The VPS ReplaySuite

Development And Evaluation Of A Novel, Internet Based Telepathology Tool

Submitted By: Dan Johnston B.Sc

For The Qualification Of Ph.D.

From The School Of Biotechnology,

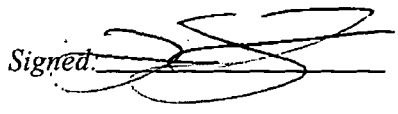
Dublin City University

Under The Supervision Of:

Dr Donal O'Shea

24th September 2005

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 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: 96494522 **Date: 24th of September 2005**

Published papers related to the work in this thesis

Johnston, DJ, Costello, SP, Dervan, PA, O'Shea, D G. (2005) Development and preliminary evaluation of the VPS ReplaySuite: a virtual double-headed microscope for pathology. *BMC medical informatics and decision making,* 5, 10.

Costello SP, Johnston DJ, Dervan PA, O' Shea DG. (2003) Development and Evaluation of the Virtual Pathology Slide: A New Tool in Telepathology. *Journal of Medical Internet Research* ;5(2):e11

Published abstracts related to the work in this thesis

Costello SP, Johnston DJ, Dervan PA, O'Shea DG. (2002) Evaluation of the virtual pathology slide: Using breast needle core biopsy specimens. *British Journal of Cancer,* 86: S1:S34

Costello SP, Johnston DJ, Dervan PA, O' Shea DG. (2002) The virtual pathology slide: A new internet telemicroscopy tool for tracing the process of microscopic diagnosis and evaluating pathologist behaviour. *Archives of Pathology and Laboratory Medicine,* 126:7 781-802.

Conferences where work related to this thesis was presented by the author

The VPS ReplaySuite: A novel application for use in pathology training. *Healthcare Informatics Society of Ireland (HISI), 8th Annual conference and scientific symposium.* Dublin. November 2003

The Development Of An Internet Based Pathology Training Tool. *6th European Congress of Telepathology, 15th Congress of International Society of Diagnostic Quantitative Pathology (ISDQP), and The 8th Congress of The European Society for Analytical Cellular Pathology (ESACP).* September 14-20 2002, Heraclion, Crete

Poster presentation of work related to this thesis

The VPS ReplaySuite: An Electronic Alternative To The Double-headed Microscope. *Irish Association For Cancer Research (IACR),* Annual meeting 2003Kilkenny. April 2003

Table of Contents

Abstract

The ReplaySuite is a web-based telepathology tool that replicates the double-headed microscope environment online, enabling a reviewing pathologist to 'replay' an archived virtual slide examination. Examination-tracking data obtained by the Virtual Pathology Slide (VPS) virtual slide viewer is exploited, allowing a remote pathologist to review an examination conducted at a different time and location. This removes temporal and spatial issues associated with double-headed microscopy.

In order to conduct a preliminary evaluation of the technology, 9 pathologists used the ReplaySuite to review examination replays and diagnostic data from archived examinations of 10 needlecore breast biopsies. Diagnostically difficult cases were most frequently evaluated, either via diagnostic concordance graphs or examination replays, and all 3 participants who replayed more than 10 examinations stated the ReplaySuite to be of some or great benefit in pathology training and quality assurance. Of those who replayed an examination by another pathologist, 83% (5/6) agreed that replays provided an insight into the examining pathologists diagnosis, and 33% (2/6) reconsidered their own diagnosis for at least one case. Of those who reconsidered their original diagnosis, all re-classified either concordant with group consensus or original glass slide diagnosis.

This study demonstrated that the ReplaySuite was of potential benefit in pathology education, however the technology required evaluation in a setting that would facilitate its impact on diagnostic performance. Accordingly, a redeveloped VPS and ReplaySuite were incorporated into the EQUALIS External Quality Assurance (EQA) study in chronic hepatitis staging and grading.

During the study, 9 Swedish pathology departments examined and scored digital representations of liver needlecore biopsies during two sessions, with 10 cases per session and two digital slides per case. Between scoring sessions, participants were provided with access to two supplementary electronic resources: the ReplaySuite, and a library of pre-selected reference images. Comparison of

concordance with gold standard (KVAST group) scoring before and after electronic resource use facilitated the elucidation of impact on diagnostic performance.

Between scoring sessions, participant concordance with KVAST staging increased by 18% (49%-67%), while concordance with KVAST grading increased by 20% (34%–54%). Mean staging un-weighted kappa improved from 0.347 to 0.554 (+0.207), or from 'fair' to 'moderate' exact agreement with KVAST staging. Linear weighted staging kappa improved from 0.603 to 0.688 (+0.085), indicating close agreement in both sessions. Mean grading un-weighted kappa increased from 0.132 to 0.412 (+0.280), or from a 'poor' to 'moderate' level of exact agreement with KVAST, while linear weighted kappa improved from 0.328 to 0.624 (+0.295), or from 'fair' to 'good' level of approximate agreement with KVAST.

Subsequent to the EQA scheme, an expert liver pathologist used the ReplaySuite to evaluate study examinations, assessing examination technique and identifying sources of error. Examinations scoring concordant with KVAST were observed to exhibit acceptable examination technique more frequently than discordantly scoring examinations. When grading, 28% (46% - 18%) more concordant than discordant examinations were considered to have viewed sufficient tissue, and at the appropriate magnification. A similar disparity of 24% (59% - 35%) was observed in staging, suggesting that examination technique was important both when determining the degree of necroinflammation within a biopsy, and when ascertaining the extent of fibrosis.

In assessing sources of error, the expert pathologist identified a potential source in 50% of grading examinations, with misinterpretation of observed pathology cited in 19%, and missed pathology (oversight) cited in 31% of grading examinations. Of the 41% of staging examinations in which a source was identified, misinterpretation of observed pathology was cited in 20% of examinations, and missed pathology (oversight) in 21% of examinations.

This study demonstrated that the use of supplementary electronic resources could result in improvements in diagnostic performance. It also illustrated the significant 'add on' value that could be provided by the ReplaySuite in EQA, by providing means to assess not only diagnostic concordance, but also diagnostic technique and identify sources of error.

In order to assess Irish trainee pathologist's perceptions of computer-assisted learning (CAL), a number of commercial systems were utilised to incorporate digital slides into a postgraduate seminar series, and provide subsequent access to seminar digital slides, diagnoses and expert annotations online. All surveyed trainees considered the use of digital slides and expert annotations of benefit in pathology training, and considered the potential implementation of expert examination replays, online self-assessment and the capability to search online for material by organ, diagnosis or pathological feature of benefit.

