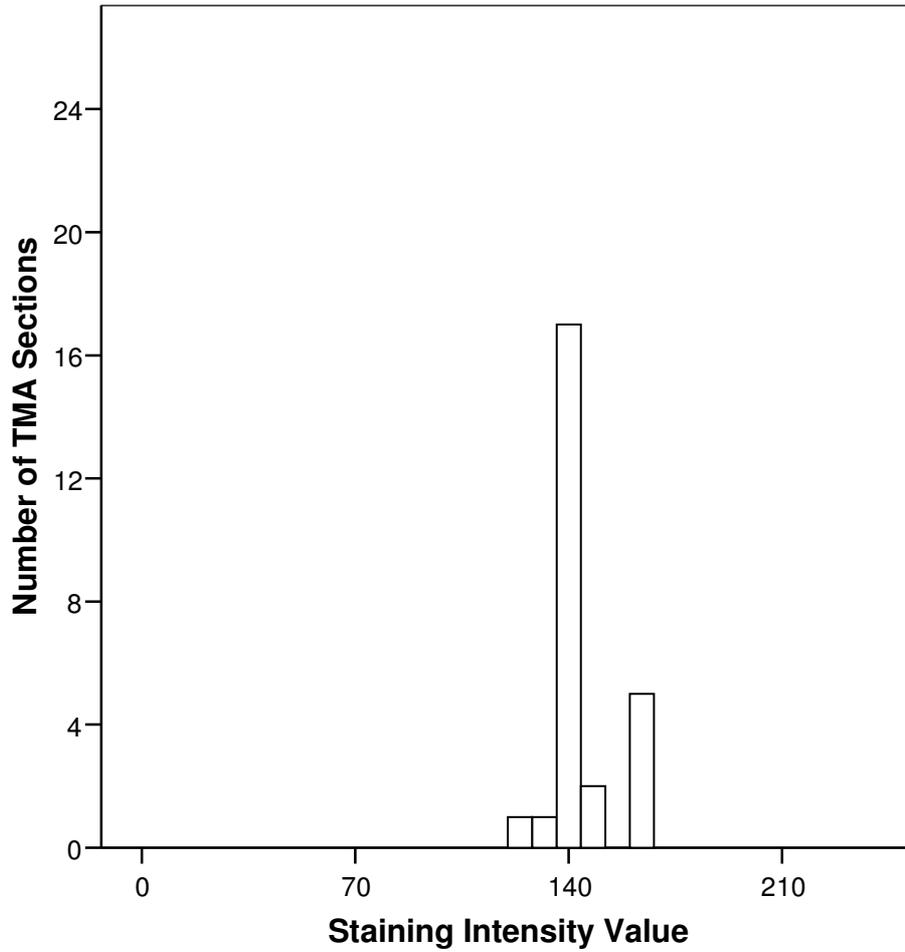
Figure 5.2-1: The level of E-cadherin staining intensity present on control tissue across 26 Tissue Microarray sections when assessed by image analysis. As a result of a certain degree of staining intensity variance image analysis results were normalised.

5.2.4   Categorisation of Staining Intensity

5.2.4.1 Human Analysis Categories

Originally, assessment of E-cadherin membrane staining intensity was performed by a single research scientist, using traditional microscope-based assessments. A three-point scale was used to quantify E-cadherin membrane staining intensity. However, on closer examination of these results the review was deemed unusable, due to the fact that 77% of all Tissue Microarray spots were categorised as moderately stained when assessed by this reviewer. The over reliance of the moderate category skewed the data and therefore, using this data to correlate E-cadherin staining intensity with prognosis would have negatively affected the results.

Therefore, one Professor of Pathology, highly experienced in the review of immunohistochemistry, scored the Tissue Microarrays, using virtual slide review methods for the assessment of E-cadherin membrane staining intensity. In order to validate standards, a research scientist performed a 10% review in conjunction with the histopathologist. A two and three-point scale was used to quantify E-cadherin membrane staining intensity, when performed by human analysis. The two and three-point scales are classified as follows: (1) weak staining and (2) strong staining / (1) weak staining, (2) moderate staining and (3) strong staining.

5.2.4.2 Image Analysis Categories

The image analysis scale ranges from 0.55 to 1.71 for E-cadherin membrane staining intensity and 0.27 to 1.55 for ß-catenin staining intensity. Utilising X-tile, cut-points were identified for E-cadherin at $\geq$ 0.8 (Strong Staining) and < 0.8 (Weak Staining). Cut points for ß-catenin were identified at $\geq$ 0.71 (Strong Staining) and < 0.71 (Weak Staining).

5.3   Results

Membrane-bound immunohistochemical staining intensity was classified using automated image analysis as follows: Twenty six Tissue Microarrays were stained for E-cadherin; membrane staining intensity was assessed by virtual-based analysis and the image analysis algorithm. The level of correlation between human and automated analysis were assessed and both methods were compared with prognostic outcome. In addition, the relevance of utilising E-cadherin expression in conjunction with ß-catenin expression as a prognostic indicator in UCB was assessed.

5.3.1   E-cadherin expression correlated with prognostic outcome

5.3.1.1 Staining Intensity Assessed by Human Analysis

As previously mentioned, E-cadherin protein expression was originally assessed by one research scientist. There were 16 more cases examined during the research scientist evaluation, than subsequent assessments. However, when staining intensity was assessed using three categories by the research scientist there was no significant difference in survival rates over 5 years (p = 0.817, 95% CI, 2 df, 1.00/1.31/1.04 RR)(Figure 5.3-1). The poor correlation between membrane E-cadherin staining and prognosis was attributed to the over use of the moderate staining intensity category. Therefore, it was deemed necessary for a more experienced individual to reassess immunohistochemical staining intensity.

Kaplan-Meier curve for Human Analysis: Research Scientist
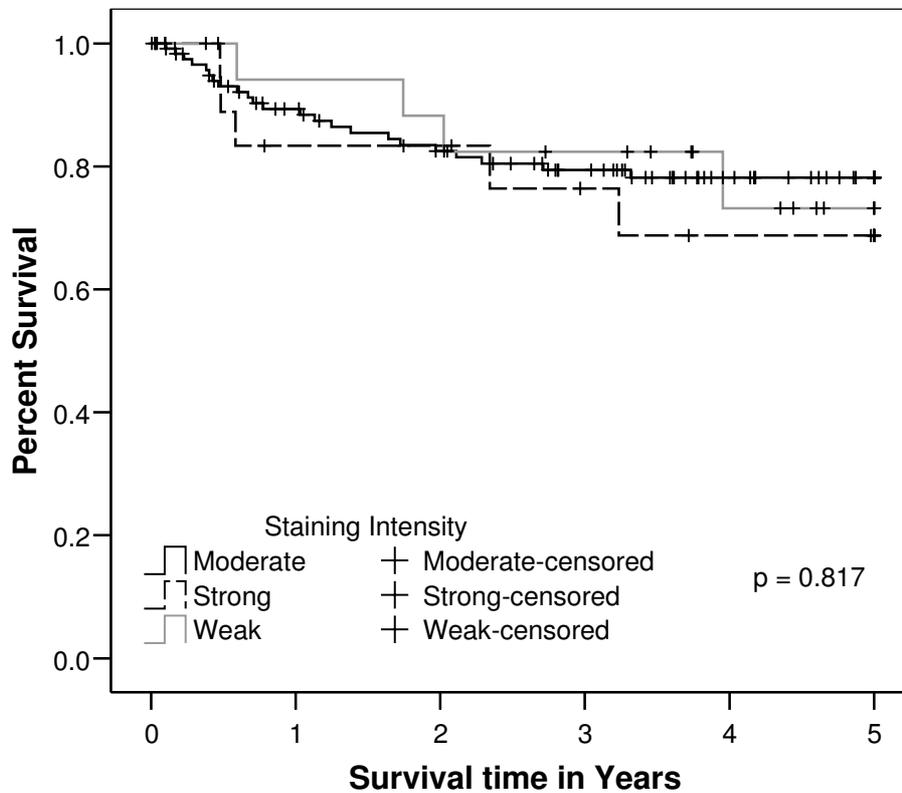
Figure 5.3-1: Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 26 UCB Tissue Microarrays, where E-cadherin staining intensity was reviewed by a research scientist utilising three categories: Weak (*n=17*), Moderate (*n=124*) and Strong (*n=20*).

The Kaplan-Meier curves illustrate the correlation between intensity of membrane E-cadherin staining with prognostic outcome over 5 years, where intensity was reviewed by human analysis (performed by a Professor of Pathology) (Figure 5.3-2). Staining intensity was assessed utilising two and three categories. When staining intensity was classified by human analysis utilising two categories, there was a significant difference in survival rates (p = 0.027, 95% CI, 1 df, 2.35/1.00 RR). In addition, when staining intensity was classified by human analysis utilising three categories, there was also a significant difference in survival rates (p = 0.010, 95% CI, 2 df, 3.44/1.59/1.00 RR).

Figure 5.3-2: Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 26 UCB Tissue Microarrays, where E-cadherin staining intensity was reviewed by human analysis utilising two and three categories. Two categories: Weak (*n=16*), and Strong (*n=129*). Three categories: Weak (*n=9*), Moderate (*n=43*) and Strong (*n=93*).

5.3.1.2 Staining Intensity Assessed by Image Analysis

The Kaplan-Meier curve illustrates the correlation between intensity of membrane E-cadherin staining with prognostic outcome over 5 years, where intensity was reviewed by image analysis (Figure 5.3-3). Within this cohort of patients, there was a significant difference in survival rates for the two classes identified by image analysis, over 5 years (p = 0.013, 95% CI, 1 df, 2.06/1.00 RR). Utilising image analysis and X-tile, a class of patients which have a strong correlation with improved survival has been identified, when staining intensity is categorised by image analysis into two groupings.

Kaplan-Meier curve for Image Analysis
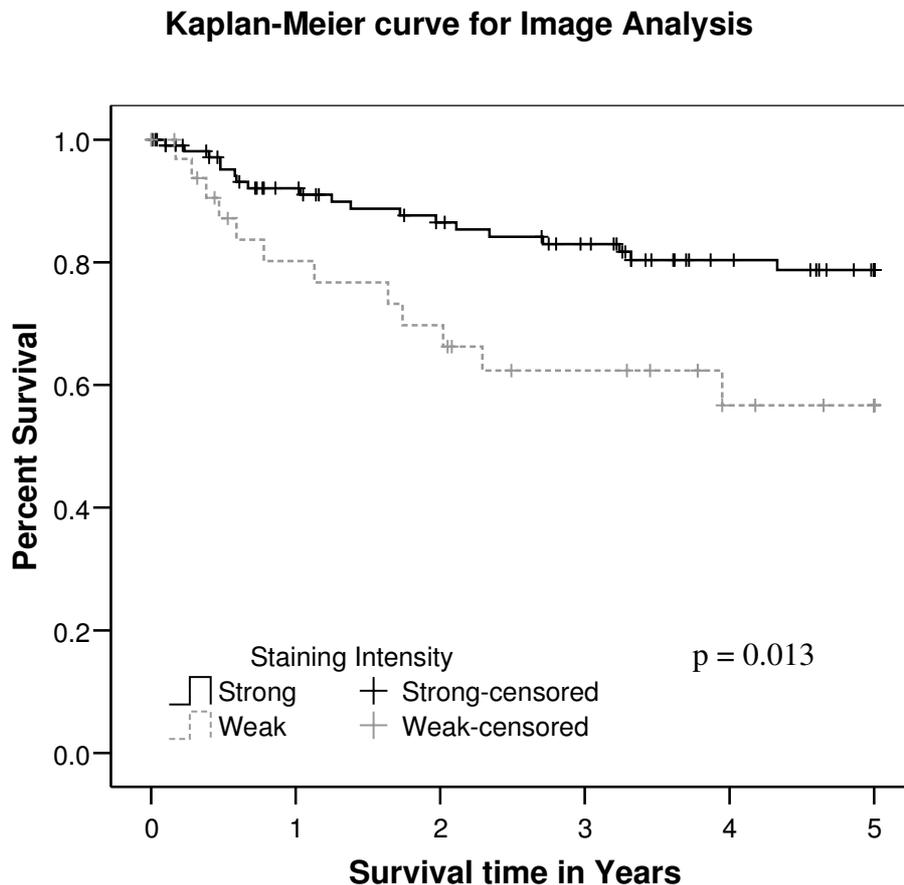
Figure 5.3-3: Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 26 UCB Tissue Microarrays, where E-cadherin staining intensity was reviewed by image analysis utilising two categories, Weak (*n=34*) and Strong (*n=111*).

5.3.2 Distribution of Staining Intensity Values: Human vs. Image Analysis

Figure 5.3-4 illustrates results from image analysis of Tissue Microarrays immunohistochemically stained with E-cadherin, sorted into human analysis intensity categories. Human analysis classified staining intensity utilising both two and three categories. The median of the staining intensity results illustrate the correlation between increasing staining intensities observed in both image and human analysis. The range of staining intensities categorised as strongly staining by human analysis, completely encompasses the alternative staining categories when intensity is classified by two or three categories. The majority of cores were classified as strongly stained when assessed by human analysis using two or three staining categories (89.0% and 64.1% respectively). There were a number of outliers present outside the normal distribution within the box plots, the majority of which were present when intensity was classified as moderately stained.

ANOVA illustrated in general there was a highly significant difference between the image analysis staining intensity values, when sorted by three human intensity categories ($p = 0.000$, 95% CI, 144 df). There was no significant difference between image analysis staining intensity values when comparing the human intensity categories weak and moderate staining ($p = 0.327$, 95% CI, 144 df). However, when comparing weak with strong staining and moderate with strong staining, there was a significant difference ($p = 0.000$ and $p = 0.000$ 95% CI, 144 df, respectively). T-tests illustrated there was a significant difference between the image analysis staining intensity values, when sorted by two human intensity categories ($p = 0.000$, 95% CI, 144 df). However, in both cases the strong category expands the entire image analysis range.
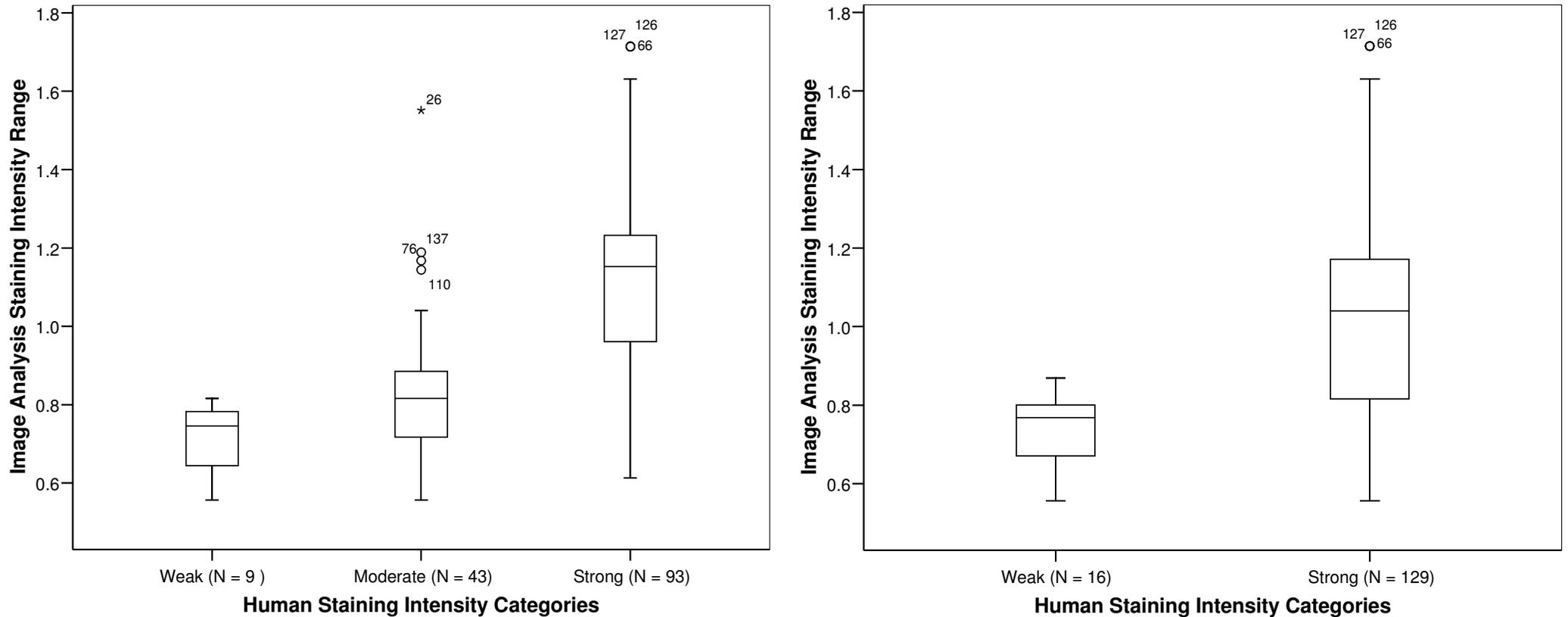
Figure 5.3-4: Box-plot representing image analysis results when sorted by human analysis categories for the quantification of E-cadherin expression across 26 UCB Tissue Microarrays. Human analysis was assessed using virtual slides by a Histopathologist utilising two (Weak or Strong) and three (Weak, Moderate or Strong) categories. Within the image analysis scale, 0.6 is the lightest staining possible and 1.8 is the strongest. The symbols (○ and *) signify the outliers present within he dataset. The numbers within the box-plots represents the patients identification number.

5.3.3   Cancer Stages and Grades

Tumour grade is the system used to classify cancer cells in terms of how abnormal they appear. Tumour grade gives an indication of how fast the tumour is likely to grow and spread. Grade is usually described as low (well differentiated), medium (intermediate differentiated) or high (poorly differentiated/undifferentiated). Differentiated cells resemble normal cells and tend to grow and spread at a lower rate than undifferentiated cells, which lack the structure and function of normal cells and grow uncontrollably. Therefore, patients with low grade cancers are expected to have a better prognosis than those with high grades. To evaluate the effect of tumour grade on prognosis within this study, cells were classified as (1/2) well/intermediate differentiated or (3) poorly/undifferentiated cells.

Tumour stage refers to the extent or severity of the cancer, based on factors such as the location of the primary tumour, number and size of tumours, lymph node involvement, and presence or absence of metastasis. Tumour stages are mostly commonly categorised by TNM system. TNM refers to primary tumour (T), regional lymph nodes (N), and distant metastasis (M). Within this cohort of patients, the primary tumour stages (T) were known, and the following stages of UCB were observed, pTa, pT1, pT2, pT3, and pT4. pTa and pT1 denoted superficial bladder cancer, or early stage bladder cancer. The cancer is in early stage of development and is situated in the innermost layer of the bladder, in the lining. pT1 is slightly more developed than pTa. pT1 tumours have started to grow further into the bladder wall than the lining. pT2 and pT3 bladder cancer denote invasive bladder cancer. pT2, the cancer has grown through the connective tissue into the muscle. pT3, the cancer has grown through the muscle layer. pT4 is known as advanced bladder cancer, the cancer in this case has spread outside of the bladder. To evaluate the effect of tumour stage on prognosis within this study, cells were classified as (pTa-pT1) superficial or (pt2/3/4) invasive. Tumour stage and grade information was only available for 141 (of the 145) patients within this dataset. Figure 5.3-5 illustrates the primary tumour stages of bladder cancer.
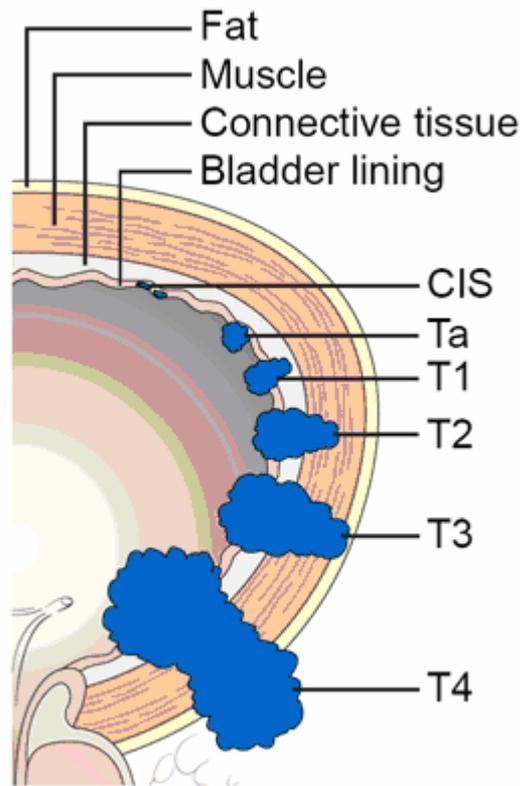
Figure 5.3-5: Primary tumour stages of bladder cancer. CIS-carcinoma *in situ*, Ta, and T1 denote superficial bladder cancer. T2 and T3 are invasive bladder cancer. T4 is known as advanced bladder cancer. Image was sourced from CancerHelp UK.

5.3.4 Stage and Grade correlated with Prognosis

The Kaplan-Meier curves illustrate the correlation between tumour stage and tumour grade of UCB with prognostic outcome over 5 years (Figure 5.3-6). Within this cohort of patients, there was a significant difference in survival rates between the two classes used for both tumour staging and grading of cancer (Figure 5.3-6). There was a significant difference in survival rates between patients with tumour grades that were well / intermediate differentiated (1/2), and poorly / undifferentiated (3), (p = 0.000, 95% CI, 1 df, 1.00/3.18 RR). There was also a highly significant difference in survival rates between patients with tumour stages that were superficial (pTa-pT1) and invasive (pT2/3/4) (p = 0.000, 95% CI, 1 df, 1.00/4.80 RR).
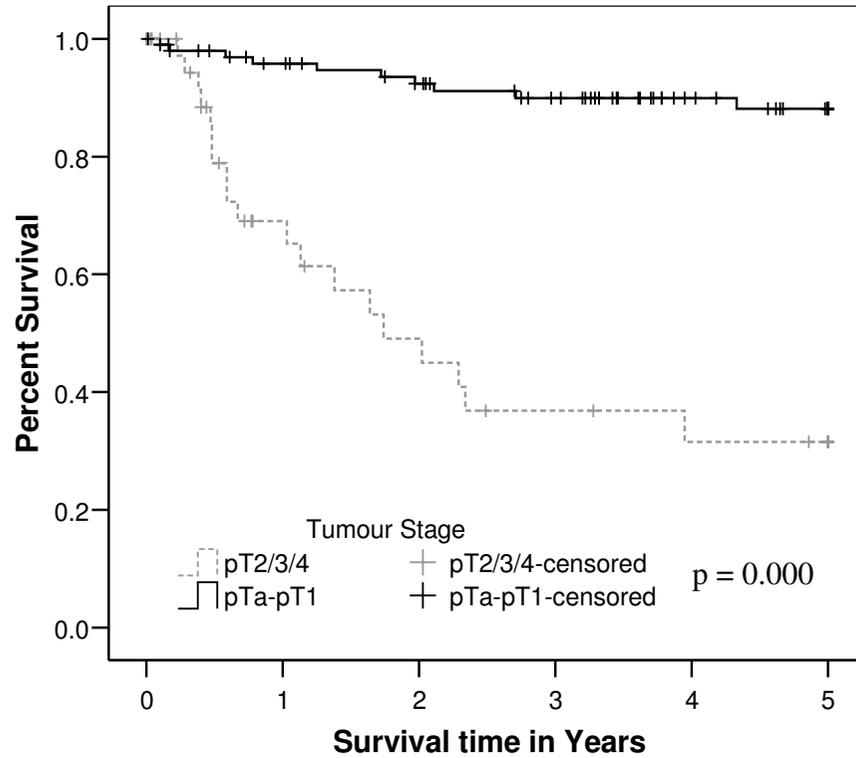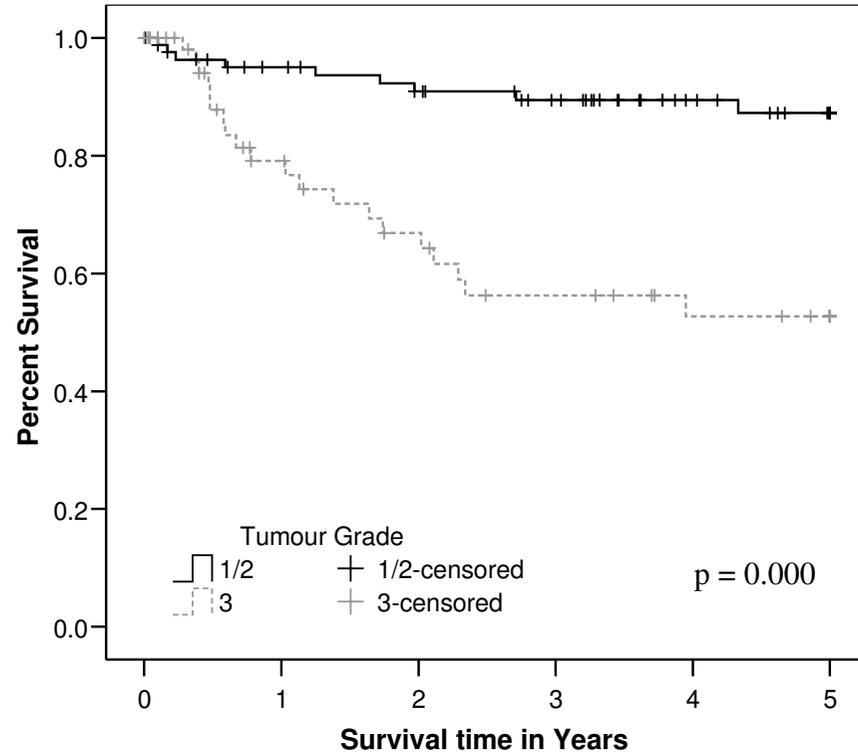
**Figure 5.3-6:** Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 26 UCB Tissue Microarrays, when outcome is categorised by stage and grade of tumours. Tumour Stage is categorised as **pTa-pT1** (Superficial, *n=101*), and **pT2/3/4** (Invasive, *n=40*). Tumour Grade is categorised as **1/2** (well/intermediate differentiated, *n=83*), and **3** (poorly/undifferentiated, *n=58*).