The work described herein illustrates that both expert and trainee pathologists alike consider the use of supplementary electronic resources of benefit in pathology education, and demonstrates that their use can improve diagnostic performance. The ability to evaluate participation in EQA studies via the ReplaySuite provides significant additional value to education schemes, providing a depth of assessment not possible with conventional microscopy.

Acknowledgements

I once heard it said that by the end, a PhD thesis becomes such a monstrous behemoth that it engulfs the unfortunate post grad's every waking moment. After four years, I can safely say this is incorrect. It even invades your dreams.

Even now, I still can't decide whether the last four years have dragged or flown. From those carefree early afternoons working (playing championship manager) in the lab, to long, long nights in the air-conditioning deficient IA room, it has certainly been an experience, but one I think I am definitely the richer for. It is said that nothing worthwhile in life is easy. After four years, I think I can definitely agree.

Firstly, my sincerest thanks to Donal, who realised I could to a PhD long before I did. His knack for knowing when to lend a helping hand, and when to leave me to my own devices helped me to grow, and I thank him for it. Maybe it is because we are so similar that we worked well together, but a finer supervisor I don't think I could have hoped for.

Secondly, my sincerest thanks to all the Irish pathologists who helped me with this work, especially Professor Peter Dervan for answering any pathology related questions that sprang to mind.

My thanks to Lennart Franzen, who saw a half-baked presentation in Crete all those years ago, yet still thought we would be worth working with. Without his contribution, and the efforts of Anita and Gunnar at EQUALIS, this work would be much the poorer.

To Sean, John, Cormac, Catherine, Lynne, Jenny, Dawn and Eimear (have I forgotten anyone?) who made the lab a great place to work. From nights in Barcelona and Budapest to simple afternoon coffee breaks and football, I am grateful for it all (especially seeing John secretly love wearing my Arsenal jersey!). And of course watching 10 Christmas sambucas in reverse, that will always be remembered fondly.

Thanks to those who make up the being that is SlidePath, their work with Digital Slidebox made my life so much easier in the final stages of my PhD.

To all my mates (who probably will never read this), Gar, Steve, Sean, Mark and everyone else who is always around at the weekend to help me kill any newly grown brain cells...my thanks. And it's not my round.

To my babe Milena, who had to put up with having half a boyfriend for the last ten months, and a tired, cranky half at that. Who would have thought id get so lucky as to meet someone as amazing as you? Now I'm finished, we have lost of time for mischief! And Nit. Yes.

Most importantly, I would like to dedicate this thesis to my parents. Over the course of this work my trips home became less and less frequent, to the point where they became biannual events. Yet my parents continued to support me, ringing and even visiting when they could. Even though I mightn't always act like it, it means so much.

And finally...to google, the best friend a web-based PhD can have.

Table of Figures

List of Tables

Chapter 1: Overview Of Pathology Education and Computer-Assisted Learning in Pathology Training

1.1 Human Diagnostic Reasoning

"I think therefore I am"- René Descartes

Philosophers and psychologists have long debated how the mind works, what constitutes intelligence, and how humans develop and refine decision-making processes. Cognitive psychology is concerned with the human mind, how it creates meaning, how it processes information it receives (input) to develop responses (output), and how those responses (output) in turn can influence subsequent input. (Anderson, 2000, Dawson and Medler, 2004).

The diagnostic process, or diagnostic reasoning, involves diverse cognitive activities including information gathering, pattern recognition, problem solving and decision-making. It underlies many intelligent activities, such as situation assessment, language understanding, interpretation of scientific observations and medical diagnosis. Much of the knowledge of human diagnostic reasoning is based on general psychological experiments about reasoning, and direct studies of the diagnostic process itself, built upon pioneering work in the 50's by researchers such as Noam Chomsky, George Miller, Alan Newell, and Herbert Simon (David *et al.*, 2004).

1.1.1 Reasoning Strategies

Kassirer (1989) identified 3 strategies of human diagnostic reasoning: Causal, Probabilistic & Deterministic. Causal reasoning seeks to establish the relationship between cause and effect, with cause generally preceding effect. Although causal inferences (reasoning from causes to consequences) can be viewed as the inverse of diagnostic inferences (reasoning from consequences to causes), studies have shown that when making judgments under uncertainty, humans assign greater impact to causal rather than diagnostic data of equal informative weight (Berner, 1999). Evans and Patel showed that experts rarely

rely directly on causal reasoning and knowledge of basic sciences, except when reasoning outside their domain of expertise (Evans and Patel, 1989).

Deterministic or categorical reasoning relates to the use of an appropriate set of routines or rules that apply to the great majority of clinical situations. For example, a categorical medical judgment is one made without significant reservations: if the patient complains of pain on urination, obtain a urine culture and consider the possibility of a urinary tract infection. These rules, as applied by the physician, are not absolutely deterministic. Although their selection and use do not involve deep reasoning, the doctor may withhold his/her full commitment from conclusions reached by even such categorical rules, establishing the flexibility to modify conclusions and re-think the problem if later difficulties arise. A categorical decision typically depends on relatively few facts, its appropriateness is easy to judge, and its result is unambiguous (Szolovits and Pauker, 1978).

However not every decision can utilise such a rule-based approach, such as when to discharge a patient or perform a liver biopsy. In such circumstances evidence is weighed up using probabilistic reasoning, and while it is known that doctors do this, it is not fully understood just how evidence that favours and opposes various hypotheses or courses of action is weighed. While doctors consider probabilistic concepts during their reasoning, observational and experimental studies show that humans are not intuitively good statisticians (Tversky, 1974).

1.1.2 Expert Vs Novice Reasoning

Norman (1982) contended that learning a complex skill takes a minimum of 10,000 hours of practice of the targeted performance. Researchers such as deGroot (1965) have shown that, on average, the achievement of expert levels of performance in any domain requires about ten years of full-time experience, where expert status (in medicine) is identified by certification by sanctioned bodies. This level of experience results in a considerable knowledge base of information and prior experiences, from which the expert can draw. However,

this process of obtaining experience through practice also alters the cognitive mechanisms through which decisions are made.

1.1.2.1 Chunking

The cognitive processes of experts and novices are quite different. Experts have a huge library of facts specific to their field, and thousands of 'organisers' to help them 'chunk' facts efficiently. Miller (1956, 1994) conducted a number of studies that led him to conclude that humans are capable of remembering only 7 discrete bits of information (± 2 bits) for a short period of time without 'processing' them. Miller concluded that people overcome this shortcoming by 'chunking' information, i.e., recoding small bits of information into larger chunks that contain the smaller, related bits of information.