5.3.5 Distribution of Tumour Stage, Grade and E-cadherin Expression

The distribution of cases when categorised by E-cadherin expression, tumour stage and grade is illustrated in Table 5.3-1 and Table 5.3-2, respectively. E-cadherin intensity was reviewed by image analysis. Overall there was a good correlation between E-cadherin expression and tumour stage and grade, (Pearson Chi Squared = 5.46, and 10.28, respectively). The majority of correlation between E-cadherin expression and tumour stage was observed when E-cadherin intensity was strong and cancers were superficial (pTa-pT1).

Table 5.3-1: Stage of UCB cancer compared with E-cadherin protein expression when quantified by image analysis.

		Image Analysis		
		Strong	**Weak**	**Total**
Stage	**pTa-pT1**	82	19	101
	pT2/3/4	25	15	40
Total		107	34	141

The majority of correlation between E-cadherin expression and tumour grade was observed when E-cadherin intensity was strong and cancers were well/intermediate differentiated (1/2, Table 5.3-2).

Table 5.3-2: Grade of UCB cancer compared with E-cadherin protein expression when quantified by image analysis.

		Image Analysis		
		Strong	**Weak**	**Total**
Grade	**1/2**	71	12	83
	3	36	22	58
Total		107	34	141

163

5.3.6 Tumour Stage (pT2/3/4) categorised by E-cadherin expression levels

As illustrated in Figure 5.3-6, patients with invasive tumours (pT2/3/4) have a considerably poorer prognosis than those with superficial tumours (pTa-pT1). The effect of E-cadherin staining intensity on invasive tumours was examined, in order to assess if a distinct sub-population with a poorer prognosis was in existence. The Kaplan-Meier curves illustrate the correlation between weak and strong E-cadherin staining of invasive UCB tumours (pT2/3/4) with prognostic outcome over 5 years, where staining intensity was assessed by image (Figure 5.3-7, A) and human analysis (Figure 5.3-7, B). Within this cohort of patients, there was not a significant difference in survival rates between the two classes used for quantifying E-cadherin staining intensity. However, the difference in survival rates between patients with weak and strong expressing tumours where at their greatest when E-cadherin staining was assessed by image analysis ($p = 0.087$) compared to human analysis ($p = 0.290$).
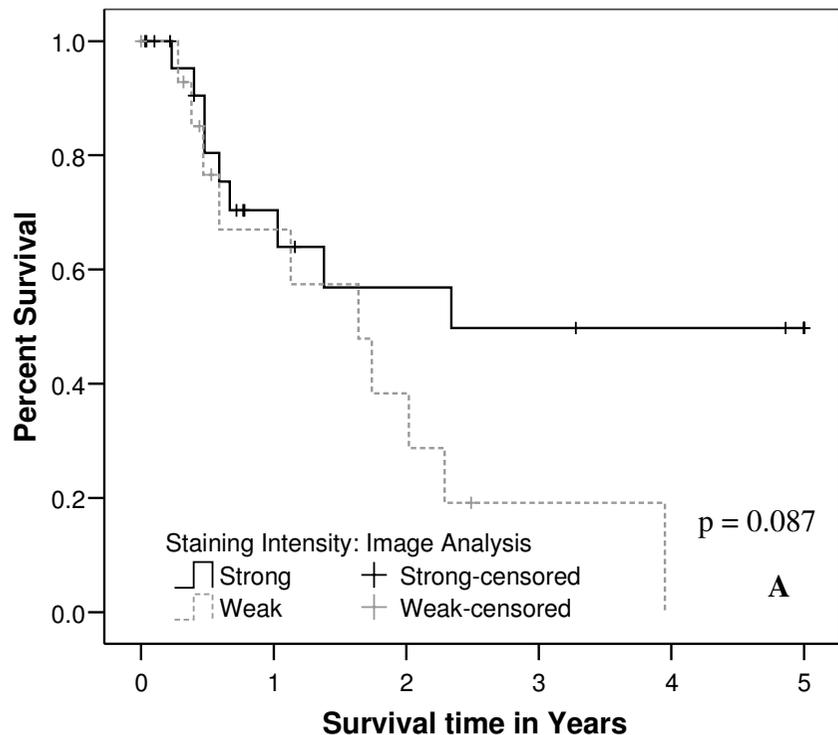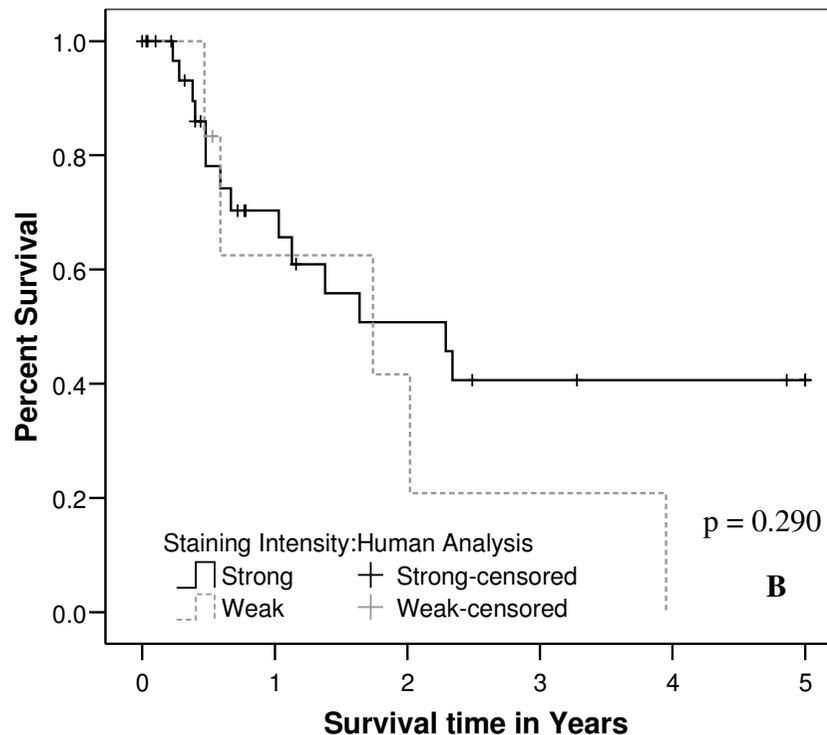
Figure 5.3-7: Kaplan-Meier curve evaluating rate of death due to UCB over 5 years across 26 Tissue Microarrays, where E-cadherin staining intensity was reviewed by image (A) and human (B) analysis. In all cases tumour stage was invasive (PT2/3/4). E-cadherin staining was categorised as weak (*n=15*) and strong (*n=25*) during image analysis (B), and weak (*n=6*) and strong (*n=34*) during human analysis (B).

5.3.7   ß-catenin Expression in Cancer

Cadherins found in all tissue, are mediators of cell-cell adhesion. A cadherin molecule on one cell binds to a cadherin molecule of the same type on another cell (Takeichi, 1991, Gontero et al., 2004). E-cadherin is expressed by all normal epithelia cells (Smith and Pignatelli, 1997). Cadherin binding occurs via members of the catenin family of molecules. Defects in catenin molecules may abolish cadherin binding at the cell surface. The three catenins that have been identified (α, ß and γ) link E-cadherin to the actin cytoskeleton and are important for maintaining its role in cell-cell adhesion (Behrens et al., 1989, Gontero et al., 2004). A reduction or loss of adhesion in human malignancy allows cells to detach from the primary site and thereby initiating the first steps in the metastatic process (Shimazui et al., 1996, Nakopoulou et al., 2000, van Oort et al., 2007, Frixen et al., 1991, Gontero et al., 2004). Therefore, a loss of either E-cadherin or a member of catenins would negatively affect prognosis.

Several studies have reported the associations between E-cadherin and ß-catenin expression levels with prognosis in bladder cancer (Nakopoulou et al., 2000, Shimazui et al., 1996, Clairotte et al., 2006, Kashibuchi et al., 2007). However, reported prognostic values of these antigens have varied. In a univariate setting, both E-cadherin and ß-catenin expression have been found to have a significant effect on prognosis (Nakopoulou et al., 2000, Shimazui et al., 1996, Clairotte et al., 2006). However, in a multivariate setting, only E-cadherin expression was a significant indicator of prognosis (Nakopoulou et al., 2000). Literature has reported that E-cadherin and ß-catenin have similar prognostic values and that staining for either of these antigens is acceptable (Shimazui et al., 1996).

The application of combining E-cadherin and ß-catenin expression levels as a prognostic indicator in bladder cancer, when intensity is assessed by image analysis, has not been reported in literature. Therefore, we extended our previous observations on E-cadherin, by combining E-cadherin and ß-catenin expression levels and correlating results with prognostic outcome. The objective is to clarify whether or not the expression levels of ß-

166

catenin could provide additional prognostic value over and above E-cadherin alone in UCB.

5.3.8 ß-catenin expression correlated with prognostic outcome

The Kaplan-Meier curve illustrates the correlation between intensity of membrane ß-catenin staining with prognostic outcome over 5 years, where intensity was assessed by image analysis (Figure 5.3-8). Within this cohort of patients, there was a significant difference in survival rates for the two classes identified by image analysis, over 5 years ($p = 0.044$, 95% CI, 1 df, 2.20/1.00 RR) As with E-cadherin expression, the optimum cut-points were generated using X-tile.

Kaplan-Meier curve for Image Analysis
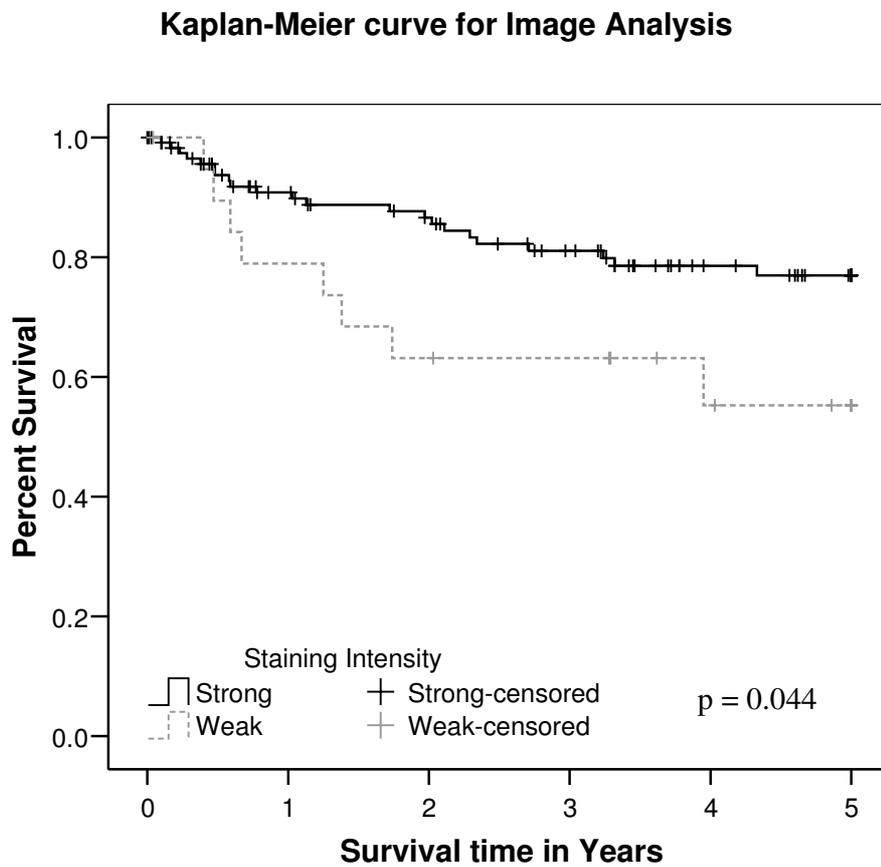
Figure 5.3-8: Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 25 Tissue Microarrays, where ß-catenin staining intensity was reviewed by image analysis utilising two categories, Weak (*n=20*) and Strong (*n=121*).

5.3.9 E-cadherin and ß-catenin expression correlated with prognostic outcome

Table 5.3-3 illustrates the distribution of E-cadherin and ß-catenin membrane staining intensity when assessed by image analysis, utilising two categories to classify staining intensity (Weak and Strong). There was a high level of correlation between E-cadherin and ß-catenin membrane expression (Kappa = 0.60), particularly when strong staining was observed.

Table 5.3-3: Correlation between E-cadherin and ß-catenin expression levels, when assessed by image analysis.

		E-cadherin Staining		
		Weak	Strong	Total
ß-catenin Staining	Weak	6	14	20
	Strong	27	94	121
Total		33	108	141

The Kaplan-Meier curve illustrates the correlation between intensity of E-cadherin and ß-catenin staining with prognostic outcome over 5 years, where intensity was assessed by image analysis (Figure 5.3-9). Membrane staining intensity was classified using three categories. Cases where both E-cadherin and ß-catenin staining was strong, cases where both E-cadherin and ß-catenin staining was weak, and cases where either E-cadherin or ß-catenin staining was strong. Within this cohort of patients, there was a significant difference in survival rates for the three classes identified by image analysis, over 5 years (p = 0.009, 95% CI, 2 df, 4.18 RR (Both Weak) / 1.68 RR (Either Strong) / 1.00 RR (Both Strong)).

Kaplan-Meier curve for Image Analysis
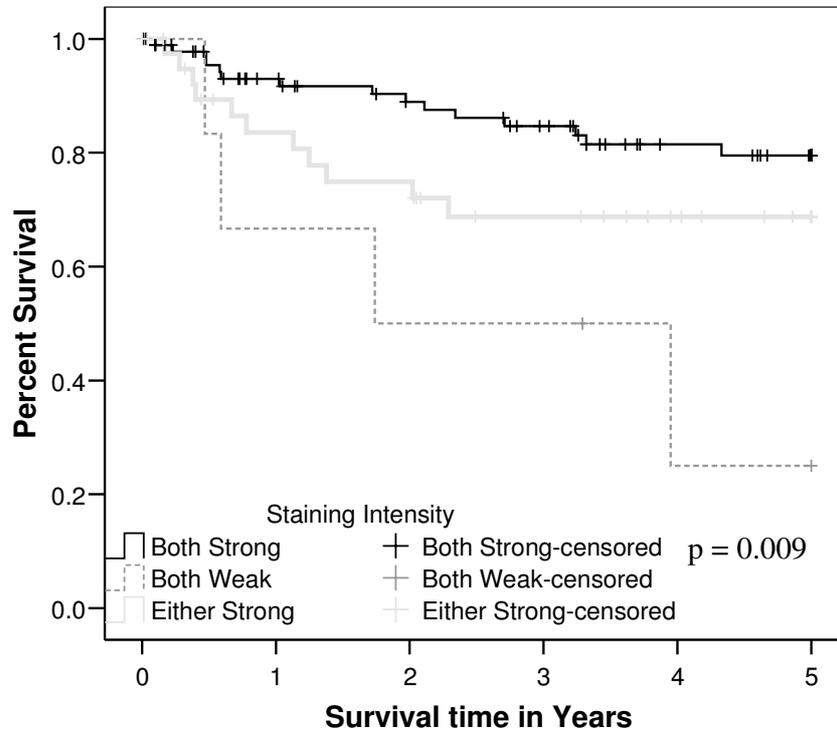
**Figure 5.3-9:** Kaplan-Meier curves evaluating rate of death due to UCB over 5 years across 25 Tissue Microarrays, where E-cadherin and ß-catenin staining intensity was reviewed by image analysis utilising three categories. Cases where E-cadherin and ß-catenin staining intensity were Both: Strong (*n=94*), Weak (*n=6*) or Either were Strong (*n=41*).

5.4　Discussion and Conclusions

The image analysis algorithm was validated using Tissue Microarrays stained for E-cadherin and ß-catenin. E-cadherin and ß-catenin have a characteristic brown tint, which is particularly difficult to isolate as it is dispersed across the colour spectrum and not ideal as such for colour thresholding. A colour cube was used to isolate colour, as it allows vastly different colours to be grouped together in the same domain. It is conceivable that the algorithm could be easily adapted to help quantify other biomarkers with predominantly membrane staining patterns. The algorithm is adapted by simply updating the colour cube. This process does not require any knowledge of computer programming and simply requires highlighting the range of colours specific to the biomarker of interest. This process is ideally suited to pathologists or clinical scientists and is therefore versatile and accessible to all those who would traditionally perform microscope-based analysis.

The concept of the algorithm is uncomplicated; areas that are considered membrane and brown are segregated from the rest of the image, and the intensity of grey in positive pixels is quantified. The creation of the algorithm in Image-Pro Plus® was straightforward as it facilitates recording and replaying of image processing sequences. However, not all laboratories have access to Image-Pro Plus®; therefore, the algorithm could be re-created with non-proprietary software, thus reducing the cost of the application and improving the algorithm's performance. For example, by utilising Java image library the speed of image analysis would be increased due to code compilation.

The algorithm alone was not enough to achieve high-throughput. A workflow was utilised to facilitate the automation of the algorithm. The workflow incorporated numerous stages including: digitisation and de-arraying of Tissue Microarrays, image analysis, storage of results, and finally statistical analysis was performed. In order to facilitate high-throughput, human intervention must be as limited as possible. In this study the first fully automated de-arraying process of its kind has been utilised. Thus fully automated image analysis of Tissue Microarrays was achieved.

When used in conjunction with Tissue Microarrays constructed solely from tissue of interest, image analysis has numerous advantages including speed, automation, accuracy and efficiency in quantifying staining intensity, providing an alternative to human analysis. Image analysis is completely objective and does not compare the image under analysis with those previously reviewed, which inevitably occurs with human analysis. Therefore, values are a true reflection of intensity, rather than a value of comparison. Once the macro has been created and recorded, analysis is fully automated. Images can be analyzed without human intervention, utilising a continuous scale. Therefore results are truly quantitative, in contrast to human analysis which at best is semi-quantitative. However, it should be noted that image analysis is only as good as the algorithm implemented.

Commercial applications are available to quantify immunohistochemical staining. The merits of utilising commercial applications are illustrated in literature (Ciampa et al., 2006, Camp et al., 2002, Divito et al., 2004, Camp et al., 2003, Berger et al., 2004, Charpin et al., 2004). However, most image analysis applications can quantify fluorescent or chromatic staining, but not both. As a result of this and high initial purchase costs, commercial applications are not always viable in research laboratories. The image analysis system described in this study utilises an algorithm, which is capable of quantifying chromatic rather than fluorescent staining. Therefore, the algorithm described provides a low cost alterative to many commercial image analysis applications.

At present, human analysis is considered the gold standard for immunostain quantification. During statistical analysis, image analysis results were grouped by human reviewing categories. The range of image analysis results within each of the human analysis categories cast doubts on the usefulness and reliability of human analysis to quantify immunostain intensity. Strong staining completely encompassed all other categories available when staining intensity was classified on a two and three grade scale, as illustrated in Figure 5.3-4. In addition, there were numerous outliers present

outside the normal distribution, when image analysis results were grouped by human reviewing categories.

The expertise of the reviewer greatly impacted the quality of results obtained within this study. Originally all tissue was assessed by a research scientist; however, no correlation between E-cadherin protein expression and prognosis was observed. However, when assessed by a Professor of Pathology correlation was observed. In addition, 77% of all cases were deemed moderate when assessed by the research scientist, whereas this value was greatly reduced during the pathologist review (43%).

Quantification of E-cadherin staining intensity performed by human and image analysis was compared with prognostic outcome, over 5 years. Within this cohort of patients, when E-cadherin staining intensity was classified into weak, moderate or strong by human analysis, there was a significant difference in survival times between the three classes over 5 years. In addition, when staining intensity was classified into two groups by human and image analysis, there was a significant difference in survival times between the two classes over 5 years. The continuous scale used by image analysis to quantify expression was divided into two categories, representing weak and strong expression. The continuous scale of intensity values was split using X-tile (Camp et al., 2004). This type of categorisation proved to be as effective as human categorisation where intensity is split into weak and high. In addition, image and human analysis proved to be almost as effective as tumour stage and grade for predicting prognosis in UCB. However, these findings only apply to this cohort of patients and require a test set for true validation of the correlation.

Both tumour stage and grade were significant indicators of prognosis in UCB. Using image analysis it was possible to identify a sub-population of invasive tumours with considerably poorer prognosis, when E-cadherin staining was weak. Although not significantly different, invasive tumours with strong E-cadherin staining when identified by image analysis had a better prognosis than strong cases identified by human analysis.

Quantification of ß-catenin staining intensity assessed by image analysis was compared with prognostic outcome, over 5 years. Within this cohort of patients, when staining intensity was classified into weak or strong by image analysis there was a significant difference in survival times between the two classes over 5 years. Correlations between the expression pattern of both E-cadherin and ß-catenin staining intensity with prognosis were also assessed. Staining intensity was categorised as follows, both proteins were: highly expressed; weakly expressed, or either proteins were strongly expressed. Results illustrate that combining E-cadherin and ß-catenin expression levels is marginally better as a prognostic indicator than either E-cadherin or ß-catenin expression alone. Utilising image analysis we found E-cadherin, ß-catenin and the combination of E-cadherin and ß-catenin expression levels to have similar prognostic values. Therefore, staining for either of these proteins could be used as a prognostic indicator in UCB. However, identifying cases with both weak E-cadherin and ß-catenin protein expression maybe of benefit when examining prognosis.

Results confirm, in this patient cohort, the importance of varying E-cadherin and ß-catenin protein expression for predicting survival in UCB. Traditionally, human analysis has been considered the gold standard for quantifying immunostaining; however, this may not be the case. Image analysis has proven to be at least as effective as human analysis for quantifying E-cadherin intensity and has identified a cut-point which results in a statistically significant correlation with survival times. UCB stained for E-cadherin has been used as a model to illustrate the fully automated image analysis of immunohistochemistry. However, this model can be applied to all types of immunohistochemically stained tissue, as illustrated by the assessment of ß-catenin. Utilising image analysis we found E-cadherin and ß-catenin to have similar prognostic values. Fully automated workflows such as the one outlined, which is the first of its type, are unbiased, fast, efficient, quantitative and automated.

The quantification of immunostain intensity is often a critical parameter in deciding the nature of cancer treatment for patients, for example in the administration of Herceptin™. Suitability for Herceptin™ is assessed by immunohistochemical analysis,

utilising a three-point scale similar to the one describe in this study. However, when samples are deemed to be moderately stained, further analysis is performed by FISH for definitive results. In light of the findings of this study, it is suggested that the application of image analysis of immunohistochemically-stained tissue could decrease the number of cases sent for FISH, therefore increasing throughput and reducing costs.

CHAPTER 6: HIGH-THROUGHPUT AUTOMATED IMAGE ANALYSIS FOR ASSESSING HER-2 STATUS IN BREAST CARCINOMA; A STUDY INVOLVING ANALYSIS OF TISSUE MICROARRAYS AND WHOLE SECTIONS

6.1   Introduction

HER-2/neu is a member of the epidermal growth factor (HER-2) receptor family. Its gene is located on the long arm of chromosome 17 and encodes a 185-kDa transmembrane glycoprotein with tyrosine kinase activity. It has been found that the 17q region is susceptible to rearrangement mutations in breast carcinomas (Popescu et al., 1989). The amplification and overexpression of HER-2 is associated with a shorter disease-free interval, shorter overall survival, higher incidence of metastasis, reduced response to certain chemotherapeutics and a more aggressive disease progression regardless of disease stage or nodal status (Andrechek and Muller, 2000, Climent et al., 2001, Camp et al., 2003, Tawfik et al., 2006). HER-2 is amplified and overexpressed in 10 – 30% of women with breast cancer (Slamon et al., 1987, Slamon et al., 1989, Borresen et al., 1990).

The clinical significance of HER-2 gene amplification as a proto-oncogene in breast cancer was first described in 1987 (Slamon et al., 1987, Winer and Burstein, 2001). Trastuzumab (Herceptin[TM]), a humanised monoclonal antibody targeting the oncogene, showed therapeutic benefits especially when combined with conventional chemotherapeutic agents in treatment of patients with metastatic breast carcinoma (Camp et al., 2003, Slamon et al., 1987, Seidman et al., 2001, Winer and Burstein, 2001, Baselga et al., 1996, Carter et al., 1992). In 1998, the Food and Drug Administration approved Herceptin[TM] (Genentech, USA) targeted to the protein product of HER-2 gene, as a breast cancer therapeutic agent. Herceptin[TM] was later approved by the European authorities in 2000 (Winer and Burstein, 2001). The association between HER-2 protein overexpression and Herceptin[TM] response has stimulated renewed interest in accurately assessing HER-2 status in breast cancer patients (Camp et al., 2003).

Slamon et al. (2001) evaluated the clinical benefits of treatment with Herceptin[TM] for women with HER-2 positive metastatic breast carcinoma, by comparing the benefits of chemotherapy alone, and chemotherapy combined with Herceptin[TM] (Slamon et al., 2001). The benefits of Herceptin™ included response rates of up to 50% when

combined with chemotherapy (paclitaxel and anthracycline plus cyclophosphamide) versus 32% with chemotherapy alone (p < 0.0001). There was also a significant increase in survival time (25 versus 20 months, p < 0.05) when assessed at 35 months of follow up, when chemotherapy and Herceptin[TM] was used (Slamon et al., 2001, Winer and Burstein, 2001).

Original investigations of the effect of Herceptin™ on metastatic breast cancer illustrated a high incidence of congestive heart failure among patients who had received Herceptin[TM] in conjunction with alternative chemotherapeutics, particularly anthracycline (Slamon et al., 1987). In addition, HER-2 amplification has been associated with a negative response to widely used hormonal drugs such as tamoxifen (Muss et al., 1994). Therefore, it is imperative that patients suitable for Herceptin[TM] treatment are correctly identified.

A correlation exists between HER-2 gene amplification and an increased level of HER-2 protein overexpression (Slamon et al., 1989, Press et al., 2002). Current methods for evaluating HER-2 protein overexpression and gene amplification include measuring protein expression by immunohistochemistry, measuring gene copy number by either fluorescent *in situ* hybridisation (FISH) or chromogenic *in situ* hybridization (CISH), measuring the antigen in the serum by enzyme-linked immunosorbent assay (ELISA), by quantitative polymerase chain reaction methods and by Southern blot analysis. Analysing for protein overexpression or gene amplification of HER-2 in surgical specimens is most commonly accomplished by either immunohistochemistry or FISH testing (Tawfik et al., 2006). Establishing tumour HER-2 status is a prerequisite when evaluating the suitability of candidates for the administration of Herceptin™ (Press et al., 1994). However, a widely accepted and standardised single test for HER-2 status is not currently available (Rampaul et al., 2002, Ellis et al., 2000).

When utilising immunohistochemical means to assess HER-2 protein overexpression, the majority of laboratories use HercepTest™ kit (DakoCytomation, UK)(Rhodes et al.,

2002). Only invasive carcinoma regions are assessed and are scored semi-quantitatively (0-3+), based on intensity and percentage of membrane-positive cells. The semi-quantitative scale used is interpreted as follows: No staining or weak membrane staining in < 10% of tumour cells (0/1+); therefore, the HER-2 protein is not overexpressed and patients do not qualify for Herceptin[TM] treatment. A strong complete membrane stain is observed in > 10% of tumour cells (3+); therefore, the HER-2 protein is overexpressed and patients qualify for Herceptin[TM] treatment. Cases with weak to moderate staining in > 10% of tumour cells are considered intermediate stained (2+) and require further definitive testing by FISH (Ellis et al., 2000, Ellis et al., 2004). As all cases are not routinely tested for FISH, there is a risk of under diagnosing cases with (0-1+) HER-2 protein expression that have gene amplification or, alternatively, over diagnosing cases with 3+ protein overexpression with no gene amplification. As a consequence, patients could be either not offered treatment that may be beneficial or exposed to inappropriate and toxic treatment (Tawfik et al., 2006).

There are several commercial systems available for FISH HER-2 testing (Abbott-Vysis, Cancer Genetics, DakoCytomation and QBioGene); however, recommendations on HER-2 testing in the UK advocate the PathVysion[®] assay (Vysis Inc, UK) for diagnostic testing (Ellis et al., 2000, Ellis et al., 2004). The scoring protocol with the PathVysion[®] assay recommends the number of chromosome 17 and HER-2 gene signals are scored for 60 cells, where possible from three distinct tumour fields, and the mean HER-2 gene to chromosome 17 copy ratio is calculated. Where tumour heterogeneity is seen in 1 – 2% of cases, more cells from the amplified regions should be scored. Samples with $\geq$ 2.0 copies of HER-2 gene for each chromosome 17 are considered to be amplified and those with < 2.0 are not gene amplified (Ellis et al., 2000). If the ratio of HER-2 gene to chromosome 17 is close to 2.0 (1.5 - 2.3) a larger number of cells should be assessed.

There is no consensus on which method of assessment of HER-2 status is most predictive of patients' response to Herceptin[TM] in a clinical setting (Tawfik et al., 2006). Generally, FISH is considered the gold standard, as it is truly quantitative; and therefore, less susceptible to inter- and intra-observer variability. However, FISH is a highly

skilled technique, which is not widely available in routine pathology laboratories. Fluorescent probes are expensive, and signal is lost over time; therefore, expensive fluorescent image capturing devices must be used if results need to be preserved for future reference. In contrast, immunohistochemistry is relatively inexpensive and widely available in routine pathology laboratories. However, immunohistochemistry is susceptible to inter- and intra-observer variability and is semi-quantitative at best. There has been some concordance identified between the two assays, particularly with the strongly positive (3+) cases. However, interpretational problems still exist for the 1+ and 2+ cases (Vera-Roman and Rubio-Martinez, 2004, Tawfik et al., 2006). The majority of the tumours that overexpression HER-2 protein (3+) will also have gene amplification (85 – 90%), and only a small percentage of the cases (5 – 25%) with borderline HER-2 overexpression (2+) have gene amplification (Ridolfi et al., 2000, Vincent-Salomon et al., 2003, Tawfik et al., 2006).

Inter-observer agreement is significantly higher when assessing HER-2 status using FISH testing rather than immunohistochemistry (Bartlett et al., 2001). Bartlett et al. (2001) illustrated inter-observer agreement was lower for the immunohistochemical HercepTest™ (kappa = 0.67) than for FISH testing by PathVysion$^{®}$ assay (kappa = 0.97), when assessing the same cohort of patients. Using immunohistochemistry, 81% concordance was observed between two reviewers when classifying normal expression (0/1+) or borderline/overexpression (2+/3+). Whereas using FISH, concordance when classifying amplified ($\geq$ 2) and non-amplified (< 2) cases was greater than 99% (Bartlett et al., 2001, Ellis et al., 2004).

Lacroix-Triki et al. (2006) evaluated the inter-observer agreement when interpreting HER-2 immunostains performed in different laboratories, according to their in house techniques. A total of 74 HER-2 immunostains were evaluated by 16 pathologists and by a central review committee. The overall agreement was good (kappa = 0.75). Agreement was excellent for the 0/1+ group (kappa = 0.85) and for the 3+ group (kappa = 0.82). As expected, when borderline expression was present (2+), poor agreement was observed (kappa = 0.38) (Lacroix-Triki et al., 2006).

Commercially available systems are in existence which quantify HER-2 gene amplification and protein overexpression, when evaluated by FISH and immunohistochemistry. High levels of correlation have been found between HercepTest™ and FISH analysis when immunohistochemical staining was assessed by automated systems (Wang et al., 2001, Luftner et al., 2004). Micrometastasis Detection System (MDS™, Applied Imaging, USA) is a computer controlled scanning microscope which captures either brightfield or fluorescence. In addition, MDS™ provides quantitative analysis of immunohistochemical or FISH images. Utilising MDS™, it was found that FISH proved to be a more accurate and consistent scoring system for determining HER-2 status compared to immunohistochemistry. In general, the authors found MDS™ to be more reliable and more consistent than visual FISH and immunohistochemical scoring systems (Ellis et al., 2005).

Automated Cellular Imaging Systems (ACIS) and Automated Quantitative Analysis (AQUA) have been used to evaluate HER-2 status. Tawfik et al. (2006) found a correlation of 94% between ACIS analysis of HercepTest™ and FISH (Tawfik et al., 2006). The number of patients identified with HER-2 protein overexpression and gene amplification differed only by 1 case, when immunohistochemical staining was assessed by ACIS (Tawfik et al., 2006). In addition, high correlation between human scoring and the ACIS system has been observed (95.1%, kappa = 0.85) (Luftner et al., 2004). Using image analysis, it is possible to identify extremely weak staining, which is not observed when immunohistochemistry is assessed by human scoring (Camp et al., 2003). In addition, studies have suggested that automated image analysis could be used to identify potential cases requiring FISH testing, rather than sending all borderline immunohistochemical cases (Ciampa et al., 2006).

Within this chapter, the application of the previously described image analysis algorithm to quantify HER-2 protein overexpression will be assessed. HER-2 protein overexpression, when assessed by human and image analysis, will be compared to HER-2 gene amplification data. In order to create and validate optimal cut-points, a training set and validation set of data will be used. Tissue Microarrays will be utilised as a

training set to create optimal cut-points within the continuous image analysis dataset. Whole section biopsies will be used as a validation set in order to validate the application of the algorithm. Results will indicate if a low cost algorithm could potentially be used to evaluate HER-2 protein overexpression.

6.2 Image Analysis Considerations

As previously described the image analysis algorithm was created in Image-Pro Plus[®]; however, due to file size restrictions it is not possible to perform truly high-throughput analysis across large tissue sections using Image-Pro Plus[®]. Distiller (SlidePath, Ireland) provides the capabilities of rapidly processing large images; however, due to software restrictions it is not possible to use Image-Pro Plus[®] in combination with Distiller. Therefore, the algorithm was re-created using an open-access software language (Java). However, morphological filters that are common to both applications may differ slightly, and a number of image manipulations available in Image-Pro Plus[®] are not available in Java. Therefore, the algorithms are not identical. Utilising the code created in Java, two scenarios occurred that were not observed when using the algorithm generated in Image-Pro Plus[®]. The two scenarios incurred include, the presence of staining artifacts and the presence of noise within the final output image, a description of both follow:

6.2.1 Staining Artifacts

Prior to dissection, edges of whole section biopsies are stained with tissue dye. The biopsies are then paraffin-embedded, sectioned using a microtome and transferred to a glass slide. On the surface of the slide, tissue dye is clearly visible around the boundary of the biopsy. The presence of tumour cells near the boundary suggests not all tumours regions were removed during surgery. The colour of the dye used to stain tumours varies with tumour type. Breast cancer is typically stained using a black/brown dye. This procedure for processing biopsies influenced image analysis assessment during this study.

Figure 6.2-1 (A) illustrates a section of immunohistochemically stained tissue; a verge of dark black/brown tissue dye is clearly visible. During image analysis, the black/brown verge was considered "edge" and "brown" and, therefore, interpreted as membrane staining. As a result, the colour cubed used within the algorithm was altered in order to exclude areas of black/brown tissue dye and image analysis results were corrected to account for this (see section 6.2.2). Therefore, the utility of tissue dye did not negatively

affect image analysis results, as evident from Figure 6.2-1 (B), where none of the tissue dye has been isolated as membrane staining when assessed by image analysis.
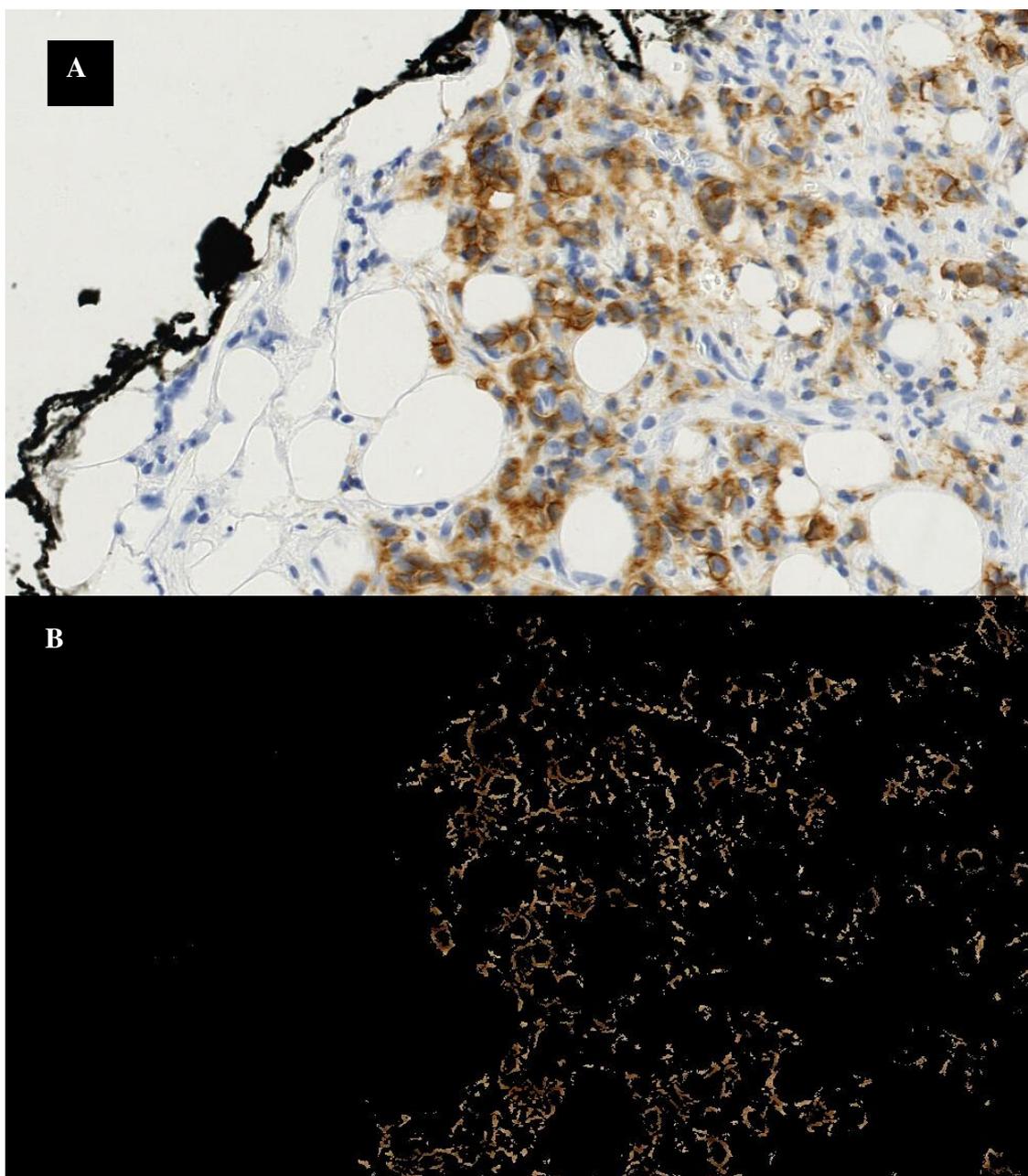
Figure 6.2-1: (A) Illustrates tissue immunohistochemically stained for HER-2. A black/brown verge of tissue dye that is routinely applied to tumour boundaries prior to tissue processing is visible within this slide. (B) Illustrates the final output image after image analysis. The colour cube was adjusted to exclude black/brown pixels, and therefore exclude the tissue dye. As a result, only areas of membrane brown staining are recorded.

6.2.2   Noise

The range of "brown" staining which encompasses HER-2 immunohistochemical staining was much greater than previously assessed antigens within this thesis. As a result, the colour cube was adjusted to encompass a large range of colour intensities. This introduced a certain level of noise into the final output image, when assessed by the algorithm. Noise is any staining that was present in non-membrane regions. However, the staining was observed in regions of brown edges and, therefore, is interpreted as membrane during image analysis. Within this study, where noise was observed, the staining intensity was mostly weak. Weak staining is essential in HER-2 status assessment and is classified as 0/1+ status. However, within strongly stained tissue the weak staining appeared as noise, as is evident from Figure 6.2-2. Figure 6.2-2 illustrates strong HER-2 immunohistochemical staining (A), and the final output after image analysis (B). Both images are digitally zoomed to in order to visualise the level of noise present within the tissue. The red box highlights an area of staining which is not true membrane; however, it has been identified as such by the image analysis algorithm.
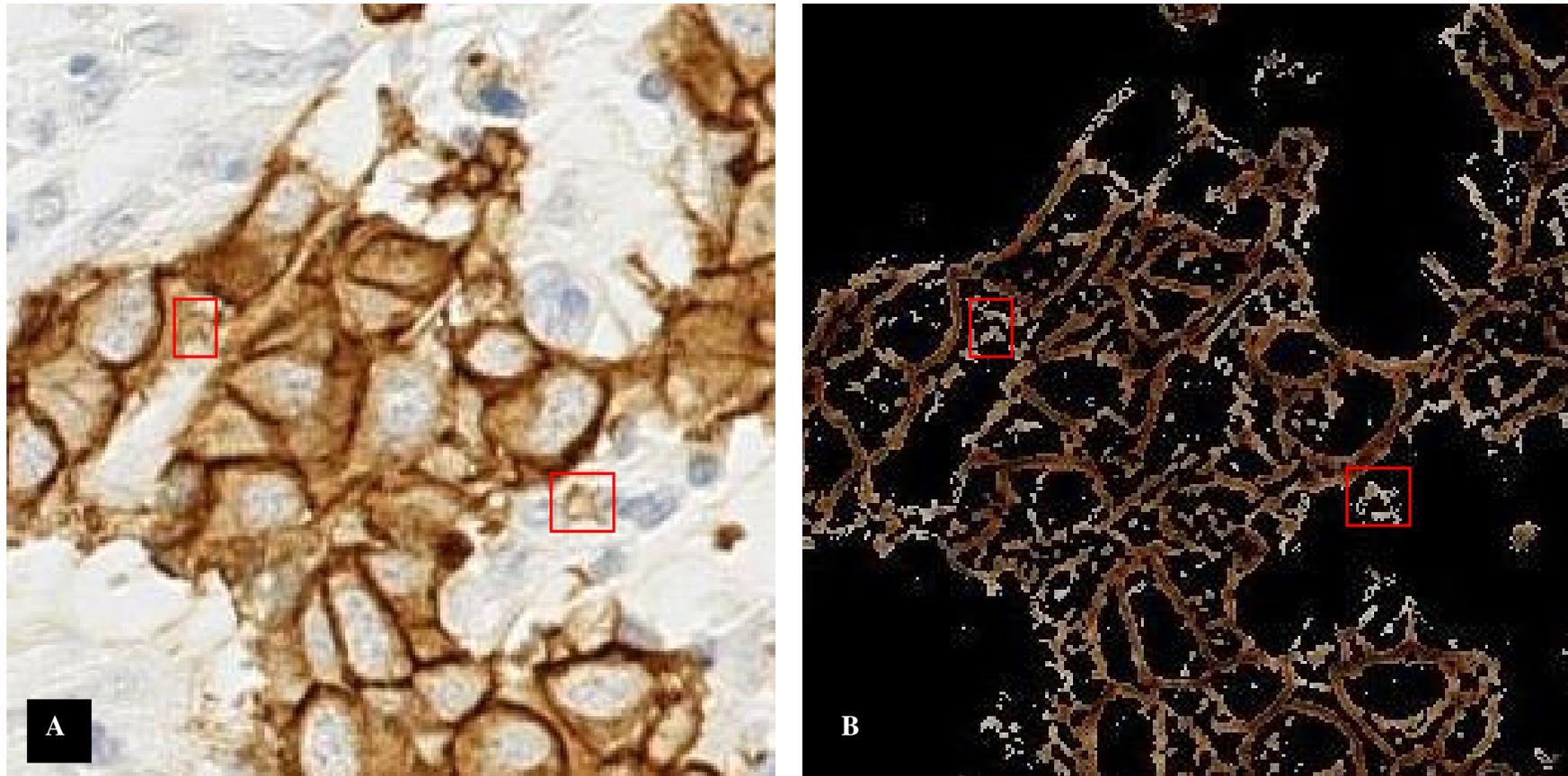
Figure 6.2-2: (A) Tissue immunohistochemically stained for HER-2. (B) Final output after image analysis. Red boxes highlight areas that are identified by image analysis as membrane staining; however, the areas of staining are clearly not membrane. Therefore, signal within these boxes is incorrectly identified as membrane-bound immunohistochemical staining.

Based on these observations it was deemed necessary to exclude the noise from image analysis results. Typically, histograms of staining intensity values are outputted for each sample assessed by image analysis. The modal value (peak) within the histogram represents the overall staining intensity value for the tissue under assessment. Figure 6.2-3 illustrates a typical histogram outputted during image analysis assessment. The x-axis represents the staining intensity scale, which quantifies the intensity of staining, and the y-axis represents the number of pixels staining positive. Figure 6.2-3 illustrates a normal distribution of staining intensity, where there is no evidence of noise, as only a single peak within the histogram is present. Therefore, the maximum bin within the histogram is recorded as the staining intensity value for the tissue under assessment.
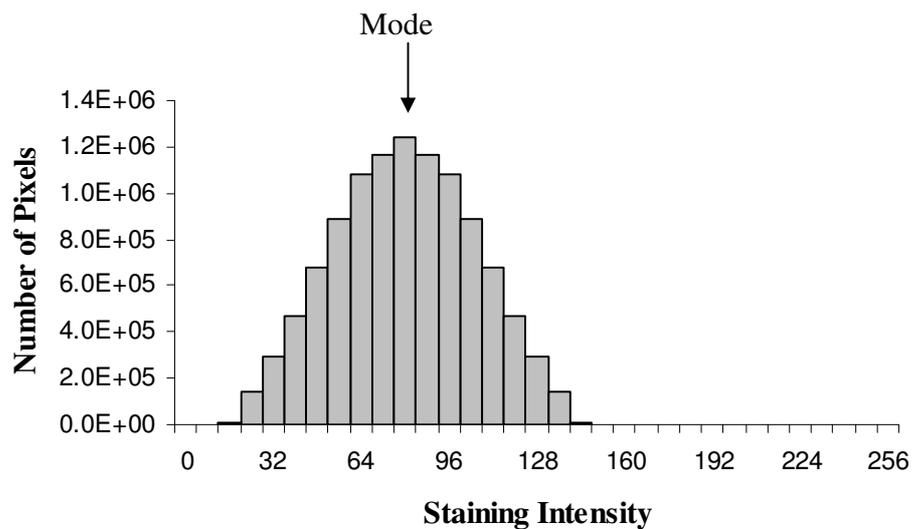
Figure 6.2-3: Histogram illustrating the number and staining intensity of membrane-pixels when assessed by image analysis. The histogram has normal distribution, with a single peak. Therefore, the peak of the histogram is recorded as staining intensity value for the tissue under assessment.