1.1.2.2 Schema

Facts alone, however, are insufficient for adept problem solving and diagnostic reasoning. Experts have developed 'fluid' memories such that the necessary facts are connected to the specific context of the problem-solving process. Novices cannot behave like experts because they do not have the organisers or fluid connections that are developed through practice. They do not see the same problems, so they must follow a linear procedure or algorithm. Their attempts to perform a complex process are deliberate and require tremendous mental effort; whereas, the expert has automated the task to the point of not having to think about its individual components or steps (Bender *et al.*, 2000).

Bartlett (1932, 1958) is credited with first proposing the concept of schema (plural: schemata), based on his observations of the memory recall of stories by subjects that were not actually there. He proposed that memory takes the form of a mental set or representation (schema), which provides a framework for understanding and remembering information. According to this view, cognitive processes are greatly influenced by 'maps' or structures of knowledge stored in the long-term memory (Psybox, 2003). Research on novice versus expert

performance suggests that the nature of expertise is largely due to the possession of schemas that guide perception and problem-solving (Chi *et al.*, 1988).

1.1.2.3 Forward/Backward Reasoning

Most models of diagnostic reasoning include the activation of working hypotheses, the testing of these hypotheses, the acquisition and interpretation of additional information, and confirming, rejecting, or adding of new hypotheses as information is gathered over time (Berner, 1999). Patel *et al* (2001) observed differences in the direction of reasoning, based on levels of expertise. Novices utilise 'top-down' or 'backward' approaches, reasoning from hypothesis to evidence, while experts employ 'bottom-up' or 'forward' approaches to diagnosis, from evidence to hypothesis. Forward reasoning is highly error prone in the absence of adequate domain knowledge as there are no built-in checks of legitimacy. Backward reasoning is best used when domain knowledge is inadequate, as reasoning will be minimally hampered by this lack of knowledge (Cuthbert *et al.*, 1999). However some researchers suggest that mixed strategies are commonly utilised (Azevedo and Lajoie, 1998, Lesgold et al., 1988), and it is not certain how or when competing approaches are swapped as experience increases.

1.2 Diagnostic Reasoning in Pathology

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 judgement under uncertainty. Clinical diagnosis fits the criteria of being an ill-structured problem, similar to the task an architect faces in creatively designing a house 'from scratch' - the realm of solutions encompasses a great variety of applicable methods and a broad set of alternative outcomes (Berner, 1999).

Pathology is the study of disease by looking at changes in the tissues, blood and other fluids of the body (The Royal College of Pathologists of Australasia, 2004). Like radiology (use of electromagnetic radiation such as x-rays to diagnose and treat disease), it is founded on the visual interpretation of images. However unlike radiology, in which images examined are relatively small, the microscopic area available to pathologists for examination can be enormous, in the order of tens of thousands of fields.

While macroscopic examination plays a part, pathology is predominately concerned with the interpretation of microscopic images, the light microscope being the main diagnostic tool in pathology (The Royal College of Pathologists, 2005). Thinly sliced samples of tissue or cells are placed on glass slides and stained with various dyes to make the cells and tissue architecture visible. The diagnostic procedure in microscopic examination begins with qualitative evaluation of a specimen at low power magnification, in order to get an impression of the disturbance of the original tissue. Areas that appear to deviate from normal tissue are then checked at higher power magnification. In the case of malignant growth, additional cell parameters will be checked, such as nucleus shape, inclusion bodies etc, and some parameters semi-quantitatively assessed (Kayser and Schlegel, 1982). It takes many years of experience to learn to distinguish normal cells and tissues from abnormal, and to identify the disease process correctly, therefore, practicing pathologists must possess highly evolved searching, perception and identification skills.

1.2.1 Expert Vs Novice Visual Diagnostic Processes

Most histopathology information is based on subjective judgements and is therefore vulnerable to the limitations of human fallibility. This inherent subjectivity means that a degree of inter-observer variability is unavoidable in histological assessment, even amongst experts. However the unique skills used during histological examination introduce additional potential sources of error that can result in discordant diagnoses. Tiersma *et al* (2003) cited 3 potential sources of error in diagnosis based on images; (1) visually searching the image

(2) interpreting the perceived visual information (3) the method of combining the collected information to form a diagnosis. Previous research by Lesgold (1988) into the development of expertise in Radiology showed that experts report more findings and show more and longer reasoning chains than novices. Experts were observed to evoke pertinent schema quickly, and exhibit more flexibility in their use of schemata.

In their work on evaluation of the development of visual expertise in pathology, Crowley *et al* (2003) attempted to differentiate between the diagnostic processes of novices and experts. 28 pathologists of varying levels of experience (novice, intermediate & expert) microscopically examined breast cases while verbalising their diagnostic process using 'think aloud' protocols, which were then coded to assess the cognitive processes used and identify errors.

Significant differences in search, detection, feature identification and data interpretation were observed along the continuum of expertise. Intermediates appeared to apply explicit strategies in searching slides, while experts appeared to use the microscope as a direct extension of their perceptual processes. Novices had difficulty locating or classifying lesions, while intermediates were able to accurately find lesions, but not classify them (Crowley *et al.*, 2003). Crowley aligned these findings with Lesgold's (1988) hypothesis that perceptual skills (identifying that something does not belong) develop earlier than cognitive processes associated with inference (associating pathological changes with diagnoses).

When processing observed pathological changes, intermediates were observed to identify and interpret individual features, while experts arrived at diagnosis sooner using a higher level, 'pattern-matching' approach. Pattern matching in visual diagnosis may reflect the compilation of processes that convert longer sequences of feature-identification and evidence-hypothesis matching into shorter sequences of non-verbalised, higher level pattern matching (Schmidt *et al.*, 1990).

Using 'eye-tracking' equipment to track the scanning patterns of 5 pathologists grading cervical intraepithelial neoplasia (CIN), Tiersma (2003) illustrated that even on single static images, different types of visual scanning patterns are utilised by different pathologists. Two patterns were identified; a 'scanning style' in which many points are focused upon, but each for only a short period of time, and a 'selective style,' in which the observer limits their search to specific points that are studied for longer periods. For a single image, alternate scanning styles were utilised, and a wide spread of diagnoses observed. Examination of an entire glass slide, at multiple magnifications would exponentially increase the number of fields potentially viewed, compounding further the variability in scanning styles.