More than one peak within the histogram signifies the presence of noise within the final output image when assessed by the algorithm. Figure 6.2-4 illustrates a histogram with non-normal distribution, where three peaks of staining intensity are visible. The three peaks and their distribution have been coloured coded within the figure. A second peak

located toward the lighter end of the staining intensity scale (256) suggests noise due to non-membrane staining detection, visualised in white in Figure 6.2-4. Non-membrane noise was illustrated in Figure 6.2-2 (B). A peak located toward the stronger end of the staining intensity scale (0), suggests noise due to the identification of ink that is applied to tumour borders, visualised in black in Figure 6.2-4. Tumour ink was illustrated in Figure 6.2-1 (A). However, as the colour cube was adjusted to discount extremely black pixels, the occurrence of this type of noise was extremely infrequent. The central peak within Figure 6.2-4 (grey) represents the actual membrane staining presence. Within Figure 6.2-4, the two modal values of the noise were less than the actual membrane-bound immunohistochemical signal. Therefore, in this case the presence of noise did not affect the overall staining intensity value; however, the presence of noise did affect the total positive pixel count.
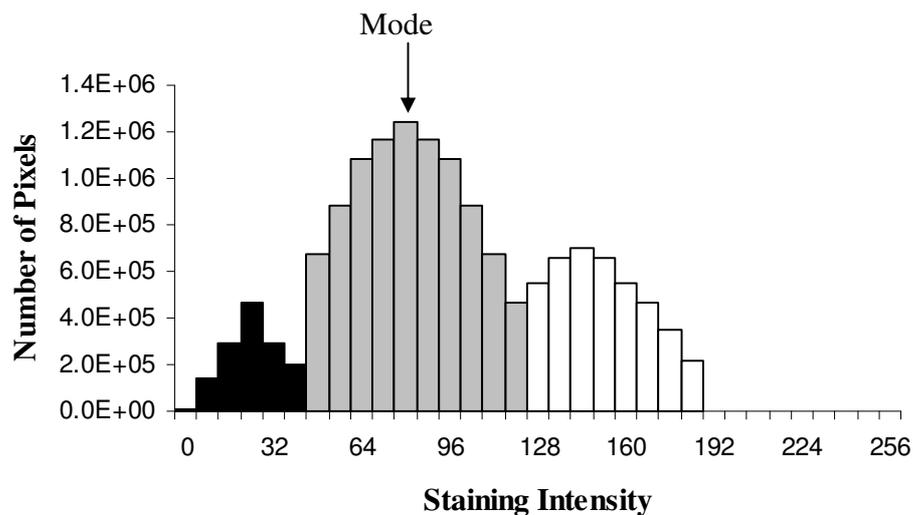
Figure 6.2-4: Tri-modal distributed histogram. Within the histogram there are 3 distributions membrane-bound immunohistochemical staining (grey), and noise due to the presence of: tissue dye (black) and non-membrane staining (white). In this case the mode of membrane-bound staining is greater than the noise; therefore, the overall mode value in this is correct. However, overall the number of membrane positive pixels is artificially inflated due to the number of noise pixels.

Figure 6.2-5 illustrates a case where the mode of membrane-bound staining intensity (grey) is considerably less than either of the noise signals (black and white). In this case, taking the overall modal value of the histogram would incorrectly classify the membrane-bound immunohistochemical staining intensity value.
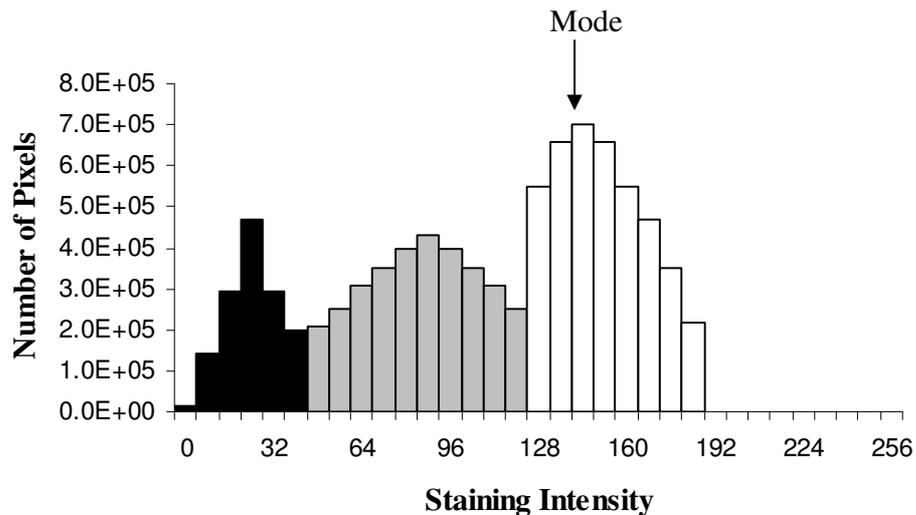
Figure 6.2-5: Non-normal distributed histogram. Within the histogram there are 3 distributions membrane-bound immunohistochemical staining (grey), and noise due to the presence of: tissue dye (black) and non-membrane staining (white). In this case the mode of membrane-bound staining is less than the noise; therefore, the overall mode value in this is incorrect.

Therefore, in order to account for non-normal distribution, correction of data was performed. Correction was performed by modelling and subtracting the noise pixels from the histograms, therefore removing any non-membrane bound signal from the dataset. This resulted in only membrane positive pixels, remaining in the histogram. Therefore, the mode and number of positive pixels could be utilised to record the amount and intensity of HER-2 membrane bound staining.

6.3   Study Design

6.3.1   Patient Cohort

Forty-seven consecutive cases of breast carcinoma were selected from the files of the Histopathology Department of Beaumont Hospital. Three Tissue Microarrays were constructed by sampling 2 x, 4 x and 6 x cores (0.6 mm diameter) from the individual blocks using the technique pioneered by Kononen et al. (1998) and described in Chapter 2 section 2.2.1.2. A total of 188 cores were included in this study and the distribution of the replica cores is illustrated in Table 6.3-1. One orientation spot composed of liver tissue was present on all slides. Replica cores were sampled from different tumour areas to account for heterogeneity in any one tumour and to minimise the number of lost cases during subsequent processing of the Tissue Microarrays. Tissue Microarrays were utilised as a training set to create optimal cut-points, based on comparisons with FISH data. A total of 46 different biopsies were also sectioned. Whole section biopsies were used to blindly test the application of the image analysis algorithm. Tissue Microarrays and whole sections were created as described in Chapter 2 section 2.2.1.1, and all slides underwent immunohistochemical staining and fluorescent *in situ* hybridisation as described in Chapter 2 section 2.3.1.2 and 2.3.4, respectively. Tissue Microarrays and whole sections were digitised using a NanoZoomer Digital Pathology (NDP) System (Hamamatsu, UK) as described in Chapter 2 section (2.4.3). Fluorescent *in situ* hybridisation was visualised using a Nikon Eclipse E600 microscope, as described in Chapter 2 section (2.4.4).

Table 6.3-1: Three Tissue Microarrays were assessed within this study. Replica cores were extracted from 47 biopsies to generate the Tissue Microarrays. The number of replica cores extracted from each biopsy is illustrated within this Table.

Replication	Number of Patients	Number of Cores
2 x	1	2
4 x	45	180
6 x	1	6
Total	47	188

6.3.2   Human Classification

Scoring of the HercepTest™ immunohistochemically stained Tissue Microarrays was performed by one research scientist and one pathologist both were experienced in the review of immunohistochemistry, on a three point scale (0-3+) using the manual for interpretation provided by the manufacturer. Reviewer 1 (research scientist) assessed all Tissue Microarrays using microscope-based review. Reviewer 2 (pathologist) assessed all Tissue Microarrays using virtual slide methods of review. A combination of two individual reviews was used as the Tissue Microarray scores. In cases were the two individuals evaluations differed, the stronger staining intensity value was recorded as the overall result. However, inter-observer variability for the assessment of Tissue Microarrays was considerable less than those reported in previous chapters (kappa = 0.845). Scoring of whole section biopsies was performed by a Professor of Pathology and a research scientist. Reviews were performed concurrently using virtual slides. For the purposes of Herceptin™ treatment, a HercepTest™ score of 0 /1+ is interpreted as negative for HER-2 protein overexpression. A score of 3+ is interpreted as positive for HER-2 protein overexpression, and a score of 2+ is inconclusive and must undergo FISH testing, for definitive results.

Scoring of the tissue which underwent FISH testing was performed by one research scientist (Reviewer 1), across all whole sections and Tissue Microarrays. PathVysion$^®$ manual states that a ratio of $\geq$ 2 is evidence of HER-2 gene amplification, and < 2 is evidence of non-amplification. Recommendations for interpretation of FISH state that if the ratio of HER-2 gene to chromosome 17 is close to 2.0 (1.5 - 2.3), a larger number of cells should be scored (Ellis et al., 2004).

6.3.3   Image Analysis Classification

Tissue Microarrays generated from 47 patients biopsies were used to train the image analysis algorithm to classify HER-2 protein expression. Staining intensity values were quantified using image analysis, and the normalised image analysis results were compared with gene amplification data quantified by FISH. Utilising FISH data it was possible to create optimal cut-points within the normalised image analysis data. The output from image analysis was categorised into positive or negative for HER-2 protein overexpression, based on the modal staining intensity value and the number of positive membrane pixels. The modal staining intensity value was normalised across all slides, by obtaining the natural log (Ln) of the modal staining intensity value (I) divided by the background lighting intensity (Io), (Ln I/Io). Samples scoring a normalised staining intensity value of $\geq$ 0.54 were considered positive for HER-2 protein overexpression. Samples scoring normalised staining intensity value of < 0.54 were considered negative for HER-2 overexpression. The value of 0.54 was obtained as it is the mid-point between 0.47 and 0.59. Which is the split point where the least amount of cases were in disagreement between FISH and image analysis data. Tissue Microarray cores and whole biopsies with < 700 and < 100,000 membrane positive pixels respectively, were considered negative for HER-2 protein overexpression. The classification system used by HercepTest™, FISH and image analysis to quantify HER-2 expression is illustrated in Figure 6.3-1.
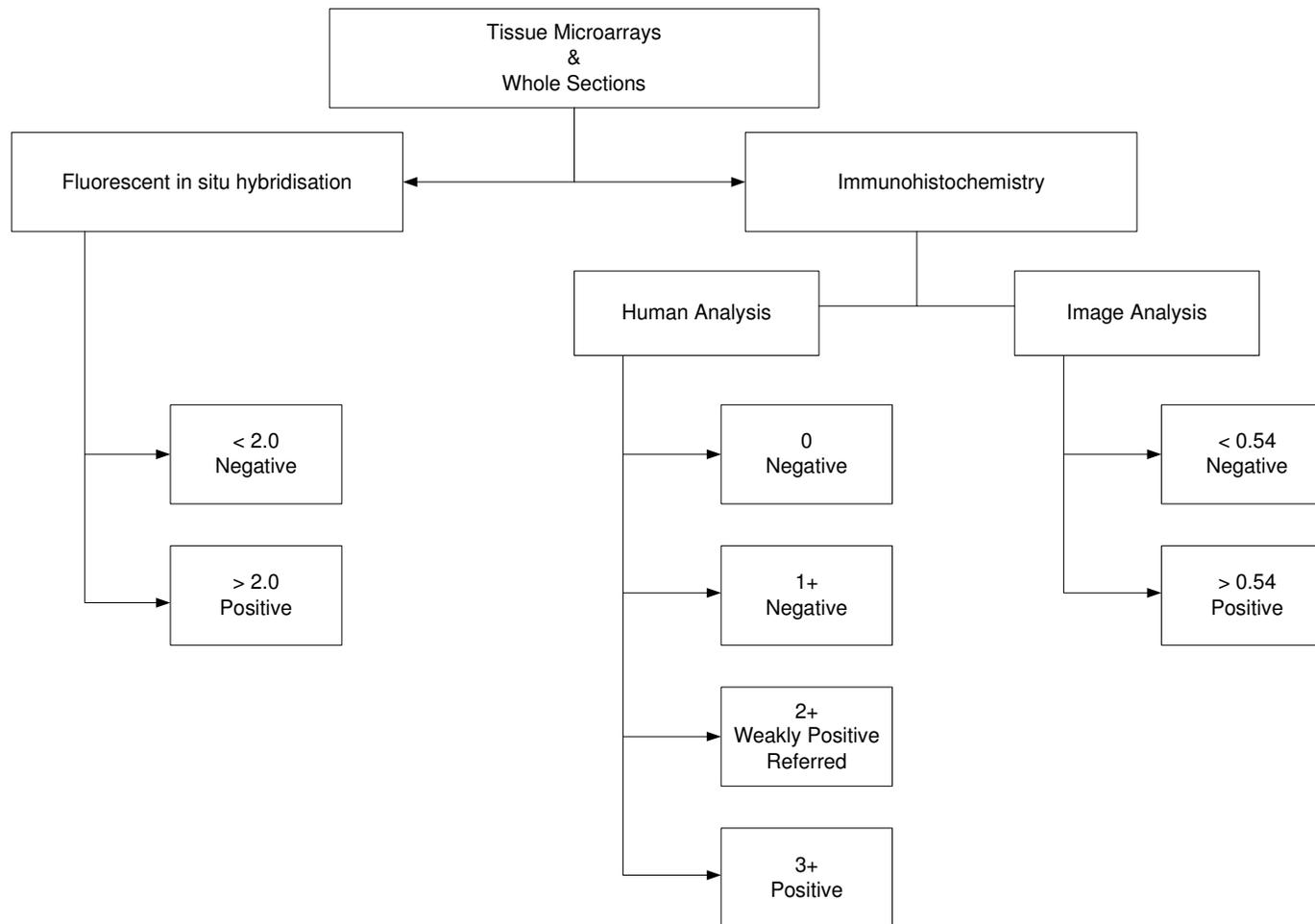
Figure 6.3-1: Classification of HER-2 gene amplification when assessed by fluorescent *in situ* hybridisation (FISH), and HER-2 protein overexpression when assessed by immunohistochemistry and interpreted by human and image analysis, across all Tissue Microarrays and whole sections within this study.

6.4 Results

The ability of the image analysis algorithm to classify HER-2 status within breast carcinoma, across Tissue Microarrays and whole sections were assessed. All slides were immunohistochemically stained with HercepTest™, and levels of HER-2 protein overexpression were assessed by four individuals and image analysis. Fluorescent *in situ* hybridisation was performed across all slides and levels of HER-2 gene amplification were assessed by one individual. Tissue Microarrays were used as a training set to create optimal cut-points within the continuous image analysis dataset. Whole sections were used as a validation set to verify the optimum cut-point to create.

6.4.1 Tissue Microarrays - Training Material

Cut-points were selected by comparing image analysis results with FISH data generated from Tissue Microarray reviews. Immunohistochemical reviews of Tissue Microarrays were performed by two individuals, and overall results were an average of the analyses performed. HER-2 gene amplification was assessed by FISH and quantified by a single individual.

Figure 6.4-1 illustrates the correlation between FISH and image analysis across three Tissue Microarrays. The two cases where disagreement between FISH and image analysis assessment of HER-2 occurred is highlighted within the black box. In both cases, HER-2 gene amplification was observed by FISH; however, HER-2 protein overexpression was not present when immunohistochemical staining was assessed by human or image analysis. In both cases, HER-2 gene amplification was extremely close to the FISH analysis cut-point (2.2 and 2.11).
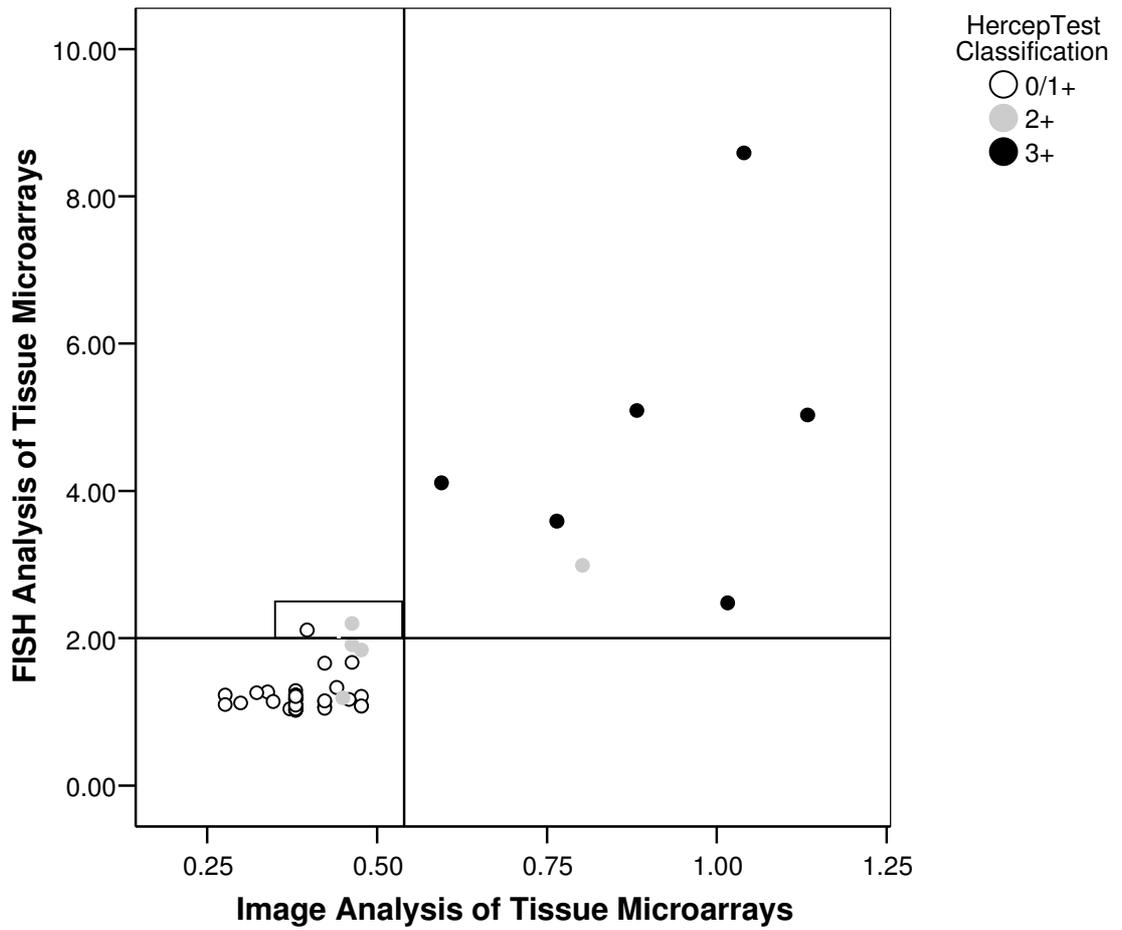
Figure 6.4-1: Correlation between image analysis and fluorescent *in situ* hybridisation (FISH) assessment of HER-2 status across three Tissue Microarrays. FISH analysis classifies HER-2 gene amplification as: < 2.0 non-amplified and ≥ 2.0 amplified. FISH cut-point (2.0) is illustrated by the horizontal line. Image analysis classified immunohistochemical HER-2 protein expression as: < 0.54 non-overexpressed and ≥ 0.54 overexpressed. Image analysis cut-point (0.54) is illustrated by the vertical line. Data points have been colour coded according to the human HercepTest™ review. Cases where FISH and image analysis disagreed are highlighted within the black box.

Agreement between HER-2 protein overexpression and gene amplification across three Tissue Microarrays and whole section biopsies is illustrated in Table 6.4-1 and Table 6.4-2, respectively. Depending solely on immunohistochemical staining to classify HER-2 status resulted in three scenarios with respect to patient treatments. "Correct treatment" occurred when HER-2 status assessed by immunohistochemistry and FISH agreed. "Referred cases" occurred when moderate immunohistochemical staining (2+) was observed and definitive testing was required to clarify HER-2 status. Referred cases would results in HER-2 status being accurately assessed; therefore referred cases have been classified with the "Correct Treatment". "Over treated" occurred when HER-2 protein overexpression was observed but the gene was not amplified, therefore, patients would incorrectly receive Herceptin™. "Under treated" occurred when HER-2 protein was not overexpressed and the gene was amplified, therefore, patients would incorrectly not receive Herceptin™.

During Tissue Microarray analysis, there were no instances of over treatment when immunohistochemical staining was assessed by either human or image analysis. There was a slightly higher occurrence of under treatment when immunohistochemistry was assessed by image analysis (4.3%, n = 2) compared to human analysis (2.1%, n = 1). The number of correctly treated cases was marginally greater when immunohistochemistry was assessed by human analysis (97.9%, n = 46) compared to image analysis (95.7%, n = 45). The number of referred cases (n = 5) was relatively low.

197

Table 6.4-1: Agreement between HER-2 gene amplification (FISH) and protein overexpression (immunohistochemistry) when assessed by human and image analysis across three Tissue Microarrays. Tissue Microarrays were used as a training set to create optimal cut-points within the image analysis dataset. Over treated occurred when protein overexpression was observed and gene amplification was not present. Correct treatment occurred when HER-2 status assessed by immunohistochemistry and FISH agreed. Under treated occurred when gene amplification was observed but protein overexpression was not present. Referred cases occurred when borderline immunohistochemical staining (2+) was observed. Referred cases were categorised under Correct Treatment.

TMAs	Over Treated		Correct Treatment			Under Treated	
	%	n	*%*	n	*Referred Cases*	*%*	n
Human Review	0.0	(0/47)	97.9	(46/47)	*5 of 46 (10.6 %)*	2.1	(1/47)
Image Analysis	0.0	(0/47)	95.7	(45/47)	*N/A*	4.3	(2/47)

6.4.2   Whole Sections – Validation Material

In order to classify HER-2 protein overexpression across whole section biopsies, the cut-points generated in the assessment of Tissue Microarrays were utilised. Assessment of immunohistochemical staining was performed concurrently by two individuals using a microscope. Figure 6.4-2 illustrates the correlation between FISH and image analysis across 46 whole sections. In total there are 8 cases (17.4%) where FISH and image analysis classification of HER-2 status disagreed, these cases are highlighted within the boxes in Figure 6.4-2.
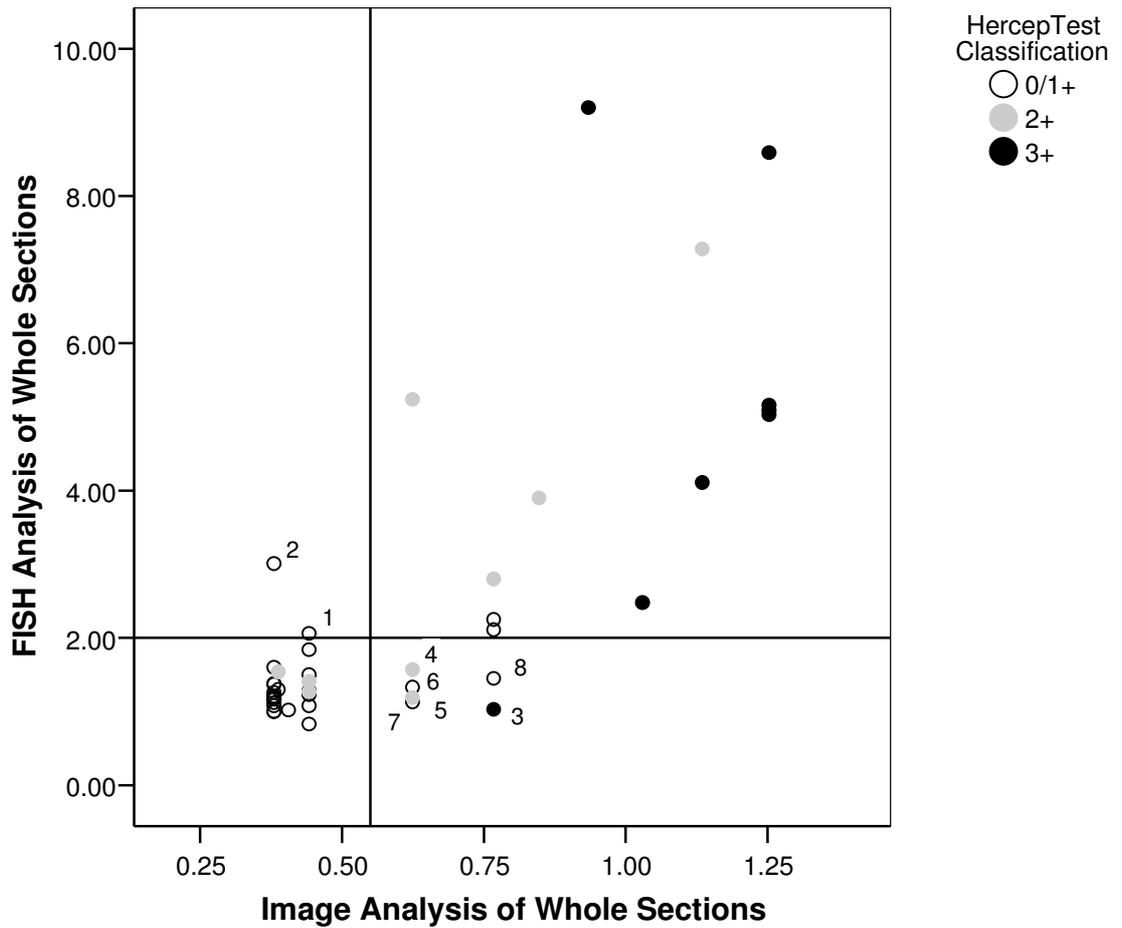
Figure 6.4-2: Correlation between image analysis and fluorescent in situ hybridisation (FISH) assessment of HER-2 status across 46 whole section biopsies. FISH analysis classifies HER-2 gene amplification as: < 2.0 non-amplified and ≥ 2.0 amplified. FISH cut-point (2.0) is illustrated by the horizontal line. Image analysis classified immunohistochemical HER-2 protein expression as: < 0.54 non-overexpressed and ≥ 0.54 overexpressed. Image analysis cut-point (0.54) is illustrated by the vertical line. Data points have been colour coded according to the human HercepTest™ review. Cases where FISH and image analysis disagreed are highlighted within the black box.

The majority of disagreement between FISH and image analysis occurred where HER-2 protein overexpression was observed and gene amplification was not (n = 6). Disagreement between FISH and image analysis classification occurred to due the following reasons:

Case 1:

HER-2 gene amplification was borderline present (2.06) and HER-2 protein not overexpressed (0.44). Recommendations for interpretation of FISH state that if the ratio of HER-2 gene to chromosome 17 is close to 2.0 (1.5-2.3), a larger number of cells should be scored (Ellis et al., 2004). As a result, the FISH analysis of this specimen would be repeated, perhaps changing the HER-2 status.

Case 2:

HER-2 gene amplification was present (3.01); however, clearly no immunohistochemical staining was present when assessed by human and image analysis. Immunohistochemical staining for this case is illustrated in Figure 6.4-3.
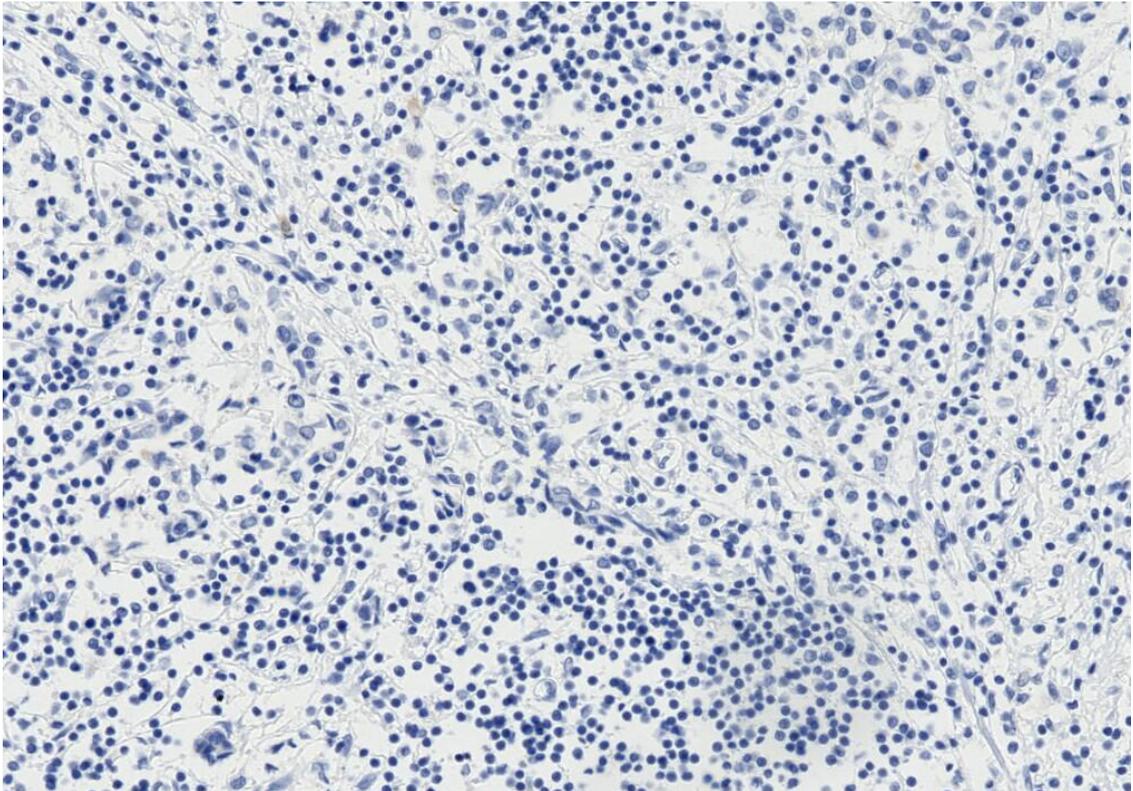
Figure 6.4-3: Tissue recorded as having HER-2 gene amplification when assessed by fluorescent *in situ* hybridisation (FISH) (3.01). However, when immunohistochemical staining was assessed by human (0) and image analysis (< 0.54) protein overexpression was not observed.

Case 3:

HER-2 gene amplification was not observed (1.03); however, HER-2 protein overexpression was clearly visible when assessed by human (3+) and image analysis (0.77). Immunohistochemical staining for this case is illustrated in Figure 6.4-4.
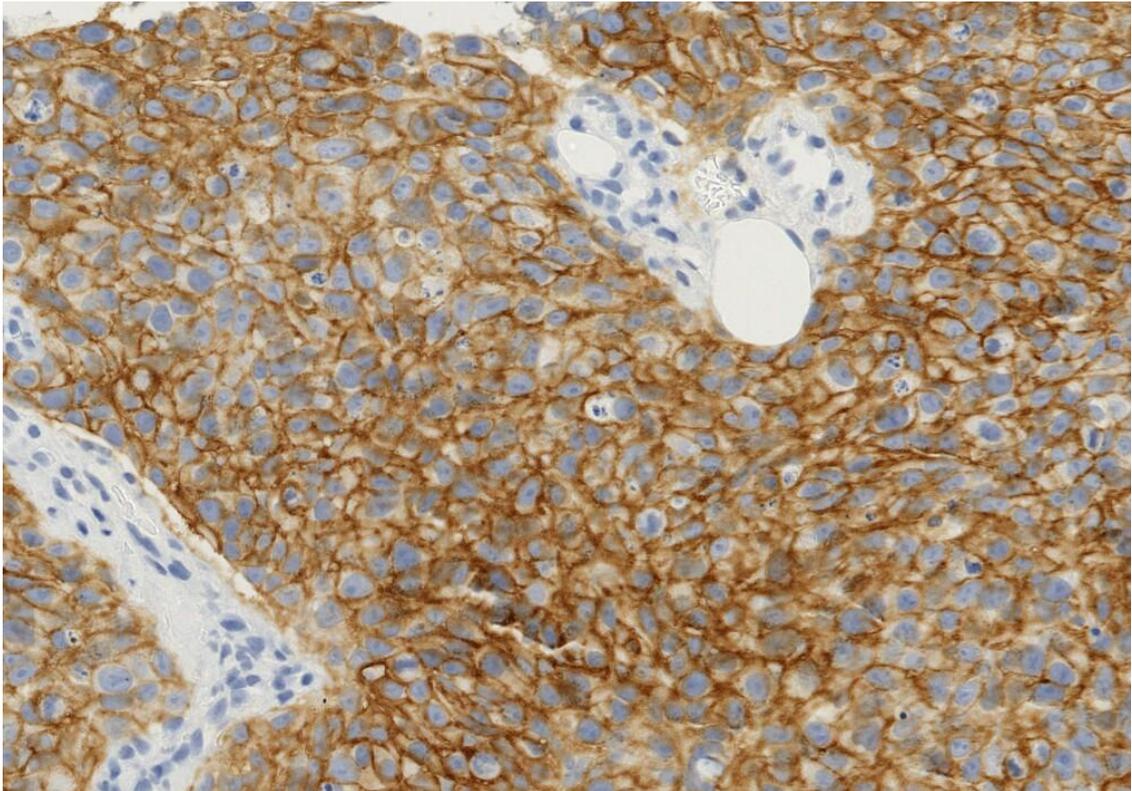
Figure 6.4-4: Tissue recorded as having HER-2 protein overexpression when assessed by human (3+) and image analysis (≥ 0.54). However, when assessed by fluorescent *in situ* hybridisation (FISH) HER-2 gene amplification was not present (1.03).

Case 4/5/6/7:

There were four cases where HER-2 gene amplification was not present; however HER-2 protein overexpressed was observed, and in all cases was recorded as 0.62 when assessed by image analysis. Human analysis classified 2 of these cases as 2+ and 2 cases as 1+. In all 4 cases staining intensity appears to be highly variable with staining intensity ranging from weak to moderate, as is illustrated in Figure 6.4-5 and Figure 6.4-6.
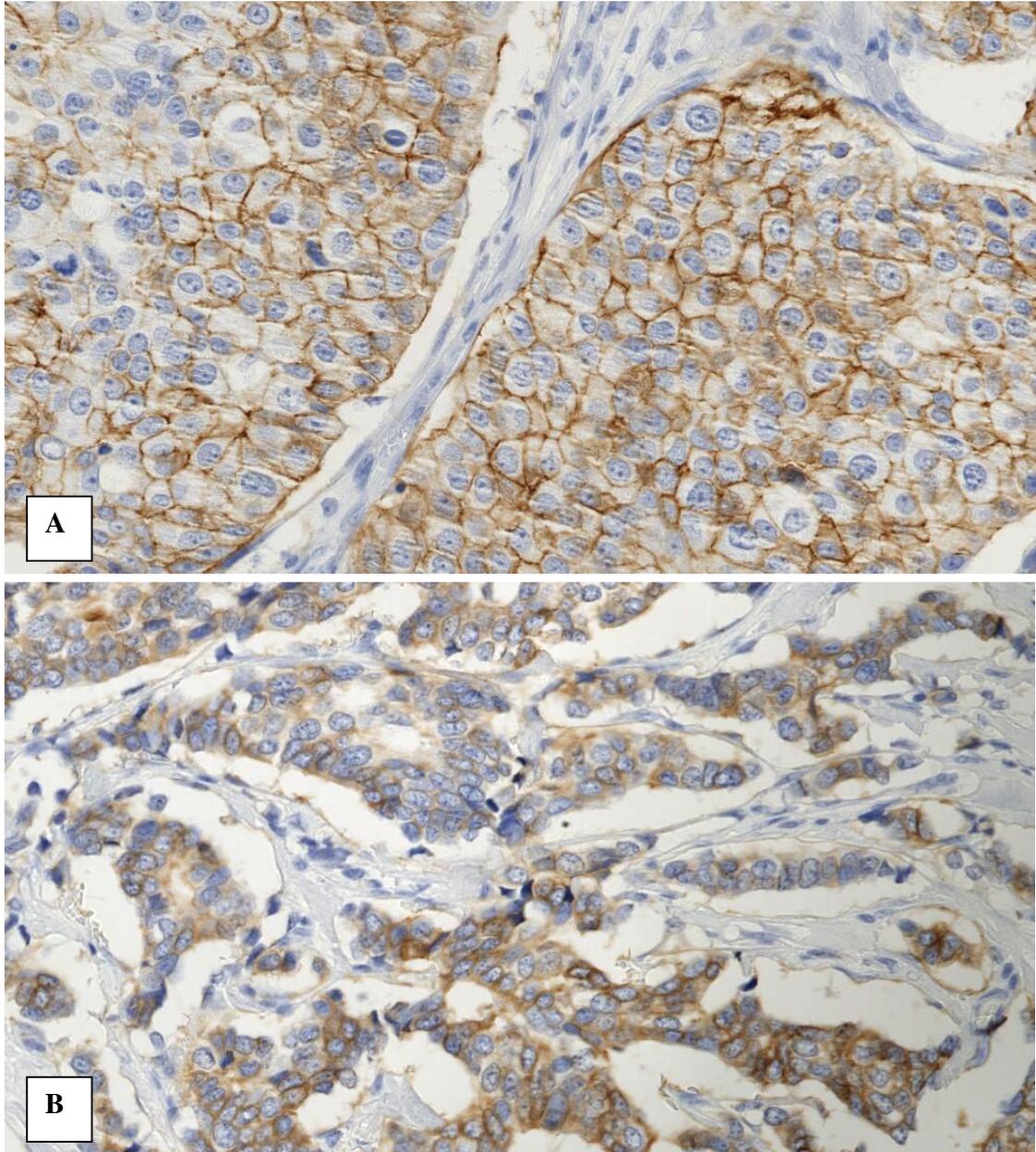
Figure 6.4-5: (A) Case 4: HER-2 gene amplification was not observed (1.57); however, HER-2 protein overexpression was observed when assessed by image analysis (0.62). **(B) Case 5:** HER-2 gene amplification was not observed (1.19); however, HER-2 protein overexpression was observed when assessed by image analysis (0.62). Both cases were referred (2+) when immunohistochemistry was assessed by human analysis.
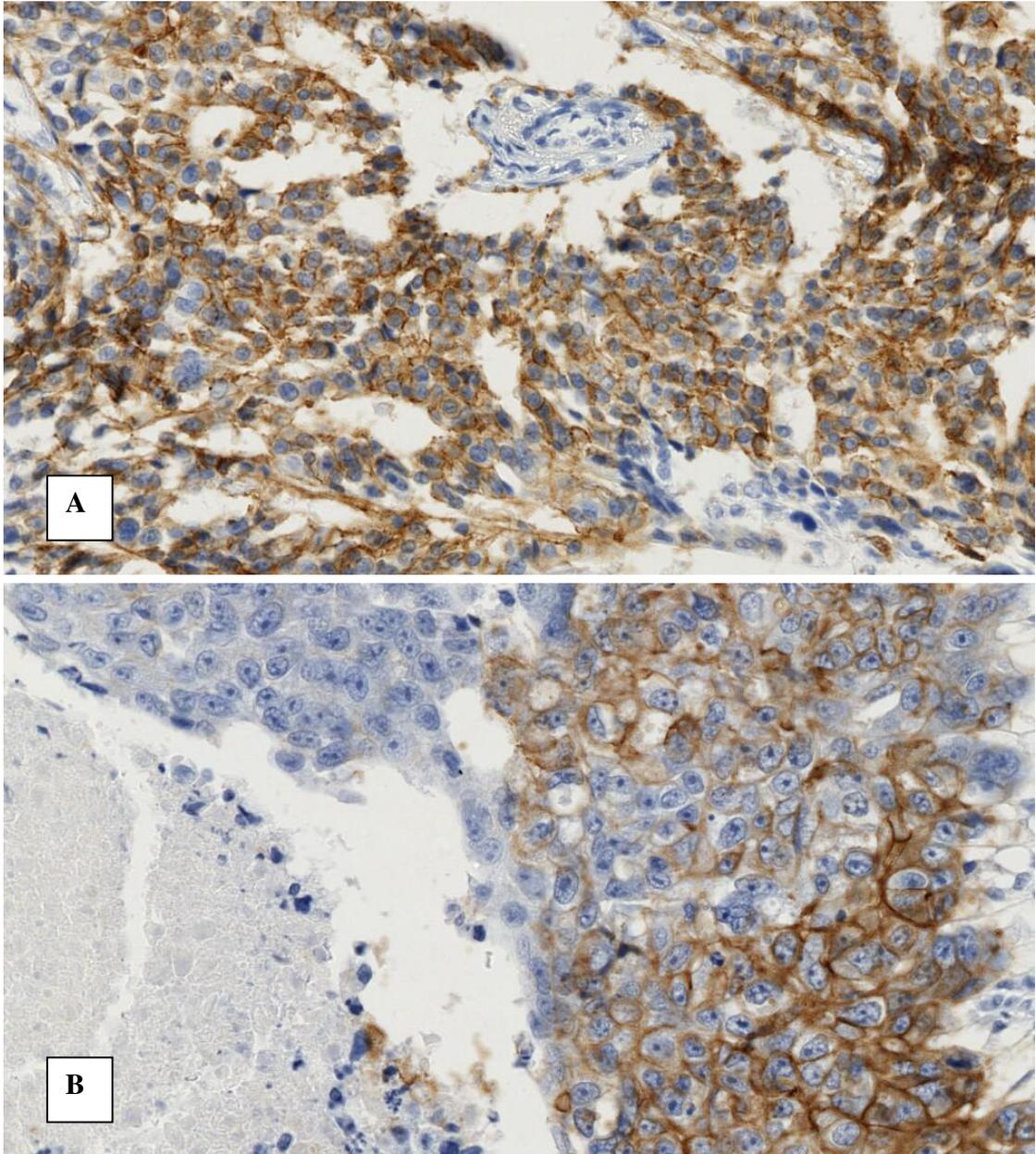
Figure 6.4-6: (A) Case 6: HER-2 gene amplification was not observed (1.13); however, HER-2 protein overexpression was observed when assessed by image analysis (0.62). **(B) Case 7:** HER-2 gene amplification was not observed (1.33); however, HER-2 protein overexpression was observed when assessed by image analysis (0.62). Both cases were considered weak (1+) when immunohistochemistry was assessed by human analysis.

Case 8:

In addition to whole section biopsy analysis, areas of invasive tumour within the biopsies were isolated and immunohistochemical staining was assessed using image analysis. Figure 6.4-7 illustrates a whole section biopsy immunohistochemically stained for HER-2; areas of invasive tumour within the biopsy and are highlighted in red. FISH results performed across the whole biopsy recorded HER-2 gene amplification was not present (1.45); however, assessment of immunohistochemical staining by image analysis (0.76) observed protein overexpression. When image analysis was repeated on areas of invasive tumour only within the biopsy, protein overexpression was not recorded (0.50). Therefore, invasive tumour region selection agreed with FISH assessment.
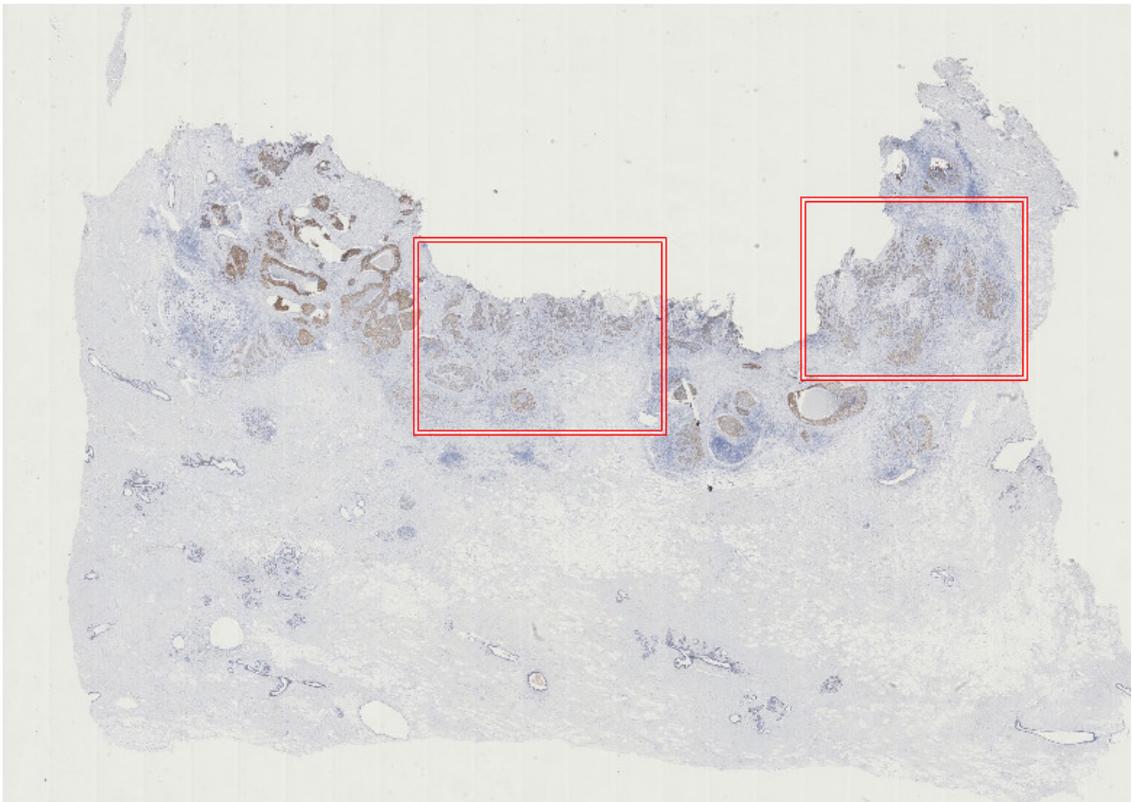
Figure 6.4-7: Whole section biopsy with no HER-2 gene amplification (1.45), and protein overexpression (0.76). However, when areas of invasive tumour only were assessed (red boxes), protein overexpression was not observed (0.50). HER-2 protein overexpression was assessed by immunohistochemistry and was quantified by image analysis. HER-2 gene amplification was assessed by FISH and quantified by one individual.