1.2.2 Development of Visual Diagnostic Expertise

Histopathology training differs to that of most medical specialities in that it involves extensive one-to-one tutoring. While formal teaching components such as seminars are incorporated into training, much of a trainee's time is spent in what might be called an 'apprenticeship;' engaging in intensive one-to-one sessions with an expert pathologist. These meetings provide trainees with an opportunity to develop and hone their diagnostic skills with cases encountered in clinical practice, under the guidance of an experienced pathologist (Dervan, 2005).

Trainees are provided with access to case material prior to sessions, allowing them to examine tissue sections independently and conclude a diagnosis. During sessions, they must then demonstrate their understanding of observed pathology, and present microscopic evidence that supports their conclusions. To assess the trainee's diagnostic reasoning, the attending expert must be able to observe the trainee's microscopic examination as evidence is presented. This is facilitated by the use of a double-headed microscope, a light microscope fitted with two eyepiece heads, which enables both trainee and expert to examine the same glass slide simultaneously. In addition to enabling the assessment of a trainee's examination technique and diagnostic process, the double-headed microscope

allows the expert to identify and highlight pathological features that may have been missed or misinterpreted by the trainee. This is a valuable learning resource for trainee pathologists, as it illustrates how an expert microscopically examines tissue sections, and acts as a reference model for appropriate diagnostic technique.

However, excessive expert pathologist's workloads restrict the time available for these sessions. The requirement for both the expert and trainee to be present at the same time means that trainees only have access to this valuable resource for short periods of time, limiting its effectiveness. If the capability to review expert assessment of cases anytime, anywhere were available, it would be a considerable resource for trainee pathologists.

1.3 Pathology Education in UK and Ireland

The level of complexity and specialisation in pathology requires trainees to undergo extensive training before they are proficient enough in their chosen speciality to practice as clinicians. Pathologists play a pivotal role in the assessment and treatment of disease; misdiagnosis can have serious consequences for patient care. To ensure that high standards are maintained in pathology teaching and training, national and international legislation has been put in place that has led to formalised training programs by recognised bodies.

1.3.1 UK and Ireland Pathology Educating Bodies

Under European legislation (European Medical Directive 93/16/EC and European Specialist Medical Qualifications Order 1995 and its amendments), postgraduate medical education must be structured and undertaken under supervision in approved institutions. In the UK, the Specialist Training Authority (STA) is responsible for maintaining standards of postgraduate medical training (except general practice). In Ireland, this responsibility rests with training bodies recognised by the Irish Medical Council, such as the Faculty of Pathology in the Royal College of Physicians of Ireland (RCPI). These bodies specify curricula

that must be adhered to during the training process, although there is some degree of variability from region to region.

1.3.2 Histopathology Training in the Republic of Ireland

In the Republic of Ireland, candidates wishing to enter Higher Specialist Training (HST) in Histopathology (after internship) are required to possess a minimum of one year experience in a histopathology post, at either Senior House Officer (SHO) or Registrar level, and are required to pass a Faculty of Pathology aptitude test. Successful candidates undergo 5 years histopathology training at an approved institute, in line with Royal College of Pathologists (RCPath) guidelines. The main methods of objectively assessing progress are two Membership of the Royal College of Pathologists (MRCPath) examinations. These are taken during training, with the first MRCPath examination taken after a minimum of one and a half years training in the Specialist Registrar (SpR) grade and the second (MRCPath P2) after three and a half years training in the Specialist Registrar (SpR) grade.

Much of trainee's time is spent widening their experience of handling and examining different types of specimens, cervical screening and autopsies. Training programmes include suitable rotations to cover all the necessary areas of experience and include an appropriate balance between teaching hospitals, district hospitals and specialised units to provide trainees with a breadth of experience. Trainees work under consultant supervision in histopathology, cytopathology and autopsy services until they have passed the MRCPath Part 2, after which they work largely independently.

More formal teaching, such as "black box" sessions and regional/national training courses supplement day-to-day supervised training, although the structure of teaching varies from region to region. Trainees are advised to avail of a period of supervised research as part of specialist registrar training in histopathology, which may be approved prospectively as an accepted component of the training programme.

On completion of 5 years HST under the supervision of a recognised training body, the Faculty of Pathology in conjunction with The Irish Committee on Higher Medical Training (ICHMT) awards a Certificate of Satisfactory Completion of Specialist Training (CSCST). Those who hold citizenship of any Member State of the EU and whose primary qualification (in medicine) was also issued by a Member State, can, if they have completed the majority of their training in Ireland apply to the Medical Council (in Ireland) for a Certificate of Specialist Doctor (CSD). This entitles the doctor to gain entry on the Specialist Register of any EU Member State (The Irish Committee on Higher Medical Training, 2004, Royal College of Physicians of Ireland, 2004).

The objective of specialist training in histopathology is to produce clinicians who are competent to practice at the consultant level in the speciality (and sub-speciality) of histopathology. Specialists develop interpretative skills at both macroscopic and microscopic levels such that clinically useful opinions can be produced from surgical, biopsy and cytology specimens and from the findings of post mortem examinations. However, due to pathology's inherently complex nature, pathologists must keep abreast of new research and recommendations that may impact on their day-to-day practice. Training and education is not a 'one off' experience, but a continuing process in which pathologists evolve, hone and refine their skills, even when experienced enough to have attained 'expert' status.

1.3.3 Continuing Medical Education

Medical practitioners are held to unprecedented levels of accountability in modern medicine, and lifelong learning skills are critical to being able to adapt to the changing needs of both patients and the healthcare profession. Continuing Medical Education (CME) attempts to maintain high standards of practice by reinforcing existing clinical knowledge and presenting new developments in medical care (Conn, 1992, Du Boulay, 1997, Russell, 1966).

Participation in CME is acknowledged via accreditation by bodies, such as the United Kingdom Accreditation Service (UKAS), who are recognised by government to assess against standards set by European (CEN) or International (ISO) standardisation bodies. In the UK, Clinical Pathology Accreditation (UK) Ltd (CPA), in partnership with UKAS, provide a means to accredit Clinical Pathology Services and External Quality Assessment Schemes (EQA) (Clinical Pathology Accreditation (UK) Ltd, 2005).

In Ireland, CME involves a 5-year accreditation cycle in which a minimum of 250 credits must be acquired, with one hour of CME activity achieving one credit. Credits are achieved by participating in External Quality Assurance (EQA) schemes; attending subsequent EQA review meetings, publishing articles, attending seminars and conferences, and participating in technology-based programs. At the end of the five year cycle, those who has attained the necessary credits for CME receive a 'Certificate of CME' from the Royal College of Physicians Ireland (RCPI) (Royal College of Physicians of Ireland, 2004).