Agreement between HER-2 status within whole section biopsies and regions of invasive tumours was extremely high when assessed by image analysis ($r^2 = 0.942$). HER-2 status differed for only one case (Case 8 above) when whole biopsies analysis was compared with regions of invasive tumour only, as illustrated in Figure 6.4-8.
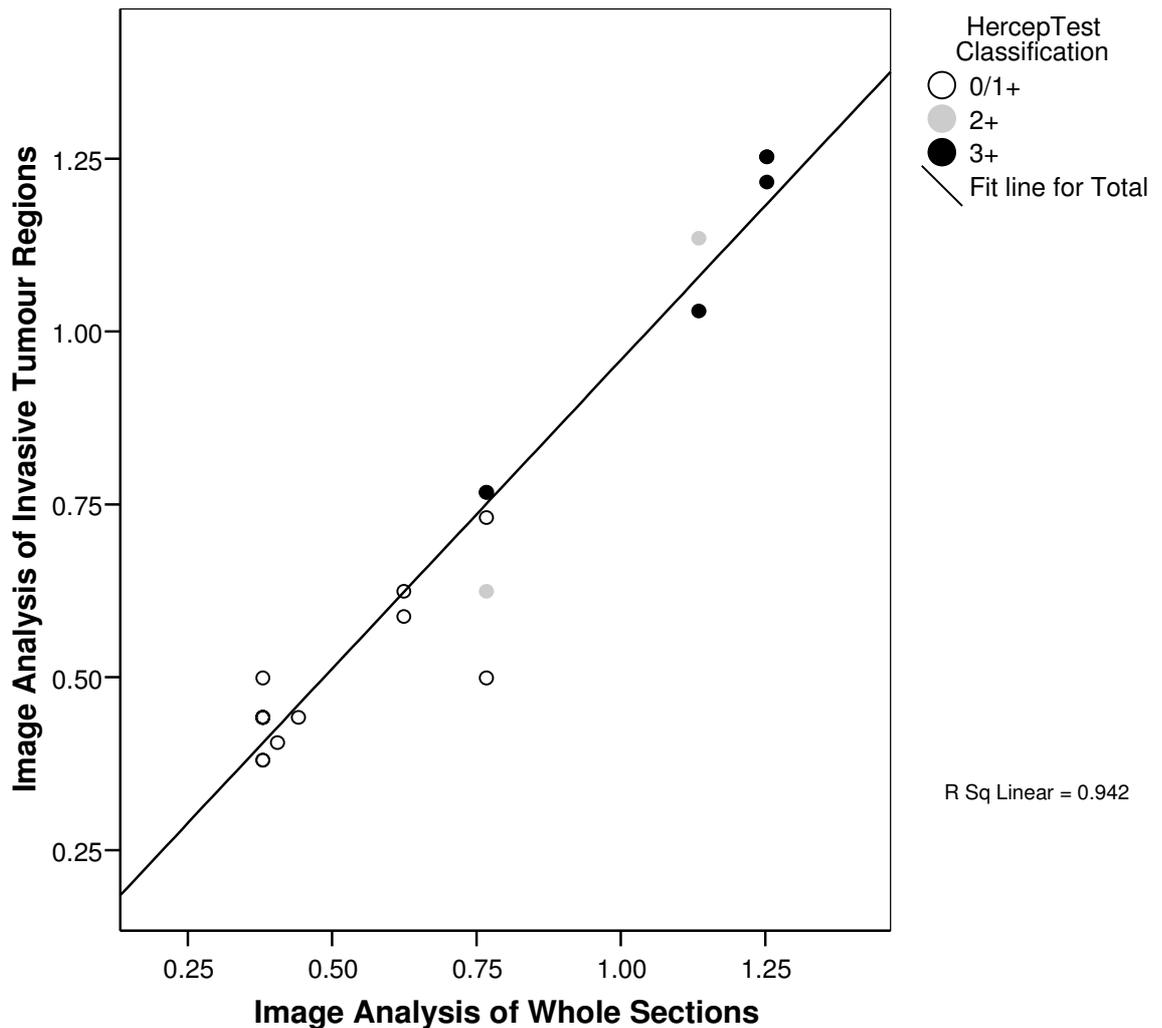
Figure 6.4-8: Correlation between image analysis assessment of HER-2 status within areas of invasive tumours only and across the entire whole section (n = 46). Image analysis classified immunohistochemical HER-2 protein expression as: < 0.54 non-overexpressed and ≥ 0.54 overexpressed. Data points have been colour coded according to the human HercepTest™ review.

Table 6.4-2 illustrates the level of agreement between HER-2 gene amplification and HER-2 protein overexpression across 46 whole sections. Both under and over treatment was observed when HER-2 status was classified using immunohistochemistry across 46 whole section biopsies. A considerably smaller number of patients were over treated when immunohistochemical staining was assessed by human (2.2%, n = 1) rather than image analysis (13.0%, n = 6). Compared to over treatment, a slightly higher percentage of patients were under treated, when assessed by human (8.7%, n = 4) and image analysis (4.3%, n = 2). The number of correctly treated cases was slightly higher when immunohistochemistry was assessed by human analysis (89.1%, n = 41) compared to image analysis (82.6%, n = 38). However, the agreement observed during the human analysis review was greatly inflated due to the number of referred cases (19.6%, n = 9).

On average 10.6% (Tissue Microarrays, n = 5) and 19.6% (Whole section biopsies, n = 9) of all immunohistochemically stained cases within this study were considered borderline overexpressed (2+) when assessed by human analysis. Therefore, FISH testing was required to definitively assess the HER-2 status within these cases. The majority of borderline cases (2+) did not have HER-2 gene amplification when assessed by FISH (Tissue Microarrays n = 3; Biopsies n = 5). Across all borderline (2+) cases within this study (n = 14), FISH and image analysis assessments differed in three cases.

Table 6.4-2: Agreement between HER-2 gene amplification (FISH) and protein overexpression (immunohistochemistry) when assessed by human and image analysis across whole section biopsies. Over treated occurred when protein overexpression was observed and gene amplification was not present. Correct treatment occurred when HER-2 status assessed by immunohistochemistry and FISH agreed. Under treated occurred when gene amplification was observed but protein overexpression was not present. Referred cases occurred when borderline immunohistochemical staining (2+) was observed. Referred cases were categorised under Correct Treatment.

Biopsies	Over Treated		Correct Treatment			Under Treatment	
	%	n	*%*	n	*Referred Cases*	*%*	n
Human Review	2.2	(1/46)	89.1	(41/46)	*9 of 41 (21.9%)*	8.7	(4/46)
Image Analysis	13.0	(6/46)	82.6	(38/46)	*N/A*	4.3	(2/46)

6.5 Discussion and Conclusions

HER-2 protein overexpression and gene amplification has been found to contribute to oncogenic transformation, and is associated with shorter disease free and overall survival in breast cancer. Therefore, accurate classification of HER-2 status within invasive breast carcinoma is crucial. Currently, the most common assay for evaluating HER-2 status is immunohistochemistry, with definitive FISH testing performed on borderline cases.

Results illustrate image analysis is at least as accurate as human interpretation of immunohistochemistry when classifying HER-2 status. Image analysis classified staining intensity on a two-point scale, whereas human analysis utilised a three-point scale. Referred cases (2+) within this study were grouped with the correct treatment results. The utilisation of an intermediate staining intensity category (2+) clearly inflated the level of agreement observed between HER-2 gene amplification and HER-2 protein overexpression when assessed by human analysis. However, disagreement between FISH and immunohistochemistry when assessed by human and image analysis were comparable.

When assessing HER-2 status within whole section biopsies, there was a high correlation between FISH testing and image analysis. Disagreement between methods occurred in 17.4% of cases (n = 8). Disagreement was attributed to results bordering the cut-points used in FISH classification (n = 1); false negative or positive HercepTest™ results (n = 2); non-invasive tumour regions influencing staining intensity values (n = 1), and image analysis recording protein overexpression as 0.62 (n = 4).

The HercepTest™ has been shown to give false negative results in up to 28% of HER-2 gene amplified cases (Ellis et al., 2005). False negative HercepTest™ occurred when gene amplification was present; but immunohistochemical staining was negative when assessed by image and human analysis. It has been suggested that this may be due to the destruction of the HER-2 epitope (Ellis et al., 2005). False positive HercepTest™ results

have been reported in up to 12% of cases (Bartlett et al., 2001, Hammock et al., 2003). False positive HercepTest™ occurred when gene amplification was not detected but immunohistochemical staining was very strong when assessed by human and image analysis. It is attributed to the overexpression of a single copy of the HER-2 gene. However, due to the heterogeneous nature of tumours in these instances it is possible that FISH analysis was performed on areas within the whole sections where no gene amplification or protein overexpression was observed.

Areas of invasive tumours were marked and assessed by image analysis. Image analysis results from invasive tumour areas and whole sections differed in one case. In this case, protein overexpression was observed when the whole biopsy was assessed; however, protein overexpression was not present when invasive tumour areas within the biopsy were examined. HER-2 gene amplification was not present for this one case. Therefore, this cohort of patients would suggest that pre-selecting areas of interest is of no real additional benefit with regard to patients' prognosis. However, this should be further validated for a definitive conclusion.

Only 14.9% (n = 7) of the Tissue Microarray dataset observed HER-2 protein overexpression (≥ 0.54). Also during Tissue Microarray analysis there were no cases with an image analysis result within the range of 0.60-0.76. Therefore, it is possible that within the range of 0.60-0.76 values a more optimum cut-point exists; however, due to the limited number of cases with HER-2 protein overexpression, it is not possible to verify this. Ideally, in future, a more comprehensive training set with a full range of image analysis results should be used to create optimal cut-points.

The algorithm recreated in ImageJ was utilised to quantify HER-2 protein expression within this Chapter. The algorithm created with ImageJ utilised the same approach as the Image-Pro Plus® algorithm. However, the implementation was slightly different and as a result, a degree of noise was present in the final image. Therefore, correction of noise signal was required during statistical analysis of the data. In addition, the procedure of

staining edges of tumours with tissue dye post-surgical extraction to identify tumour boundaries was found to influence image analysis assessment. These artifacts were not observed during Tissue Microarray analysis, and therefore were not a concern in previous Chapters. However, to avoid possible false positive results, black/brown tissue dye should be replaced with other colours, when image analysis assessment of immunohistochemistry is intended. The algorithm was used to assess whole section biopsies, therefore providing an automated and high-throughout analysis method. As a result, the application of the algorithm in a clinical context is possible.

The application of image analysis to quantify HER-2 protein overexpression has been assessed within this study. Image analysis proved to be as accurate as human analysis when assessing HER-2 protein overexpression. Correlation between image analysis and quantitative FISH testing was high. In addition, the application of image analysis assessment of whole section biopsies was demonstrated. Results suggest image analysis could increase throughput and precision and decrease subjectivity when assessing HER-2 protein overexpression.

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

7.1 Discussion and Conclusion

Human analysis of immunohistochemical staining intensity was proven to be inherently flawed, during the variety of assessments performed as part of this body of work. Levels of intra- and inter-observer agreements were found to be extremely low, in particular, when assessing membrane-bound E-cadherin immunohistochemical staining intensity (32.4% and 25.3% agreement, respectively). The level of expertise of the reviewers performing the analysis was found to hugely affect inter-observer agreement when assessing membrane-bound immunohistochemical staining. As expected, the quality of results was greatly increased when assessments were performed by a Professor of Pathology rather than a research scientist. Also, using reference tissue which illustrated a range of staining intensities during immunohistochemical analysis increased inter-observer agreements. The number of classifiers used to quantify membrane-bound immunohistochemical staining did not affect the quality of results when assessed by a Professor of Pathology. Traditionally, at least three classifiers are used to quantify staining intensity; however using two classifiers was proven to be as accurate.

There are a number of factors that affect human interpretation of immunohistochemical staining intensity which are inherent to traditional microscope-based reviews. The human eye is extremely sensitive to contrasts in colour. When interpreting staining intensity, the extent of cytoplasmic staining can hugely affect the perception of the colour within the membrane. In addition, adjustments in lighting settings can introduce huge variability in the reviewing process. Consistency of microscope configuration between slides and reviews is extremely difficult during microscope-based assessments. Also the analysis of Tissue Microarrays under a microscope can lead to misplaced spot orientation. Accurate manual tracking of the Tissue Microarray cores during microscope-based reviews is challenging and prone to error.

The Virtual Tissue Matrix (VTM) was a software application created to facilitate review and scoring of Tissue Microarrays via the Internet. The VTM alleviated variance of background lighting and disorientation within the Tissue Microarray slide. However,

inter- and intra-observer was still present when immunohistochemistry was assessed using the VTM. Ideally the assessments undergone within the VTM should have been performed by two experienced pathologist rather than comparing inter- and intra-observer agreements between a research scientist and a pathologist. However, results illustrated that inter- and intra-observer agreements were particularly low when assessing membrane-bound immunohistochemical staining intensity over other parameters assessed. Therefore, it was concluded that computer-assessed methods maybe of benefit when assessing this parameter.

For this reason, an image analysis algorithm was created to quantify membrane-bound immunohistochemical staining intensity. The algorithm was created using two different platforms (Image-Pro Plus® and ImageJ/Java). The algorithm proved simple to implement and facilitated effective analysis through integration into two different high-throughput workflows. However, the algorithm had a number of limitations. The algorithm did not distinguish between tumour and non-tumour regions nor could it discriminate between invasive or non-invasive tumour components. This was not a concern when Tissue Microarrays constructed from tumour regions only, in the case of the Bladder cancer arrays; however, when whole section analysis was required, for example in HER-2 analysis, this limitation was discovered. In addition, the percentage of positive membrane staining could be quantified more accurately. Currently, the number of positive pixels within the image is utilised; however, a value based on the percentage of positive staining rather than pixel counts should be utilised. The creation of the algorithm using the second development platform (ImageJ) introduced noise within the final image which introduced a need for additional downstream data analysis. The algorithm can be further improved to remove noise from the final image rather relying on statistical analysis of the data to remove the noise.

Cell Microarrays were constructed as part of this body of work. This enabled the Author to obtain invaluable practical laboratory experience in cell culture, tissue array design and construction within the National Institute for Cellular Biotechnology, Ireland and the National Cancer Institute, National Institute of Health, USA. Due to the absence of

comparative data, Cell Microarrays were assessed predominantly to validate the presence or absence of immunohistochemical staining for each of the biomarkers evaluated. Human and image analysis strongly agreed when qualifying staining presence or absence. The Author was dependant on retrospective data in relation to protein expression in these cell lines. Western blots were the predominant assay conducted on these cell lines. Greater use of ELISA in the original studies for protein quantification would have facilitated the correlation of image analysis intensity data with the extent of protein expression within the cell lines studied. In addition, protein expression was not evident in a large number of cases examined, and therefore, the level of agreement observed between methods was inflated by the number of negative cases present.

Image analysis of Tissue Microarrays composed of Urothelial Cell Carcinoma of the bladder (UCB) illustrated that the level of E-cadherin and ß-catenin membrane expression had a significant effect on prognosis. However, no such correlation was observed when the Tissue Microarrays were originally assessed by a research scientist during routine laboratory examination. Tissue Microarrays were constructed, immunohistochemically stained and reviewed prior to the creation of the image analysis algorithm as part of a wider study examining the effect of E-cadherin / ß-catenin complex on prognosis. When no correlation was observed between E-cadherin / ß-catenin protein expression and prognosis, Tissue Microarrays were assessed by image analysis. Once image analysis observed a significant correlation between expression and prognosis, the Tissue Microarrays were reassessed by a highly experienced Professor of Pathology resulting in the uncovering of a significant correlation between outcome and human review. That said, the cohort of low expression cases identified by the experienced reviewer was much smaller and statistically less significant than the cohort identified by image analysis.

However, as this was a retrospective study and all the materials existed prior to the Authors involvement, there were numerous difficulties encountered which could have been avoided if the study was designed optimally prior to Tissue Microarray construction. Firstly, the Tissue Microarrays were designed poorly. There was not

216

enough control standards present on each slide, which could be used to adequately calibrate the image analysis assessment. The Author had to make use of a single standard per slide, therefore forcing the assumption of an origin intercept and a linearity of relationship between signal and concentration. Cases were distributed across 26 Tissue Microarrays when a single Tissue Microarray would have sufficed which introduced unnecessary slide-to-slide variability. Finally, due to the limited number of samples assessed, the study lacked statistical power. As the study required both training and test sets to validate findings in the general population, a larger cohort of patients was required. This was also noted when assessing the correlation between E-cadherin expression and outcome with respect to tumour stage. Despite the limitations of this study, image analysis proved to be a better method for identifying the correlation between E-cadherin / ß-catenin protein expression and prognosis in UCB.

The HER-2 study allowed the Author better control of study parameters therefore allowing the application of training and test set principles lacking in the previous body of work. The quality of human review was considerably better when assessing HER-2 protein expression, compared to E-cadherin assessments. However, image analysis proved to be as effective as human analysis when assessing HER-2 status. Especially, considering all borderline cases (2+) observed during human analyses were interpreted as correctly assessed. If borderline cases were either excluded from the review or classified as 3+, image analysis would have surpassed human assessment in both cases. In addition, image analysis had the additional benefit of being high-throughput. However, the fact that regions of tumour / non-tumour and invasive / non-invasive were not distinguishable by the algorithm was clearly a limitation within this study.

In conclusion, an image analysis algorithm was created which compensates for inter-observer variability in human interpretation of immunohistochemistry. The algorithm whilst simplistic clearly has room for improvement has proven to perform beyond the limit of human capabilities in the quantification of membrane staining intensity. When used in conjunction with Tissue Microarray constructed solely from regions of tumour,

the algorithm offers an effective alternative to human assessment of immunohistochemical staining intensity.

7.2   Future Work

Moving forward the sophistication of the analysis performed by the algorithm could be greatly increased. In particular, the capabilities to identify regions of invasive / non invasive and tumour / non-tumour are of the utmost importance. The identification of tumour regions prior to quantification of staining intensity would result in reducing the time of analysis and increasing the accuracy of results obtained during tumour profiling. One proposed method for identifying tumour regions is to perform unsupervised clustering. There are numerous methods used to perform unsupervised clustering; however, in general the premise is to group entities with similar features together. For example, within tumour regions the nuclei are typically larger and more pleomorphic compared to normal tissue. Using unsupervised clustering, all similar shaped and size nuclei would be grouped together, therefore creating two distinct classes, normal and tumour regions. Ideally tissue types would be grouped by the algorithm, and an external review by a pathologist would be performed in order to classify the type of tissue within each grouping identified. Once a group has been classified by a pathologist, the criteria could be saved for future analysis. This method of tumour identification benefits from the experience of pathologist and the high-throughput nature of image analysis.

The assessments performed on the Cell Microarrays should be further developed by comparing ELISA results with the image analysis results for the complete dataset. In addition, a Cell Microarray block has been created which can provide numerous sections for further immunohistochemical analysis thereby facilitating the application of this array to the determination of expression of a variety of new biomarkers.

Using the UCB Tissue Microarrays as a training set, a larger cohort of samples should be assessed for E-cadherin protein expression. In addition, the Tissue Microarray blocks used within this study have been immunohistochemically stained with multiple markers. Membrane protein expression within these Tissue Microarrays should be assessed and correlated with prognosis. Only univariate analysis was performed within this body of

work. Multivariate analysis should be performed to examine the effect alternative parameters has on prognosis.

Once the algorithm has been updated to account for noise signal and tumour region identification, the tissue examined during the assessment of HER-2 protein expression, should be reassessed. In addition, a larger cohort of patients should be examined, with a wide range of FISH data available for correlation with image analysis results. Finally, the whole sections pertaining to the Tissue Microarrays used within this study should be assessed, to examine if the range of HER-2 protein expression is truly represented within the Tissue Microarrays.

BIBLIOGRAPHY

AL KURAYA, K., SIMON, R. & SAUTER, G. (2004) Tissue microarrays for high-throughput molecular pathology. *Ann Saudi Med,* 24**,** 169-74.

AMBUDKAR, S. V., DEY, S., HRYCYNA, C. A., RAMACHANDRA, M., PASTAN, I. & GOTTESMAN, M. M. (1999) Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu Rev Pharmacol Toxicol,* 39**,** 361-98.

ANDERSEN, C. L., HOSTETTER, G., GRIGORYAN, A., SAUTER, G. & KALLIONIEMI, A. (2001) Improved procedure for fluorescence in situ hybridization on tissue microarrays. *Cytometry,* 45**,** 83-6.

ANDRECHEK, E. R. & MULLER, W. J. (2000) Tyrosine kinase signalling in breast cancer: tyrosine kinase-mediated signal transduction in transgenic mouse models of human breast cancer. *Breast Cancer Res,* 2**,** 211-6.

BANI, M. R., NICOLETTI, M. I., ALKHAROUF, N. W., GHILARDI, C., PETERSEN, D., ERBA, E., SAUSVILLE, E. A., LIU, E. T. & GIAVAZZI, R. (2004) Gene expression correlating with response to paclitaxel in ovarian carcinoma xenografts. *Mol Cancer Ther,* 3**,** 111-21.

BARTLETT, J., MALLON, E. & COOKE, T. (2003) The clinical evaluation of HER-2 status: which test to use? *J Pathol,* 199**,** 411-7.

BARTLETT, J. M., GOING, J. J., MALLON, E. A., WATTERS, A. D., REEVES, J. R., STANTON, P., RICHMOND, J., DONALD, B., FERRIER, R. & COOKE, T. G. (2001) Evaluating HER2 amplification and overexpression in breast cancer. *J Pathol,* 195**,** 422-8.

BASELGA, J. & ARTEAGA, C. L. (2005) Critical update and emerging trends in epidermal growth factor receptor targeting in cancer. *J Clin Oncol,* 23**,** 2445-59.

BASELGA, J., TRIPATHY, D., MENDELSOHN, J., BAUGHMAN, S., BENZ, C. C., DANTIS, L., SKLARIN, N. T., SEIDMAN, A. D., HUDIS, C. A., MOORE, J., ROSEN, P. P., TWADDELL, T., HENDERSON, I. C. & NORTON, L. (1996) Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. *J Clin Oncol,* 14**,** 737-44.

BATTIFORA, H. (1986) The multitumor (sausage) tissue block: novel method for immunohistochemical antibody testing. *Lab Invest,* 55**,** 244-8.

BATTIFORA, H. & MEHTA, P. (1990) The checkerboard tissue block. An improved multitissue control block. *Lab Invest,* 63**,** 722-4.

BEHRENS, J., MAREEL, M. M., VAN ROY, F. M. & BIRCHMEIER, W. (1989) Dissecting tumor cell invasion: epithelial cells acquire invasive properties after the loss of uvomorulin-mediated cell-cell adhesion. *J Cell Biol,* 108**,** 2435-47.

BERGER, A. J., CAMP, R. L., DIVITO, K. A., KLUGER, H. M., HALABAN, R. & RIMM, D. L. (2004) Automated quantitative analysis of HDM2 expression in malignant melanoma shows association with early-stage disease and improved outcome. *Cancer Res,* 64**,** 8767-72.

BERTHEAU, P., CAZALS-HATEM, D., MEIGNIN, V., DE ROQUANCOURT, A., VEROLA, O., LESOURD, A., SENE, C., BROCHERIOU, C. & JANIN, A. (1998) Variability of immunohistochemical reactivity on stored paraffin slides. *J Clin Pathol,* 51**,** 370-4.

BORRESEN, A. L., OTTESTAD, L., GAUSTAD, A., ANDERSEN, T. I., HEIKKILA, R., JAHNSEN, T., TVEIT, K. M. & NESLAND, J. M. (1990) Amplification and protein over-expression of the neu/HER-2/c-erbB-2 protooncogene in human breast carcinomas: relationship to loss of gene sequences on chromosome 17, family history and prognosis. *Br J Cancer,* 62**,** 585-90.

BOVA, G. S., PARMIGIANI, G., EPSTEIN, J. I., WHEELER, T., MUCCI, N. R. & RUBIN, M. A. (2001) Web-based tissue microarray image data analysis: initial validation testing through prostate cancer Gleason grading. *Hum Pathol,* 32**,** 417-27.

BRAUNSCHWEIG, T., CHUNG, J. Y. & HEWITT, S. M. (2004) Perspectives in tissue microarrays. *Comb Chem High Throughput Screen,* 7**,** 575-85.

BRECELJ, E., FRKOVIC GRAZIO, S., AUERSPERG, M. & BRACKO, M. (2005) Prognostic value of E-cadherin expression in thyroid follicular carcinoma. *Eur J Surg Oncol,* 31**,** 544-8.

BREMNES, R. M., VEVE, R., GABRIELSON, E., HIRSCH, F. R., BARON, A., BEMIS, L., GEMMILL, R. M., DRABKIN, H. A. & FRANKLIN, W. A. (2002a) High-throughput tissue microarray analysis used to evaluate biology and prognostic significance of the E-cadherin pathway in non-small-cell lung cancer. *J Clin Oncol,* 20**,** 2417-28.

BREMNES, R. M., VEVE, R., HIRSCH, F. R. & FRANKLIN, W. A. (2002b) The E-cadherin cell-cell adhesion complex and lung cancer invasion, metastasis, and prognosis. *Lung Cancer,* 36**,** 115-24.

BRINGUIER, P. P., UMBAS, R., SCHAAFSMA, H. E., KARTHAUS, H. F., DEBRUYNE, F. M. & SCHALKEN, J. A. (1993) Decreased E-cadherin immunoreactivity correlates with poor survival in patients with bladder tumors. *Cancer Res,* 53**,** 3241-5.

BROOKS, S. (2005) Cell Imaging Techniques. *The Biomedical Scientist*.

BROWN, R. E., LUN, M., PRICHARD, J. W., BLASICK, T. M. & ZHANG, P. L. (2004) Morphoproteomic and pharmacoproteomic correlates in hormone-receptor-negative breast carcinoma cell lines. *Ann Clin Lab Sci,* 34**,** 251-62.

BUKHOLM, I. K., NESLAND, J. M., KARESEN, R., JACOBSEN, U. & BORRESEN-DALE, A. L. (1998) E-cadherin and alpha-, beta-, and gamma-catenin protein expression in relation to metastasis in human breast carcinoma. *J Pathol,* 185**,** 262-6.