1.3.3.1 External Quality Assurance

External Quality Assurance (EQA) can be described as:

> *"an external audit of the ability to provide a service of high quality by declaring a defined standard of practice, which is confirmed by peer review (Clinical Pathology Accreditation (UK) Ltd, 2005)"*

Clinical laboratories participating in EQA schemes are sent samples on a regular basis, which they test as if they had been obtained from patients. Results are then returned to EQA centres, which subsequently feedback on participant's performance compared to other laboratories and/or groups of laboratories using the same test method(s) (UKNEQAS, 2001). This usually involves the analysis of identical specimens at many laboratories, and the comparison of results with those of other sites and a 'correct' answer. They may be conducted nationally,

such as the National Liver Histopathology scheme; or regionally, as in the case of the Yorkshire General Histopathology EQA, and may evaluate diagnostic proficiency in a specific system (e.g. the national renal pathology EQAS) or general pathology (e.g. North-West Region Histopathology Scheme) (Clinical Pathology Accreditation (UK) Ltd, 2005).

EQA schemes serve two important functions, (1) they provide an effective means of assessing the quality of work in pathology laboratories (2) they constitute an important educational resource that provides pathologists with access to interesting/rare cases that they may not regularly encounter in clinical practice. Subsequent review meetings facilitate discussion of cases and provide 'add-on value' to such schemes.

1.3.4 Educational and Teaching Strategies in Medicine

In 1993 (revised in 2003), the UK General Medical Council's education committee recommended a paradigm shift in the delivery of medical education from a teaching centred approach, in which the emphasis is on the teacher and what they teach, to a student-centred approach, were the focuses lies on students and what they learn. The development of skills for self-directed learning was considered pivotal to this objective, while the incorporation of modern educational theory and research practices into teaching and learning systems was also recommended. The use of modern technologies, where evidence shows that these are effective, was also considered beneficial (General Medical Council, 1993). This follows a more general trend in education towards encouraging independent learning.

1.3.4.1 Self-Directed Learning

Self-directed learning requires students to be responsible for organising and managing their own learning activities and needs. It involves the learner as an active participant, and encourages the development of a deep approach to learning. Key features of self-directed learning are:

- Identifying learning needs
- Formulating goals
- Identifying resources
- Implementing appropriate activities
- Evaluating outcomes

Intrinsic motivation (e.g. desire to learn) is a greater incentive to self-directed learning that extrinsic motivation (e.g. desire to pass exams). Studies have shown that active learning, such as problem solving and group work, leads to students retaining more information and having a better understanding of topic than passive learning, e.g. lecture attendance and reading. Self directed learning is considered the most likely educational strategy for producing doctors prepared for lifelong learning, and capable of adapting to the changing needs of patients (Spencer and Jordan, 1999, General Medical Council, 1993).

While trainee pathologists develop diagnostic skills during tutoring session, a significant proportion of their time is spent preparing for these sessions. This preparation requires self-directed learning that conforms to the key features described: identification of goals (identify diagnosis and supporting evidence), identification of resources (diagnostic cues), implementation of appropriate activities (hypothesis formation) and evaluation of outcomes (critical assessment of diagnostic reasoning).

1.3.4.2 Deep Vs Surface Learning

Self directed learning encourages deep learning of topics, in which the underlying processes and mechanisms are understood, rather than surface learning, which merely encourages students to reproduce what has been learnt (Coles, 1998). Research has identified the student's approach to learning—surface or deep—as the crucial factor in determining the quality of learning outcomes (Oxford Centre for Staff Development, 1992). A surface approach is common in learning environments that have a heavy workload, an excessive

amount of material, little opportunity to pursue subjects in depth, and an assessment system that provokes anxiety and mainly rewards reproduction of factual information. Courses that foster deep learning, however, commonly provide a context in which students are motivated by the need to know, active learning and exploratory work in small groups (Spencer and Jordan, 1999). Pathologists must have an underlying comprehension of pathological processes in order to diagnostically reason to conclusion (diagnosis). Deep learning is essential to the acquisition of this knowledge, and critical evaluation of the diagnostic process in trainee-expert tutoring sessions can facilitate this.

1.3.4.3 Problem Based Learning

> *I hear, and I forget; I see, and I remember; I do, and I understand*
> (Chinese Proverb)

Problem Based Learning (PBL) describes instructional strategies in which students identify and address issues raised by specific problems, to help develop understanding about underlying concepts and principles. New knowledge and understanding arise through working on the problem, rather than in the traditional approaches in which the new knowledge is a prerequisite for working on the problem. This approach requires students to be self-directed and utilise deep-learning strategies. The underlying assumption of PBL is that the closer the resemblance between the situation in which something is learned and the situation in which it will be applied, the more likely it is that transfer of learning will occur, a phenomenon known as "encoding specificity." (Schmidt, 1983). Trainee-expert tutoring sessions are a good example of the application of PBL in pathology training; allowing trainees to develop diagnostic skills using real cases encountered in clinical practice.

Problem based learning is gaining in popularity as both an educational method and a curricular philosophy, and has been endorsed by bodies such as the World Health Organisation (Spencer and Jordan, 1999). However its impact has yet to

be conclusively assessed. In a review of problem based learning, Eschach (2003) observed that while there is much evidence to support the claim that PBL provides a more challenging, motivating, and enjoyable approach to medical education, it is difficult to conclusively state whether it improves performance. In response to this, Eschach highlighted comments made by Normal and Schmidt (2000), that suggested this lack of 'better performance' may not be due to a lack of improvement, but to the way improvements are assessed.

1.4 Computer-Assisted Pathology

The microscope remains the primary diagnostic tool in pathology, however the utilisation of computing facilities is rapidly gaining significance in modern pathology. Advancements in computing and telecommunications have enabled pathologists to provide a more efficient, rapid and superior service, and facilitate a level of communication with other health-care professionals not previously possible. The development of these facilities has significant ramifications not only for how pathology is practiced on a day-to-day basis, but also on the 'creation' of pathologists; the channels through which trainees learn, the development of diagnostic skills and the assessment of diagnostic profiency.

The invention of the Internet is considered by many to be one of the most important scientific developments of the 20th Century. In terms of advances in communications, it can certainly be held in the same regard as the invention of the telephone by Alexander Graham Bell in 1876. The result of visionary thinking in the early 1960's, the Internet's origins reside with researchers who saw the potential value of allowing computers to share information on research and development in scientific and military fields. Early use of the Internet was mainly restricted to computer experts, engineers, scientists, and librarians. However the development of the World Wide Web in 1991, and subsequent development of user-friendly browsers, hypertext and server-side languages opened the Internet up to the world at large (Howe, 2004, Kristula, 2001, World Wide Web Consortium, 2005).