BYRNE, R. R., SHARIAT, S. F., BROWN, R., KATTAN, M. W., MORTON, R. J., WHEELER, T. M. & LERNER, S. P. (2001) E-cadherin immunostaining of bladder transitional cell carcinoma, carcinoma in situ and lymph node metastases with long-term followup. *J Urol,* 165**,** 1473-9.

CAMOZZI, C. & RAZVI, E. (2004) Tissue: Microarrays: Facilitating Drug Research. *Genetic Engineering News,* 24**,** 30-39.

CAMP, R. L., CHUNG, G. G. & RIMM, D. L. (2002) Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med,* 8**,** 1323-7.

CAMP, R. L., DOLLED-FILHART, M., KING, B. L. & RIMM, D. L. (2003) Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. *Cancer Res,* 63**,** 1445-8.

CAMP, R. L., DOLLED-FILHART, M. & RIMM, D. L. (2004) X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res,* 10**,** 7252-9.

CAMP, R. L. D., K.A. (2005) Tissue Microarrays - automated analysis and future directions. *Breast Cancer Online,* 8.

CAMPOS, R. S., LOPES, A., GUIMARAES, G. C., CARVALHO, A. L. & SOARES, F. A. (2006) E-cadherin, MMP-2, and MMP-9 as prognostic markers in penile cancer: analysis of 125 patients. *Urology,* 67**,** 797-802.

CARTER, P., PRESTA, L., GORMAN, C. M., RIDGWAY, J. B., HENNER, D., WONG, W. L., ROWLAND, A. M., KOTTS, C., CARVER, M. E. & SHEPARD, H. M. (1992) Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc Natl Acad Sci U S A,* 89**,** 4285-9.

CHARPIN, C., DALES, J. P., GARCIA, S., CARPENTIER, S., DJEMLI, A., ANDRAC, L., LAVAUT, M. N., ALLASIA, C. & BONNIER, P. (2004) Tumor neoangiogenesis by CD31 and CD105 expression evaluation in breast carcinoma tissue microarrays. *Clin Cancer Res,* 10**,** 5815-9.

CHOW, V., YUEN, A. P., LAM, K. Y., TSAO, G. S., HO, W. K. & WEI, W. I. (2001) A comparative study of the clinicopathological significance of E-cadherin and catenins (alpha, beta, gamma) expression in the surgical management of oral tongue carcinoma. *J Cancer Res Clin Oncol,* 127**,** 59-63.

CIAMPA, A., XU, B., AYATA, G., BAIYEE, D., WALLACE, J., WERTHEIMER, M., EDMISTON, K. & KHAN, A. (2006) HER-2 status in breast cancer: correlation of gene amplification by FISH with immunohistochemistry expression using advanced cellular imaging system. *Appl Immunohistochem Mol Morphol,* 14**,** 132-7.

CLAIROTTE, A., LASCOMBE, I., FAUCONNET, S., MAUNY, F., FELIX, S., ALGROS, M. P., BITTARD, H. & KANTELIP, B. (2006) Expression of E-cadherin and alpha-, beta-, gamma-catenins in patients with bladder cancer: identification of gamma-catenin as a new prognostic marker of neoplastic progression in T1 superficial urothelial tumors. *Am J Clin Pathol,* 125**,** 119-26.

CLIMENT, M. A., SEGUI, M. A., PEIRO, G., MOLINA, R., LERMA, E., OJEDA, B., LOPEZ-LOPEZ, J. J. & ALONSO, C. (2001) Prognostic value of HER-2/neu and p53 expression in node-positive breast cancer. HER-2/neu effect on adjuvant tamoxifen treatment. *Breast,* 10**,** 67-77.

CONWAY, C. M., O'SHEA, D., O'BRIEN, S., LAWLER, D. K., DODRILL, G. D., O'GRADY, A., BARRETT, H., GULMANN, C., O'DRISCOLL, L., GALLAGHER, W. M., KAY, E. W. & O'SHEA, D. G. (2006) The development and validation of the Virtual Tissue Matrix, a software application that facilitates the review of tissue microarrays on line. *BMC Bioinformatics,* 7**,** 256.

COSTELLO, S. S., JOHNSTON, D. J., DERVAN, P. A. & O'SHEA, D. G. (2003) Development and evaluation of the virtual pathology slide: a new tool in telepathology. *J Med Internet Res,* 5**,** e11.

CREGGER, M., BERGER, A. J. & RIMM, D. L. (2006) Immunohistochemistry and quantitative analysis of protein expression. *Arch Pathol Lab Med,* 130**,** 1026-30.

DANIELE, L., MACRI, L., SCHENA, M., DONGIOVANNI, D., BONELLO, L., ARMANDO, E., CIUFFREDA, L., BERTETTO, O., BUSSOLATI, G. & SAPINO, A. (2007) Predicting gefitinib responsiveness in lung cancer by

fluorescence in situ hybridization/chromogenic in situ hybridization analysis of EGFR and HER2 in biopsy and cytology specimens. *Mol Cancer Ther,* 6**,** 1223-9.

DELLA MEA, V., VIEL, F. & BELTRAMI, C. A. (2005) A pixel-based autofocusing technique for digital histologic and cytologic slides. *Comput Med Imaging Graph,* 29**,** 333-41.

DIVITO, K. A., BERGER, A. J., CAMP, R. L., DOLLED-FILHART, M., RIMM, D. L. & KLUGER, H. M. (2004) Automated quantitative analysis of tissue microarrays reveals an association between high Bcl-2 expression and improved outcome in melanoma. *Cancer Res,* 64**,** 8773-7.

DOLLED-FILHART, M., MCCABE, A., GILTNANE, J., CREGGER, M., CAMP, R. L. & RIMM, D. L. (2006) Quantitative in situ analysis of beta-catenin expression in breast cancer shows decreased expression is associated with poor outcome. *Cancer Res,* 66**,** 5487-94.

ELLIS, C. M., DYSON, M. J., STEPHENSON, T. J. & MALTBY, E. L. (2005) HER2 amplification status in breast cancer: a comparison between immunohistochemical staining and fluorescence in situ hybridisation using manual and automated quantitative image analysis scoring techniques. *J Clin Pathol,* 58**,** 710-4.

ELLIS, I. O., BARTLETT, J., DOWSETT, M., HUMPHREYS, S., JASANI, B., MILLER, K., PINDER, S. E., RHODES, A. & WALKER, R. (2004) Best Practice No 176: Updated recommendations for HER2 testing in the UK. *J Clin Pathol,* 57**,** 233-7.

ELLIS, I. O., DOWSETT, M., BARTLETT, J., WALKER, R., COOKE, T., GULLICK, W., GUSTERSON, B., MALLON, E. & LEE, P. B. (2000) Recommendations for HER2 testing in the UK. *J Clin Pathol,* 53**,** 890-2.

FAITH, D. A., ISAACS, W. B., MORGAN, J. D., FEDOR, H. L., HICKS, J. L., MANGOLD, L. A., WALSH, P. C., PARTIN, A. W., PLATZ, E. A., LUO, J. & DE MARZO, A. M. (2004) Trefoil factor 3 overexpression in prostatic carcinoma: prognostic importance using tissue microarrays. *Prostate,* 61**,** 215-27.

FEJZO, M. S. & SLAMON, D. J. (2001) Frozen tumor tissue microarray technology for analysis of tumor RNA, DNA, and proteins. *Am J Pathol,* 159**,** 1645-50.

FERGENBAUM, J. H., GARCIA-CLOSAS, M., HEWITT, S. M., LISSOWSKA, J., SAKODA, L. C. & SHERMAN, M. E. (2004) Loss of antigenicity in stored sections of breast cancer tissue microarrays. *Cancer Epidemiol Biomarkers Prev,* 13**,** 667-72.

FOX, C. H., JOHNSON, F. B., WHITING, J. & ROLLER, P. P. (1985) Formaldehyde fixation. *J Histochem Cytochem,* 33**,** 845-53.

FRIXEN, U. H., BEHRENS, J., SACHS, M., EBERLE, G., VOSS, B., WARDA, A., LOCHNER, D. & BIRCHMEIER, W. (1991) E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J Cell Biol,* 113**,** 173-85.

GAMBOA-DOMINGUEZ, A., DOMINGUEZ-FONSECA, C., CHAVARRI-GUERRA, Y., VARGAS, R., REYES-GUTIERREZ, E., GREEN, D., QUINTANILLA-MARTINEZ, L., LUBER, B., BUSCH, R., BECKER, K. F., BECKER, I., HOFLER, H. & FEND, F. (2005) E-cadherin expression in sporadic gastric

cancer from Mexico: exon 8 and 9 deletions are infrequent events associated with poor survival. *Hum Pathol,* 36**,** 29-35.

GLINSKY, G. V., KRONES-HERZIG, A., GLINSKII, A. B. & GEBAUER, G. (2003) Microarray analysis of xenograft-derived cancer cell lines representing multiple experimental models of human prostate cancer. *Mol Carcinog,* 37**,** 209-21.

GLOECKLER RIES, L. A., REICHMAN, M. E., LEWIS, D. R., HANKEY, B. F. & EDWARDS, B. K. (2003) Cancer survival and incidence from the Surveillance, Epidemiology, and End Results (SEER) program. *Oncologist,* 8**,** 541-52.

GOMEZ, D. D., CARSTENSEN, J. M. & ERSBOLL, B. K. (2006) Collecting highly reproducible images to support dermatological medical diagnosis. *Image and Vision Computing,* 24**,** 186-191.

GONTERO, P., BANISADR, S., FREA, B. & BRAUSI, M. (2004) Metastasis markers in bladder cancer: a review of the literature and clinical considerations. *Eur Urol,* 46**,** 296-311.

GRENIER, J. & SORIA, J. C. (2007) [Inhibitors of the EGFR pathway in non small cell lung cancer: what's new in 2007?]. *Bull Cancer,* 94**,** 53-61.

GUESSOUS, I., CORNUZ, J. & PACCAUD, F. (2007) Lung cancer screening: current situation and perspective. *Swiss Med Wkly,* 137**,** 304-11.

GULMANN, C., LORING, P., O'GRADY, A. & KAY, E. (2004) Miniature tissue microarrays for HercepTest standardisation and analysis. *J Clin Pathol,* 57**,** 1229-31.

HAMMOCK, L., LEWIS, M., PHILLIPS, C. & COHEN, C. (2003) Strong HER-2/neu protein overexpression by immunohistochemistry often does not predict oncogene amplification by fluorescence in situ hybridization. *Hum Pathol,* 34**,** 1043-7.

HANSEN, W. H., GILMAN, G., FINNESGARD, S. J., WELLIK, T. J., NELSON, T. A., JOHNSON, M. F., SCHWENK, N. M., SEWARD, J. B. & KHANDHERIA, B. K. (2004) The transition from an analog to a digital echocardiography laboratory: The Mayo experience. *Journal of the American Society of Echocardiography,* 17**,** 1214-1224.

HARIGOPAL, M., BERGER, A. J., CAMP, R. L., RIMM, D. L. & KLUGER, H. M. (2005) Automated quantitative analysis of E-cadherin expression in lymph node metastases is predictive of survival in invasive ductal breast cancer. *Clin Cancer Res,* 11**,** 4083-9.

HEWITT, S. M. (2004) *Methods in molecular biology, chapter Design, Construction and user of tissue microarrays*, Humana Press Inc. Totowa NJ.

HEWITT, S. M. (2006) The application of tissue microarrays in the validation of microarray results. *Methods Enzymol,* 410**,** 400-15.

HEWITT, S. M., DEAR, J. & STAR, R. A. (2004) Discovery of protein biomarkers for renal diseases. *J Am Soc Nephrol,* 15**,** 1677-89.

HICKS, D. G. & TUBBS, R. R. (2005) Assessment of the HER2 status in breast cancer by fluorescence in situ hybridization: a technical review with interpretive guidelines. *Hum Pathol,* 36**,** 250-61.

HIDALGO, A., PINA, P., GUERRERO, G., LAZOS, M. & SALCEDO, M. (2003) A simple method for the construction of small format tissue arrays. *J Clin Pathol,* 56**,** 144-6.

HO, J., JUKIC, D., PARWANI, A., ANTHONY, L. & GILBERTSON, J. (2006) Use of whole slide imaging for primary diagnosis in surgical pathology: An intrainstitutional validation study. *Modern Pathology,* 19**,** 329A-329A.

HOANG, M. P., SAHIN, A. A., ORDONEZ, N. G. & SNEIGE, N. (2000) HER-2/neu gene amplification compared with HER-2/neu protein overexpression and interobserver reproducibility in invasive breast carcinoma. *Am J Clin Pathol,* 113**,** 852-9.

HOOS, A. & CORDON-CARDO, C. (2001) Tissue microarray profiling of cancer specimens and cell lines: opportunities and limitations. *Lab Invest,* 81**,** 1331-8.

HOOS, A., URIST, M. J., STOJADINOVIC, A., MASTORIDES, S., DUDAS, M. E., LEUNG, D. H., KUO, D., BRENNAN, M. F., LEWIS, J. J. & CORDON-CARDO, C. (2001) Validation of tissue microarrays for immunohistochemical profiling of cancer specimens using the example of human fibroblastic tumors. *Am J Pathol,* 158**,** 1245-51.

HSI, E. D. & TUBBS, R. R. (2004) Guidelines for HER2 testing in the UK. *J Clin Pathol,* 57**,** 241-2.

HSU, S. M., RAINE, L. & FANGER, H. (1981) Use of avidin-biotin-peroxidase complex (ABC) in immunoperoxidase techniques: a comparison between ABC and unlabeled antibody (PAP) procedures. *J Histochem Cytochem,* 29**,** 577-80.

JACOBS, T. W., PRIOLEAU, J. E., STILLMAN, I. E. & SCHNITT, S. J. (1996) Loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. *J Natl Cancer Inst,* 88**,** 1054-9.

JEMAL, A., THOMAS, A., MURRAY, T. & THUN, M. (2002) Cancer statistics, 2002. *CA Cancer J Clin,* 52**,** 23-47.

JESCHKE, U., MYLONAS, I., KUHN, C., SHABANI, N., KUNERT-KEIL, C., SCHINDLBECK, C., GERBER, B. & FRIESE, K. (2007) Expression of E-cadherin in human ductal breast cancer carcinoma in situ, invasive carcinomas, their lymph node metastases, their distant metastases, carcinomas with recurrence and in recurrence. *Anticancer Res,* 27**,** 1969-74.

JOHANSSON, A. C., VISSE, E., WIDEGREN, B., SJOGREN, H. O. & SIESJO, P. (2001) Computerized image analysis as a tool to quantify infiltrating leukocytes: a comparison between high- and low-magnification images. *J Histochem Cytochem,* 49**,** 1073-79.

JOHN MORAN EYE CENTRE UNIVERSITY OF UTAH (2005) Visual Acuity.

JOHNSTON, D. (2005a) The VPS ReplaySuite: Development ans Evaluation of a novel internet based telepathology tool. *School of Biotechnology.* Dublin, Dublin City University.

JOHNSTON, D. J. (2005b) The VPS ReplaySuite Development and evaluation of a novel, internet based telepathology tool. *Medical Informatics Group.* Dublin, Dublin City University.

JOHNSTON, D. J., COSTELLO, S. P., DERVAN, P. A. & O'SHEA, D. G. (2005) Development and preliminary evaluation of the VPS ReplaySuite: a virtual double-headed microscope for pathology. *BMC Med Inform Decis Mak,* 5**,** 10.

JOSHI, A. S., SHARANGPANI, G. M., PORTER, K., KEYHANI, S., MORRISON, C., BASU, A. S., GHOLAP, G. A., GHOLAP, A. S. & BARSKY, S. H. (2007) Semi-automated imaging system to quantitate Her-2/neu membrane receptor immunoreactivity in human breast cancer. *Cytometry A,* 71**,** 273-85.

KASE, S., SUGIO, K., YAMAZAKI, K., OKAMOTO, T., YANO, T. & SUGIMACHI, K. (2000) Expression of E-cadherin and beta-catenin in human non-small cell lung cancer and the clinical significance. *Clin Cancer Res,* 6**,** 4789-96.

KASHIBUCHI, K., TOMITA, K., SCHALKEN, J. A., KUME, H., TAKEUCHI, T. & KITAMURA, T. (2007) The prognostic value of E-cadherin, alpha-, beta- and gamma-catenin in bladder cancer patients who underwent radical cystectomy. *Int J Urol,* 14**,** 789-94.

KAY, E., O'GRADY, A., MORGAN, J. M., WOZNIAK, S. & JASANI, B. (2004) Use of tissue microarray for interlaboratory validation of HER2 immunocytochemical and FISH testing. *J Clin Pathol,* 57**,** 1140-4.

KAY, E. W., WALSH, C. J., CASSIDY, M., CURRAN, B. & LEADER, M. (1994) C-erbB-2 immunostaining: problems with interpretation. *J Clin Pathol,* 47**,** 816-22.

KLIJN, J. G., BERNS, P. M., SCHMITZ, P. I. & FOEKENS, J. A. (1992) The clinical significance of epidermal growth factor receptor (EGF-R) in human breast cancer: a review on 5232 patients. *Endocr Rev,* 13**,** 3-17.

KONECNY, G. E., PEGRAM, M. D., VENKATESAN, N., FINN, R., YANG, G., RAHMEH, M., UNTCH, M., RUSNAK, D. W., SPEHAR, G., MULLIN, R. J., KEITH, B. R., GILMER, T. M., BERGER, M., PODRATZ, K. C. & SLAMON, D. J. (2006) Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res,* 66**,** 1630-9.

KONONEN, J., BUBENDORF, L., KALLIONIEMI, A., BARLUND, M., SCHRAML, P., LEIGHTON, S., TORHORST, J., MIHATSCH, M. J., SAUTER, G. & KALLIONIEMI, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med,* 4**,** 844-7.

KROGER, N., ACHTERRATH, W., HEGEWISCH-BECKER, S., MROSS, K. & ZANDER, A. R. (1999) Current options in treatment of anthracycline-resistant breast cancer. *Cancer Treat Rev,* 25**,** 279-91.

KURDZIEL, K. A., KALEN, J. D., HIRSCH, J. I., WILSON, J. D., AGARWAL, R., BARRETT, D., BEAR, H. D. & MCCUMISKEY, J. F. (2007) Imaging multidrug resistance with 4-[(18)F]fluoropaclitaxel. *Nucl Med Biol,* 34**,** 823-31.

LACROIX-TRIKI, M., MATHOULIN-PELISSIER, S., GHNASSIA, J. P., MACGROGAN, G., VINCENT-SALOMON, A., BROUSTE, V., MATHIEU, M. C., ROGER, P., BIBEAU, F., JACQUEMIER, J., PENAULT-LLORCA, F. & ARNOULD, L. (2006) High inter-observer agreement in immunohistochemical evaluation of HER-2/neu expression in breast cancer: a multicentre GEFPICS study. *Eur J Cancer,* 42**,** 2946-53.

LANDIS, J. R. & KOCH, G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics,* 33**,** 159-74.

LARKIN, A., MORAN, E., ALEXANDER, D., DOHERTY, G., CONNOLLY, L., KENNEDY, S. M. & CLYNES, M. (1999) A new monoclonal antibody that specifically recognises the MDR-3-encoded gene product. *Int J Cancer,* 80**,** 265-71.

LARKIN, A., O'DRISCOLL, L., KENNEDY, S., PURCELL, R., MORAN, E., CROWN, J., PARKINSON, M. & CLYNES, M. (2004) Investigation of MRP-1 protein and MDR-1 P-glycoprotein expression in invasive breast cancer: a prognostic study. *Int J Cancer,* 112**,** 286-94.

LEONESSA, F. & CLARKE, R. (2003) ATP binding cassette transporters and drug resistance in breast cancer. *Endocr Relat Cancer,* 10**,** 43-73.

LI, R., NI, J., BOURNE, P. A., YEH, S., YAO, J., DI SANT'AGNESE, P. A. & HUANG, J. (2005) Cell culture block array for immunocytochemical study of protein expression in cultured cells. *Appl Immunohistochem Mol Morphol,* 13**,** 85-90.

LIANG, Y., O'DRISCOLL, L., MCDONNELL, S., DOOLAN, P., OGLESBY, I., DUFFY, K., O'CONNOR, R. & CLYNES, M. (2004) Enhanced in vitro invasiveness and drug resistance with altered gene expression patterns in a human lung carcinoma cell line after pulse selection with anticancer drugs. *Int J Cancer,* 111**,** 484-93.

LIU, C. L., PRAPONG, W., NATKUNAM, Y., ALIZADEH, A., MONTGOMERY, K., GILKS, C. B. & VAN DE RIJN, M. (2002) Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. *Am J Pathol,* 161**,** 1557-65.

LUFTNER, D., HENSCHKE, P., KAFKA, A., ANAGNOSTOPOULOS, I., WIECHEN, K., GEPPERT, R., STEIN, H., WERNECKE, K. D., KREIENBERG, R. & POSSINGER, K. (2004) Discordant results obtained for different methods of HER-2/neu testing in breast cancer--a question of standardization, automation and timing. *Int J Biol Markers,* 19**,** 1-13.

MACBEATH, G. (2002) Protein microarrays and proteomics. *Nat Genet,* 32 Suppl**,** 526-32.

MAHNKEN, A., KAUSCH, I., FELLER, A. C. & KRUGER, S. (2005) E-cadherin immunoreactivity correlates with recurrence and progression of minimally invasive transitional cell carcinomas of the urinary bladder. *Oncol Rep,* 14**,** 1065-70.

MANLEY, S., MUCCI, N. R., DE MARZO, A. M. & RUBIN, M. A. (2001) Relational database structure to manage high-density tissue microarray data and images for pathology studies focusing on clinical outcome: the prostate specialized program of research excellence model. *Am J Pathol,* 159**,** 837-43.

MATTHEWS, C., CATHERWOOD, M. A., LARKIN, A. M., CLYNES, M., MORRIS, T. C. & ALEXANDER, H. D. (2006) MDR-1, but not MDR-3 gene expression, is associated with unmutated IgVH genes and poor prognosis chromosomal aberrations in chronic lymphocytic leukemia. *Leuk Lymphoma,* 47**,** 2308-13.

MCCABE, A., DOLLED-FILHART, M., CAMP, R. L. & RIMM, D. L. (2005) Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *J Natl Cancer Inst,* 97**,** 1808-15.

MCCULLOUGH, B., YING, X., MONTICELLO, T. & BONNEFOI, M. (2004) Digital microscopy imaging and new approaches in toxicologic pathology. *Toxicol Pathol,* 32 Suppl 2**,** 49-58.

MCKAY, J. S., BIGLEY, A., BELL, A., JENKINS, R., SOMERS, R., BROCKLEHURST, S., WHITE, A. & GOODWIN, L. (2006) A pilot evaluation of the use of tissue microarrays for quantitation of target distribution in drug discovery pathology. *Exp Toxicol Pathol,* 57**,** 181-93.

MENDELSOHN, J. & BASELGA, J. (2003) Status of epidermal growth factor receptor antagonists in the biology and treatment of cancer. *J Clin Oncol,* 21**,** 2787-99.

MESSERSMITH, W., OPPENHEIMER, D., PERALBA, J., SEBASTIANI, V., AMADOR, M., JIMENO, A., EMBUSCADO, E., HIDALGO, M. & IACOBUZIO-DONAHUE, C. (2005) Assessment of Epidermal Growth Factor Receptor (EGFR) signaling in paired colorectal cancer and normal colon tissue samples using computer-aided immunohistochemical analysis. *Cancer Biol Ther,* 4**,** 1381-6.

MICROGRAPHIA (2003) The Microscope Lamp.

MILANES-YEARSLEY, M., HAMMOND, M. E., PAJAK, T. F., COOPER, J. S., CHANG, C., GRIFFIN, T., NELSON, D., LARAMORE, G. & PILEPICH, M. (2002) Tissue micro-array: a cost and time-effective method for correlative studies by regional and national cancer study groups. *Mod Pathol,* 15**,** 1366-73.

MOCH, H., KONONEN, T., KALLIONIEMI, O. P. & SAUTER, G. (2001) Tissue microarrays: what will they bring to molecular and anatomic pathology? *Adv Anat Pathol,* 8**,** 14-20.