As with medicine in general, computers and Information Technology (IT) have been incorporated into the everyday practice of pathology. This is illustrated in the Royal College of Physician of Ireland Curriculum for higher specialist training in Histopathology, in which one of the objectives of training is the:

> *"Understanding of information technology sufficient to be able to use computers for producing pathology reports and laboratory statistics, to search databases and to access e-mail and Internet services."*

A national survey of pathologists in the UK by Dennis *et al* (2005) found that 71% of surveyed pathologists had T1 lines (dedicated phone connection supporting data transmission rates of up to 1.54Mbps) in their laboratory, illustrating that adequate IT facilities are available in many laboratories. Pathologists use computer systems on a daily basis; from communicating with colleagues via email, to using presentation software such as PowerPoint (Microsoft Corporation, 2005) during meetings and producing electronic pathology reports. While it is beyond the scope of this work to detail all the uses of computer/IT systems that support pathology practice, there are a number of pertinent examples that are worth describing in greater detail.

1.4.1 Computer-based Patient-Record Systems

Patient records are documents storing patients' medical data, such as observed symptoms, diagnoses and recommended treatments. Traditionally, these records are paper based, however this creates logistical and organisational issues that can reduce their effectiveness. Availability of records, retrieval times and the sheer magnitude of managing large volumes of physical records make paper-based records time-consuming, inefficient and ultimately inhibit the effectiveness of care provided.

Computer-based Patient-Record systems (CPR) (Harmon *et al.*, 2003) are, as the name suggests, a repository of electronically maintained information about an

individuals lifetime health status and health care, stored such that it can serve the multiple legitimate users of the record (Shortliffe and Perreault, 2001). The aims of CPR systems such as the Electronic Medical Record (EMR) (OmniMD, 2004) and (Computerised Physician Order Entry) CPOE (Ash, 2003) are to improve the speed of retrieval of medical records, allow many persons to have simultaneous access to the same medical record; improve data confidentiality while tracing who has accessed it and collect routine data (Rind *et al.*, 1997, Barrows and Clayton, 1996). CPR systems are flexible, in that they can display data in different formats, to suit the information requirements of the viewer. With a paper-based patient record, the user is required to manipulate the data mentally to obtain important clinical information. However, CPR systems allow data to be manipulated electronically, in order to aid interpretation of the information. In a review of 26 CPR publications, Delpierre *et al* (2004) observed that, while not conclusive, CPR systems (or CBPRS) were perceived favourably by physicians and might lead to significant improvements in medical care.

Laboratory Information Systems (LIS) such as SoftPath (Fleege *et al.*, 1992) are database driven applications similar to CPS, but modified to incorporate additional functionality of use in pathology laboratories, such as electronic generation of reports, and image management for maintaining image archives. Trestle's Xcellerator™ suite (Trestle Corporation, 2005), for example, is a whole-slide digital image archiving and caseflow management system, intended to replace traditional paper reports and glass slides with a digitised case-specific database.

1.4.2 Clinical Decision Support Systems

Research has illustrated that medical care in industrialised nations often falls short of optimal care. In their evaluation of adverse events in two London hospitals, Vincent *et al* (2001) found that over 11% of patients experienced adverse effects, of which about half were deemed preventable with standard care, and one third of which led to disability or death. Commenting on errors in histopathology reporting of 4 audits, Ramsey (1999) observed discordance between review and original diagnosis in up to 4% of cases, with an overall rate of between 0.26% and 1.4% for cases in which the error is regarded as significant by virtue of risking patient welfare (Ramsay, 1999).

To address these deficiencies in care, healthcare organisations are increasingly turning to clinical decision support systems (CDSS), which provide clinicians with patient-specific assessments or recommendations to aid clinical decision-making (Kawamoto *et al.*, 2005). Clinical Decision Support Systems are computer programs designed to help health care professionals make clinical decisions, (Sacile *et al.*, 2003, Chang *et al.*, 1999, Hamilton *et al.*, 1996, Yamauchi and Fukatsu, 1995, Cross *et al.*, 2000, Firestone *et al.*, 1998), showing great promise for reducing medical errors and improving patient care. This is illustrated in a review of seventy studies evaluating the ability of Clinical Decision Support Systems to improve clinical practice, in which Kawamoto *et al* (2005) observed a significant improvement in clinical practice in 68% of trials.

The basic components of a CDSS include a *dynamic* (medical) knowledgebase and an *inferencing engine* (usually a set of rules derived from the experts and evidence-based medicine). The knowledge base is the compiled expert medical knowledge used by the system, consisting of literature, statistics, disease-finding relationships and other information. The inference engine is the set of computer algorithms used to process patient findings in relation to the knowledge base (Berner, 1999). Inference engines try to emulate human diagnostic reasoning to a conclusion, with those used in medical CDDSS including rule-based, heuristic, neural networks and Bayesian Belief Networks (BBN).

Different engines use different approaches to replicate human reasoning. BBNs use Bayes' theorem, which examines the probability of an event occurring, given that other events have already occurred (Montironi *et al.*, 1996, Shek, 1996, Whimster *et al.*, 1996, Wang *et al.*, 1999). Structurally, BBNs comprise a series of evidence (daughter) nodes, representing features (and outcomes), linked to a decision (parent) node that contains the possible diagnostic outcome via conditional probability matrices (CPM) (Morrison *et al.*, 2002). The CPM's role is dependant on the strategy used; backward or forward chaining. A backward chaining inference engine is 'goal-orientated' in the sense that it tries to prove a goal (diagnosis) by confirming the truth of all of its premises (evidence). Forward chaining acts in the other direction, examining the current state of the knowledge base and, finding rules whose premises (evidence) can be satisfied from known given data (Darlington, 1996).

Evaluation of the mechanism by which DSS acquire data used in their diagnostic algorithms, and the interface through which they interact with clinicians to report their results have led some observers to categorise decision support systems into the following classifications (Haug *et al.*, 1994):

- *Alerting system* that respond to clinical data by issuing and alert
- *Critiquing systems* that respond to recorded decisions to alter care by critiquing the decision and suggesting alternatives (if appropriate)
- *Suggestions systems* that respond to requests by proposing a diagnostic/therapeutic course
- *Quality Assurance systems* that retrospectively abstract clinical data from patient records and decisions about the quality of care made (Berner, 1999)

Kawamoto *et al* (2005) identified four features that are strongly associated with a decision support system's ability to improve clinical practice; (a) decision support provided automatically as part of clinician workflow, (b) decision support delivered at the time and location of decision-making, (c) actionable

recommendations provided, and (d) computer based. A common theme of all four features was that they made it easier for clinicians to use a clinical decision support system, suggesting that an effective system minimises the effort required by clinicians to receive and act on system recommendations.

Decision Support systems in the form of Bayesian belief networks have been exploited in several areas of pathology as an aid to diagnostic decision making (Morrison *et al.*, 2002, Hamilton *et al.*, 1996, Hamilton *et al.*, 1995, Bibbo *et al.*, 1993). Quick Medical Reference (QMR) (Arene *et al.*, 1998, Lemaire *et al.*, 1999, Miller *et al.*, 1986) is a diagnostic decision-support system with a knowledge base of diseases, diagnoses, findings, disease associations and lab information, with information from primary medical literature on almost 700 diseases and more than 5,000 symptoms, signs, and labs. Iliad (Lincoln *et al.*, 1991, Turner *et al.*, 1991) is a computerized, expert system for internal medical diagnosis designed to teach diagnostic skills by means of simulated patient case presentations. CytoInform is a forward-chaining DSS originally developed for breast fine-needle aspiration (FNA) cytology (Diamond *et al.*, 2002), but also applied to the classification of endometrial hyperplasia (Morrison *et al.*, 2002). Interestingly, the system maps the diagnostic process using a cumulative probability graph, allowing users to identify diagnostic clues of greatest importance to establishing a diagnosis. In the assessment of diagnostic classification reproducibility of 50 cases of endometrial hyperplasia, users achieved good intra-observer agreement between unaided and supported diagnosis (Morrison *et al.*, 2002). Two participating medical students with little or no prior knowledge of endometrial hyperplasia classification achieved admirable (0.771 & 0.560 weighted kappa) inter-observer agreement, illustrating that acceptable diagnostic performance can be achieved using a DSS, even by inexperienced observers.

1.4.3 Telepathology

Social, economic and geographical factors mean that not every medical institution is capable of providing the best possible diagnostic opinion in-house, as smaller or remote institutes may not have sub-specialists, or even pathology departments. In order to address this disparity, a sub-branch of telemedicine (practice of medicine via telecommunication system) has evolved that enables remote assessment of histological specimens using IT infrastructure, irrespective of geographical location or economic factors. Telepathology is defined as the practice of diagnostic pathology by a remote pathologist, utilising images of tissue specimens transmitted over a telecommunications network (Weinstein, 1986, Weinstein *et al.*, 1989). The concept of telepathology relates less to the existence of a distance between the specimen and the pathologist but the utilisation of a telecommunications network to transmit the histological images and a computer monitor to display them.

One of the earliest uses of telepathology was in Boston in 1968, when a microwave-based telecommunications system was used by Bird to transmit black and white peripheral blood smears images from a clinic to Logan airport (Bird, 1975). In 1973, pathology images of peripheral bone marrow smears and clinical data were transmitted via satellite from a ship moored off the coast of Brazil to a hospital in Washington DC (Riggs *et al.*). However, it wasn't until a 1986 editorial in *Human Pathology* by Weinstein that the term *telepathology* was proposed (Weinstein, 1986). From these modest beginnings, the utilisation of telepathology has grown. A PubMed search now retrieves over 400 articles containing the *telepathology* keyword.

1.4.3.1 Categorisation of Telepathology Systems

Telepathology systems, by definition, require a means by which to digitise, transmit and display images, and may be categorised by their method of achieving this. While Weinstein (2001) defined 12 different categories of systems developed over successive generations (Weinstein *et al.*, 2001),

telepathology systems are generally regarded to belong to one of three categories (Cross *et al.*, 2002):

- Static telepathology
- Dynamic telepathology
- Static-Dynamic Hybrids (Digital Slides)

1.4.3.1.1 Static Telepathology

Static telepathology (Halliday *et al.*, 1997, Weinstein *et al.*, 1997, Nordrum *et al.*, 2004), or Store-and-Forward telepathology refers to systems delivering a small number of pre-selected images over a telecommunications network, usually via E-mail, File Transfer Protocol (FTP) or HyperText Transfer Protocol (HTTP). The relatively small file size of static images also permits them to be stored on CD-ROM or DVD, either for storage or transport. Static systems are inexpensive and are relatively easy to construct, often only consisting of a microscope fitted with a digital camera and attached to an Internet connected PC. A pathologist or technician selects a small number of representative images to be digitised and transmitted. These images however only represent a small section of the entire tissue specimen. In a large scale survey of UK pathologists on telepathology, Dennis *et al* (2005) found that 64% of surveyed pathologists had a digital camera mounted on their microscope, but that only 36% had emailed digital images.

While static images are generally easy to create and transmit, their use in telepathology is not without disadvantages. Pathologists generally prefer to be able to examine the entire tissue section before concluding a diagnosis and studies have illustrated that sampling bias in image selection can be a factor in diagnostic discordance between conventional microscopy and static telemicroscopy (Weinstein *et al.*, 1997, Weinberg *et al.*, 1996). The ability of the image selector to pick appropriate fields is often regarded as crucial to reducing sampling bias, however, Della Mea (1998) illustrated that in the routine

diagnostic work of a frozen-section service, an inexperienced pathologist can select images which are sufficiently informative for a remote diagnosis.

1.4.3.1.2 Dynamic Telepathology

In contrast to static systems, dynamic telepathology provides a more accurate representation of the conventional microscopic environment, enabling examination of an entire tissue section in real-time, at a range of magnifications. Dynamic telepathology refers to systems that enable a pathologist to examine a slide on a remote microscope, over a telecommunications network. In its simplest representation, it involves the transmission of real-time images from a digital camera attached to the microscope, as the examining pathologist instructs the microscope controller (technician or referring pathologist) where to move the stage (Petersen *et al.*, 2000). In its most advanced form, a robotic microscope is remotely controlled by the examining pathologist, providing complete control of the microscope, including stage movements, lighting and contrast adjustment, and objective selection (Singh *et al.*, 2002, Kaplan *et al.*, 2002, Nordrum *et al.*, 1991).

However, in comparison to static systems, in which the financial outlay extends only to the procurement of a digital camera, dynamic systems such as Trestle's MedMicroscopy, Nikon's Coolscope and Zeiss' Axiopath are much more expensive, and require greater technical profiency. Lack of compatibility between different vendor systems also prohibits interactions between different systems, restricting the potential to develop dynamic telepathology networks.