MORAN, E., CLEARY, I., LARKIN, A. M., AMHLAOIBH, R. N., MASTERSON, A., SCHEPER, R. J., IZQUIERDO, M. A., CENTER, M., O'SULLIVAN, F. & CLYNES, M. (1997a) Co-expression of MDR-associated markers, including P-170, MRP and LRP and cytoskeletal proteins, in three resistant variants of the human ovarian carcinoma cell line, OAW42. *Eur J Cancer,* 33**,** 652-60.

MORAN, E., LARKIN, A., DOHERTY, G., KELEHAN, P., KENNEDY, S. & CLYNES, M. (1997b) A new mdr-1 encoded P-170 specific monoclonal antibody: (6/1C) on paraffin wax embedded tissue without pretreatment of sections. *J Clin Pathol,* 50**,** 465-71.

MORITA, S. Y., KOBAYASHI, A., TAKANEZAWA, Y., KIOKA, N., HANDA, T., ARAI, H., MATSUO, M. & UEDA, K. (2007) Bile salt-dependent efflux of cellular phospholipids mediated by ATP binding cassette protein B4. *Hepatology,* 46**,** 188-99.

MOYANO CALVO, J. L., BLANCO PALENCIANO, E., BEATO MORENO, A., GUTIERREZ GONZALEZ, M., PEREZ-LANZAC LORCA, A., SAMANIEGO TORRES, A., MONTANO, J. A. & FERNANDEZ CASTINEIRAS, J. (2006) [Prognostic value of E-cadherina, beta catenin, Ki-67 antigen and p53 protein in the superficial bladder tumors]. *Actas Urol Esp,* 30**,** 871-8.

MURTA-NASCIMENTO, C., SCHMITZ-DRAGER, B. J., ZEEGERS, M. P., STEINECK, G., KOGEVINAS, M., REAL, F. X. & MALATS, N. (2007) Epidemiology of urinary bladder cancer: from tumor development to patient's death. *World J Urol,* 25**,** 285-95.

MUSS, H. B., THOR, A. D., BERRY, D. A., KUTE, T., LIU, E. T., KOERNER, F., CIRRINCIONE, C. T., BUDMAN, D. R., WOOD, W. C., BARCOS, M. & ET AL. (1994) c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. *N Engl J Med,* 330**,** 1260-6.

NAKOPOULOU, L., ZERVAS, A., GAKIOPOULOU-GIVALOU, H., CONSTANTINIDES, C., DOUMANIS, G., DAVARIS, P. & DIMOPOULOS, C. (2000) Prognostic value of E-cadherin, beta-catenin, P120ctn in patients with transitional cell bladder cancer. *Anticancer Res,* 20**,** 4571-8.

NATHKE, I. S., HINCK, L. E. & NELSON, W. J. (1993) Epithelial cell adhesion and development of cell surface polarity: possible mechanisms for modulation of cadherin function, organization and distribution. *J Cell Sci Suppl,* 17**,** 139-45.

229

NOOTER, K., DE LA RIVIERE, G. B., KLIJN, J., STOTER, G. & FOEKENS, J. (1997) Multidrug resistance protein in recurrent breast cancer. *Lancet,* 349**,** 1885-6.

O'CONNOR, R., HEENAN, M., CONNOLLY, L., LARKIN, A. & CLYNES, M. (2004) Increased anti-tumour efficacy of doxorubicin when combined with sulindac in a xenograft model of an MRP-1-positive human lung cancer. *Anticancer Res,* 24**,** 457-64.

O'DRISCOLL, L. & CLYNES, M. (2006) Biomarkers and multiple drug resistance in breast cancer. *Curr Cancer Drug Targets,* 6**,** 365-84.

PARK, K., HAN, S., SHIN, E., KIM, H. J. & KIM, J. Y. (2007) EGFR gene and protein expression in breast cancers. *Eur J Surg Oncol,* 33**,** 956-60.

PATTON, N., ASLAM, T. M., MACGILLIVRAY, T., DEARY, I. J., DHILLON, B., EIKELBOOM, R. H., YOGESAN, K. & CONSTABLE, I. J. (2006) Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research,* 25**,** 99-127.

PEIFER, M. (1993) Cancer, catenins, and cuticle pattern: a complex connection. *Science,* 262**,** 1667-8.

PICK, E., KLUGER, Y., GILTNANE, J. M., MOEDER, C., CAMP, R. L., RIMM, D. L. & KLUGER, H. M. (2007) High HSP90 expression is associated with decreased survival in breast cancer. *Cancer Res,* 67**,** 2932-7.

PLODOWSKI, A. & JACKSON, S. R. (2001) Vision: getting to grips with the Ebbinghaus illusion. *Curr Biol,* 11**,** R304-6.

POPESCU, N. C., KING, C. R. & KRAUS, M. H. (1989) Localization of the human erbB-2 gene on normal and rearranged chromosomes 17 to bands q12-21.32. *Genomics,* 4**,** 362-6.

PRESS, M. F., HUNG, G., GODOLPHIN, W. & SLAMON, D. J. (1994) Sensitivity of HER-2/neu antibodies in archival tissue samples: potential source of error in immunohistochemical studies of oncogene expression. *Cancer Res,* 54**,** 2771-7.

PRESS, M. F., SLAMON, D. J., FLOM, K. J., PARK, J., ZHOU, J. Y. & BERNSTEIN, L. (2002) Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. *J Clin Oncol,* 20**,** 3095-105.

RAJPUT, A. B., TURBIN, D. A., CHEANG, M. C., VODUC, D. K., LEUNG, S., GELMON, K. A., GILKS, C. B. & HUNTSMAN, D. G. (2007) Stromal mast cells in invasive breast cancer are a marker of favourable prognosis: a study of 4,444 cases. *Breast Cancer Res Treat.*

RAMPAUL, R. S., PINDER, S. E., GULLICK, W. J., ROBERTSON, J. F. & ELLIS, I. O. (2002) HER-2 in breast cancer--methods of detection, clinical significance and future prospects for treatment. *Crit Rev Oncol Hematol,* 43**,** 231-44.

RHODES, A., JASANI, B., ANDERSON, E., DODSON, A. R. & BALATON, A. J. (2002) Evaluation of HER-2/neu immunohistochemical assay sensitivity and scoring on formalin-fixed and paraffin-processed cell lines and breast tumors: a comparative study involving results from laboratories in 21 countries. *Am J Clin Pathol,* 118**,** 408-17.

RIDOLFI, R. L., JAMEHDOR, M. R. & ARBER, J. M. (2000) HER-2/neu testing in breast carcinoma: a combined immunohistochemical and fluorescence in situ hybridization approach. *Mod Pathol,* 13**,** 866-73.

ROJO, M. G., GARCIA, G. B., MATEOS, C. P., GARCIA, J. G. & VICENTE, M. C. (2006) Critical comparison of 31 commercially available digital slide systems in pathology. *Int J Surg Pathol,* 14**,** 285-305.

ROMANENKO, A., MORIMURA, K., KINOSHITA, A., WANIBUCHI, H., VOZIANOV, A. & FUKUSHIMA, S. (2006) Aberrant expression of E-cadherin and beta-catenin in association with transforming growth factor-beta1 in urinary bladder lesions in humans after the Chernobyl accident. *Cancer Sci,* 97**,** 45-50.

ROY, S., KENNY, E., KENNEDY, S., LARKIN, A., BALLOT, J., PEREZ DE VILLARREAL, M., CROWN, J. & O'DRISCOLL, L. (2007) MDR1/P-glycoprotein and MRP-1 mRNA and protein expression in non-small cell lung cancer. *Anticancer Res,* 27**,** 1325-30.

RUSSELL, A. L., HENDERSON, C. J., SMITH, G. & WOLF, C. R. (1994) Suppression of multi-drug resistance gene expression in the mouse liver by 1,4-bis[2,(3,5-dichloropyridyloxy)]benzene. *Int J Cancer,* 58**,** 550-4.

SCHNITT, S. J., CONNOLLY, J. L., TAVASSOLI, F. A., FECHNER, R. E., KEMPSON, R. L., GELMAN, R. & PAGE, D. L. (1992) Interobserver reproducibility in the diagnosis of ductal proliferative breast lesions using standardized criteria. *Am J Surg Pathol,* 16**,** 1133-43.

SEIDMAN, A. D., FORNIER, M. N., ESTEVA, F. J., TAN, L., KAPTAIN, S., BACH, A., PANAGEAS, K. S., ARROYO, C., VALERO, V., CURRIE, V., GILEWSKI, T., THEODOULOU, M., MOYNAHAN, M. E., MOASSER, M., SKLARIN, N., DICKLER, M., D'ANDREA, G., CRISTOFANILLI, M., RIVERA, E., HORTOBAGYI, G. N., NORTON, L. & HUDIS, C. A. (2001) Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol,* 19**,** 2587-95.

SHARIAT, S. F., PAHLAVAN, S., BASEMAN, A. G., BROWN, R. M., GREEN, A. E., WHEELER, T. M. & LERNER, S. P. (2001) E-cadherin expression predicts clinical outcome in carcinoma in situ of the urinary bladder. *Urology,* 57**,** 60-5.

SHARMA-OATES, A., QUIRKE, P. & WESTHEAD, D. R. (2005) TmaDB: a repository for tissue microarray data. *BMC Bioinformatics,* 6**,** 218.

SHERGILL, I. S., SHERGILL, N. K., ARYA, M. & PATEL, H. R. (2004) Tissue microarrays: a current medical research tool. *Curr Med Res Opin,* 20**,** 707-12.

SHIBANUMA, H., HIRANO, T., TSUJI, K., WU, Q., SHRESTHA, B., KONAKA, C., EBIHARA, Y. & KATO, H. (1998) Influence of E-cadherin dysfunction upon local invasion and metastasis in non-small cell lung cancer. *Lung Cancer,* 22**,** 85-95.

SHIMAZUI, T., SCHALKEN, J. A., GIROLDI, L. A., JANSEN, C. F., AKAZA, H., KOISO, K., DEBRUYNE, F. M. & BRINGUIER, P. P. (1996) Prognostic value of cadherin-associated molecules (alpha-, beta-, and gamma-catenins and p120cas) in bladder tumors. *Cancer Res,* 56**,** 4154-8.

SIMON, R., MIRLACHER, M. & SAUTER, G. (2003) Tissue microarrays in cancer diagnosis. *Expert Rev Mol Diagn,* 3**,** 421-30.

SIMON, R. & SAUTER, G. (2002) Tissue microarrays for miniaturized high-throughput molecular profiling of tumors. *Exp Hematol,* 30**,** 1365-72.

SLAMON, D. J., CLARK, G. M., WONG, S. G., LEVIN, W. J., ULLRICH, A. & MCGUIRE, W. L. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science,* 235**,** 177-82.

SLAMON, D. J., GODOLPHIN, W., JONES, L. A., HOLT, J. A., WONG, S. G., KEITH, D. E., LEVIN, W. J., STUART, S. G., UDOVE, J., ULLRICH, A. & ET AL. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science,* 244**,** 707-12.

SLAMON, D. J., LEYLAND-JONES, B., SHAK, S., FUCHS, H., PATON, V., BAJAMONDE, A., FLEMING, T., EIERMANN, W., WOLTER, J., PEGRAM, M., BASELGA, J. & NORTON, L. (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med,* 344**,** 783-92.

SMITH, M. E. & PIGNATELLI, M. (1997) The molecular histology of neoplasia: the role of the cadherin/catenin complex. *Histopathology,* 31**,** 107-11.

SONKA, H., BOYLE (1998) *Image Processing Analysis, and Machine Vision*, PWS Publishing.

STELLDINGER, P. & KOTHE, U. (2006) Connectivity preserving digitization of blurred binary images in 2D and 3D. *Computers & Graphics-Uk,* 30**,** 70-76.

STROMBERG, S., BJORKLUND, M. G., ASPLUND, C., SKOLLERMO, A., PERSSON, A., WESTER, K., KAMPF, C., NILSSON, P., ANDERSSON, A. C., UHLEN, M., KONONEN, J., PONTEN, F. & ASPLUND, A. (2007) A high-throughput strategy for protein profiling in cell microarrays using automated image analysis. *Proteomics,* 7**,** 2142-50.

SULZER, M. A., LEERS, M. P., VAN NOORD, J. A., BOLLEN, E. C. & THEUNISSEN, P. H. (1998) Reduced E-cadherin expression is associated with increased lymph node metastasis and unfavorable prognosis in non-small cell lung cancer. *Am J Respir Crit Care Med,* 157**,** 1319-23.

SWALES, K. E., KORBONITS, M., CARPENTER, R., WALSH, D. T., WARNER, T. D. & BISHOP-BAILEY, D. (2006) The farnesoid X receptor is expressed in breast cancer and regulates apoptosis and aromatase expression. *Cancer Res,* 66**,** 10120-6.

SWANTON, C., FUTREAL, A. & EISEN, T. (2006) Her2-targeted therapies in non-small cell lung cancer. *Clin Cancer Res,* 12**,** 4377s-4383s.

SZELACHOWSKA, J., JELEN, M. & KORNAFEL, J. (2006) Prognostic significance of intracellular laminin and Her2/neu overexpression in non-small cell lung cancer. *Anticancer Res,* 26**,** 3871-6.

TAKEICHI, M. (1991) Cadherin cell adhesion receptors as a morphogenetic regulator. *Science,* 251**,** 1451-5.

TAMURA, M., OHTA, Y., TSUNEZUKA, Y., MATSUMOTO, I., KAWAKAMI, K., ODA, M. & WATANABE, G. (2005) Prognostic significance of dysadherin expression in patients with non-small cell lung cancer. *J Thorac Cardiovasc Surg,* 130**,** 740-5.

TAWFIK, O. W., KIMLER, B. F., DAVIS, M., DONAHUE, J. K., PERSONS, D. L., FAN, F., HAGEMEISTER, S., THOMAS, P., CONNOR, C., JEWELL, W. & FABIAN, C. J. (2006) Comparison of immunohistochemistry by automated cellular imaging system (ACIS) versus fluorescence in-situ hybridization in the

evaluation of HER-2/neu expression in primary breast carcinoma. *Histopathology*, 48**,** 258-67.

TORHORST, J., BUCHER, C., KONONEN, J., HAAS, P., ZUBER, M., KOCHLI, O. R., MROSS, F., DIETERICH, H., MOCH, H., MIHATSCH, M., KALLIONIEMI, O. P. & SAUTER, G. (2001) Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol,* 159**,** 2249-56.

TROCK, B. J., LEONESSA, F. & CLARKE, R. (1997) Multidrug resistance in breast cancer: a meta-analysis of MDR1/gp170 expression and its possible functional significance. *J Natl Cancer Inst,* 89**,** 917-31.

TSUDA, H., SASANO, H., AKIYAMA, F., KUROSUMI, M., HASEGAWA, T., OSAMURA, R. Y. & SAKAMOTO, G. (2002) Evaluation of interobserver agreement in scoring immunohistochemical results of HER-2/neu (c-erbB-2) expression detected by HercepTest, Nichirei polyclonal antibody, CB11 and TAB250 in breast carcinoma. *Pathol Int,* 52**,** 126-34.

TSUTSUI, S., OHNO, S., MURAKAMI, S., HACHITANDA, Y. & ODA, S. (2002) Prognostic value of epidermal growth factor receptor (EGFR) and its relationship to the estrogen receptor status in 1029 patients with breast cancer. *Breast Cancer Res Treat,* 71**,** 67-75.

TUBBS, R. R., SWAIN, E., PETTAY, J. D. & HICKS, D. G. (2007) An approach to the validation of novel molecular markers of breast cancer via TMA-based FISH scanning. *J Mol Histol,* 38**,** 141-50.

TURASHVILI, G., BOUCHALOVA, K., BOUCHAL, J. & KOLAR, Z. (2007) Expression of E-cadherin and c-erbB-2/HER-2/neu oncoprotein in high-grade breast cancer. *Cesk Patol,* 43**,** 87-92.

TZANKOV, A., WENT, P., ZIMPFER, A. & DIRNHOFER, S. (2005) Tissue microarray technology: principles, pitfalls and perspectives--lessons learned from hematological malignancies. *Exp Gerontol,* 40**,** 737-44.

VAN DEN BROEK, L. J. & VAN DE VIJVER, M. J. (2000) Assessment of problems in diagnostic and research immunohistochemistry associated with epitope instability in stored paraffin sections. *Appl Immunohistochem Mol Morphol,* 8**,** 316-21.

VAN OORT, I. M., TOMITA, K., VAN BOKHOVEN, A., BUSSEMAKERS, M. J., KIEMENEY, L. A., KARTHAUS, H. F., WITJES, J. A. & SCHALKEN, J. A. (2007) The prognostic value of E-cadherin and the cadherin-associated molecules alpha-, beta-, gamma-catenin and p120ctn in prostate cancer specific survival: a long-term follow-up study. *Prostate,* 67**,** 1432-8.

VERA-ROMAN, J. M. & RUBIO-MARTINEZ, L. A. (2004) Comparative assays for the HER-2/neu oncogene status in breast cancer. *Arch Pathol Lab Med,* 128**,** 627-33.

VINCENT-SALOMON, A., MACGROGAN, G., COUTURIER, J., ARNOULD, L., DENOUX, Y., FICHE, M., JACQUEMIER, J., MATHIEU, M. C., PENAULT-LLORCA, F., RIGAUD, C., ROGER, P., TREILLEUX, I., VILAIN, M. O., MATHOULIN-PELISSIER, S. & LE DOUSSAL, V. (2003) Calibration of immunohistochemistry for assessment of HER2 in breast cancer: results of the French multicentre GEFPICS study. *Histopathology,* 42**,** 337-47.

VLEMINCKX, K., VAKAET, L., JR., MAREEL, M., FIERS, W. & VAN ROY, F. (1991) Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell,* 66**,** 107-19.

VOGEL, U. F. & BUELTMANN, B. D. (2006) Simple, inexpensive, and precise paraffin tissue microarrays constructed with a conventional microcompound table and a drill grinder. *Am J Clin Pathol,* 126**,** 342-8.

WANG, H., WANG, H., ZHANG, W. & FULLER, G. N. (2002a) Tissue microarrays: applications in neuropathology research, diagnosis, and education. *Brain Pathol,* 12**,** 95-107.

WANG, S., HOSSEIN SABOORIAN, M., FRENKEL, E. P., HALEY, B. B., SIDDIQUI, M. T., GOKASLAN, S., HYNAN, L. & ASHFAQ, R. (2002b) Aneusomy 17 in breast cancer: its role in HER-2/neu protein expression and implication for clinical assessment of HER-2/neu status. *Mod Pathol,* 15**,** 137-45.

WANG, S., SABOORIAN, M. H., FRENKEL, E. P., HALEY, B. B., SIDDIQUI, M. T., GOKASLAN, S., WIANS, F. H., JR., HYNAN, L. & ASHFAQ, R. (2001) Assessment of HER-2/neu status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence in situ hybridization assay as the standard. *Am J Clin Pathol,* 116**,** 495-503.

WARANABE, I. (2007) Laboratory of Isao Waranabe, Visual Illusion.

WARFORD, A., HOWAT, W. & MCCAFFERTY, J. (2004) Expression profiling by high-throughput immunohistochemistry. *J Immunol Methods,* 290**,** 81-92.

WATERWORTH, A., HANBY, A. & SPEIRS, V. (2005) A novel cell array technique for high-throughput, cell-based analysis. *In Vitro Cell Dev Biol Anim,* 41**,** 185-7.

WEAVER, D. L., KRAG, D. N., MANNA, E. A., ASHIKAGA, T., HARLOW, S. P. & BAUER, K. D. (2003) Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. *Mod Pathol,* 16**,** 1159-63.

WEI, B., BU, H., ZHU, C. R., GUO, L. X., CHEN, H. J., ZHAO, C., ZHANG, P., CHEN, D. Y., TANG, Y. & JIANG, Y. (2004) [Interobserver reproducibility in the pathologic diagnosis of borderline ductal proliferative breast diseases]. *Sichuan Da Xue Xue Bao Yi Xue Ban,* 35**,** 849-53.

WEINSTEIN, R. S., DESCOUR, M. R., LIANG, C., BARKER, G., SCOTT, K. M., RICHTER, L., KRUPINSKI, E. A., BHATTACHARYYA, A. K., DAVIS, J. R., GRAHAM, A. R., RENNELS, M., RUSSUM, W. C., GOODALL, J. F., ZHOU, P., OLSZAK, A. G., WILLIAMS, B. H., WYANT, J. C. & BARTELS, P. H. (2004) An array microscope for ultrarapid virtual slide processing and telepathology. Design, fabrication, and validation study. *Hum Pathol,* 35**,** 1303-14.

WERNER, M., CHOTT, A., FABIANO, A. & BATTIFORA, H. (2000) Effect of formalin tissue fixation and processing on immunohistochemistry. *Am J Surg Pathol,* 24**,** 1016-9.

WESTER, K., WAHLUND, E., SUNDSTROM, C., RANEFALL, P., BENGTSSON, E., RUSSELL, P. J., OW, K. T., MALMSTROM, P. U. & BUSCH, C. (2000) Paraffin section storage and immunohistochemistry. Effects of time, temperature, fixation, and retrieval protocol with emphasis on p53 protein and MIB1 antigen. *Appl Immunohistochem Mol Morphol,* 8**,** 61-70.

WHITEFORD, C. C., BILKE, S., GREER, B. T., CHEN, Q., BRAUNSCHWEIG, T. A., CENACCHI, N., WEI, J. S., SMITH, M. A., HOUGHTON, P., MORTON, C., REYNOLDS, C. P., LOCK, R., GORLICK, R., KHANNA, C., THIELE, C. J., TAKIKITA, M., CATCHPOOLE, D., HEWITT, S. M. & KHAN, J. (2007) Credentialing preclinical pediatric xenograft models using gene expression and tissue microarray analysis. *Cancer Res,* 67**,** 32-40.

WIJNHOVEN, B. P., DINJENS, W. N. & PIGNATELLI, M. (2000) E-cadherin-catenin cell-cell adhesion complex and human cancer. *Br J Surg,* 87**,** 992-1005.

WINER, E. P. & BURSTEIN, H. J. (2001) New combinations with Herceptin in metastatic breast cancer. *Oncology,* 61 Suppl 2**,** 50-7.

YASUDA, N., NAMIKI, K., HONMA, Y., UMESHIMA, Y., MARUMO, Y., ISHII, H. & BENTON, E. R. (2005) Development of a high speed imaging microscope and new software for nuclear track detector analysis. *Radiation Measurements,* 40**,** 311-315.

ZENG, X., WU, S. F., ZHOU, W. X., LI, D. J., GAO, J., LIANG, Z. Y. & LIU, T. H. (2006) [EGFR and HER2 gene expression status and their correlation in non-small cell lung cancer]. *Zhonghua Bing Li Xue Za Zhi,* 35**,** 398-402.

ZERKOWSKI, M. P., CAMP, R. L., BURTNESS, B. A., RIMM, D. L. & CHUNG, G. G. (2007) Quantitative analysis of breast cancer tissue microarrays shows high cox-2 expression is associated with poor outcome. *Cancer Invest,* 25**,** 19-26.

ZHOU, Y. N., XU, C. P., HAN, B., LI, M., QIAO, L., FANG, D. C. & YANG, J. M. (2002) Expression of E-cadherin and beta-catenin in gastric carcinoma and its correlation with the clinicopathological features and patient survival. *World J Gastroenterol,* 8**,** 987-93.

ZIMPFER, A., SCHONBERG, S., LUGLI, A., AGOSTINELLI, C., PILERI, S. A., WENT, P. & DIRNHOFER, S. (2007) Construction and validation of a bone marrow tissue microarray. *J Clin Pathol,* 60**,** 57-61.

ZU, Y., STEINBERG, S. M., CAMPO, E., HANS, C. P., WEISENBURGER, D. D., BRAZIEL, R. M., DELABIE, J., GASCOYNE, R. D., MULLER-HERMLINK, K., PITTALUGA, S., RAFFELD, M., CHAN, W. C. & JAFFE, E. S. (2005) Validation of tissue microarray immunohistochemistry staining and interpretation in diffuse large B-cell lymphoma. *Leuk Lymphoma,* 46**,** 693-701.