Image quality is always an issue regarding telepathology systems. While 57% of UK pathologists surveyed by Dennis *et al* (2005) considered telepathology image quality to be of insufficient quality to make a diagnosis, others studies have observed that image quality in robotic microscopy is not an obstacle to diagnosis (Singh *et al.*, 2002, Nordrum *et al.*, 1991). It should be noted that of those surveyed by Dennis, only 7% had the use of a full telepathology workstation in their own laboratory, suggesting that this unfavourable perception

of telepathology image quality is not based on experience. Finally, in order to remotely examine a slide, the tissue section, the referring pathologist/technician and examining pathologist must be in their appropriate positions at the same time. For pathologists with heavy workloads this is often a difficult proposition, and successful implementation of remote consultation services requires planning and scheduling of telepathology sessions in which to examine cases.

1.4.3.1.3 Static-Dynamic Hybrids

Static-Dynamic hybrids on the other hand, remove this prerequisite. Hybrids, or Digital/Virtual Slides (Catalyurek *et al.*, 2003, Costello *et al.*, 2003, Dee *et al.*, 2003, Steinberg and Ali, 2001) (Costello *et al.*, 2003, Catalyurek *et al.*, 2003, Steinberg and Ali, 2001, Dee *et al.*, 2003) as they are more commonly referred to, are digital representations of entire glass slides that are examined using customised viewers. They are considered hybrids in that they display static images (rather than real-time images), but mimic light microscopy by permitting the examiner to navigate between images representing the tissue at different magnifications. As with dynamic systems, the entire section may be examined, but unlike dynamic telemicroscopy, there is no requirement for the referring pathologist/technician or even the glass slide to be present during the examination. Importantly, the ability to examine the entire section also removes the potential for sampling bias often cited in static telepathology.

While earlier digital slides were created using modified standard microscopy systems with motorised stages (Costello *et al.*, 2003), the increasing availability of purpose built ultra-fast slide scanners by manufacturers such as Trestle (Trestle Corporation, 2005), Aperio (Aperio Technologies, 2005), 3Dhistech (3DHistech Ltd, 2002) and Dmetrix (Dmetrix Inc, 2005) have enabled the rapid creation of digital slides. Some observers claim that an estimated scanning time of 1min/slide would be sufficient to handle the daily throughput of an average pathology institute (Saeger *et al.*, 2003), and Dmetrix claim that this is feasible with their Miniaturised Microscope Array (MMA) technology (Descour *et al.*, 2002, Dmetrix Inc, 2005).

Digital slides however are not without their disadvantages. Because images are captured prior to virtual slide examination, real-time adjustment of lighting and focusing parameters is not feasible, within the context of what is permissible with light and dynamic microscopy. While adherence to scanning protocols to ensure optimum digitisation should reduce most requirements for these capabilities, some pathologists still prefer to have complete control over their microscopic environment. Some virtual slide viewers such as Aperio's ImageScope allow the digital adjustment of lighting parameters, but this is performed on the static image, post-acquisition, rather than real-time. The inability to 'focus through' histological specimens in the z-plane is also a limitation of current virtual slide systems, especially in examination of thick tissue sections and cytology. However, the continuing decrease in the cost of computer storage will negate the increased storage requirements of digitisation in multiple focal planes, and 'focus through' capabilities will be incorporated into new and existing scanning systems.

1.4.3.2 Applications of Telepathology

The utilisation of telepathology systems has expanded beyond that which it was originally conceived for. The inherent flexibility of telecommunication systems makes telepathology systems ideal for application in a plethora of pathology related fields, in both training and practice of pathology.

1.4.3.2.1 Remote Primary Diagnosis

For small hospitals, telepathology provides a means of obtaining a diagnosis without the time-consuming process of sending glass slides to another institute. This is of particular relevance in the diagnosis of Frozen Sections (FS), which are an important factor in intraoperative decision-making. One of the earliest uses of robotic telepathology was in the remote diagnosis of intraoperative

frozen sections in Norway. Pathologists from the University Hospital of Tromsø remotely examined 17 FS at two rural hospitals and achieved correct benign versus malignant diagnoses in all cases, compared with final glass slide diagnoses (Nordrum *et al.*, 1991). Other studies have achieved similarly favourable results, with Kaplan *et al* (2002) observing 100% agreement between telepathology and glass slide diagnosis of 120 frozen section cases, using the Trestle MedMicro system. Winoker *et al* (2000) found no statistical difference between telepathology and conventional microscopy of 99 intraoperative consultations, while Singh *et al* (2002) achieved 80.9% diagnostic correlation between glass slide and remote robotic microscopy examinations of 47 fine needle aspirate breast smears.

1.4.3.2.2 Teleconsultation

Pathology is a complex sub-discipline of medicine. Different diagnosis can result in significantly different courses of treatment; accurate diagnosis is paramount. Pathologists may often consult with a specialist on difficult cases before concluding a diagnosis. This is a major application of telepathology, as it removes the need to physically deliver glass slides for a second opinion. Studies have shown the use of static images is a suitable means of obtaining a second opinion (Nordrum *et al.*, 2004, Weinstein *et al.*, 1997), and a number of laboratories in the US have telepathology referral facilities via email attachment of histological images (Mullick *et al.*, 1996, Halliday *et al.*, 1997). Some systems have now been in use for over 10 years, with the iPath system providing remote services to countries such as Laos, Cambodia, Nepal and Poland (Brauchli *et al.*, 2004).

However, as discussed earlier, sampling bias is a concern when using static images. Telemic, in contrast, is an open-source telepathology system for remote consultation that allows real-time visualisation of microscopic images, and textual discussion between the referring and consulting pathologist (Petersen *et al.*, 2000). It enables remote consultation sessions using a conventional light microscope with an attached video camera and framegrabber, and does not

restrict examination to pre-selected fields, removing the potential for sampling bias.

1.4.3.2.3 Quality Assurance

The practice of performing Quality Assurance (QA) in pathology is intended to improve the quality of work and increase patient safety, by assessing pathologist's ability to accurately diagnose sets of circulated cases (UKNEQAS, 2001). Undertaking such QA work however incurs significant amounts of administrative overheads. Preparation and distribution of large quantities of slides, collection and analysis of diagnostic data are all expensive and time-intensive tasks. Additionally, participants do not examine identical slides. While very similar, different sections from the same paraffin block are never identical, and specimens such as needlecore biopsies are not suitable for use in QA studies due to the limited quantity of tissue available for distribution. The use of telepathology for Quality Assurance (Burthem *et al.*, 2005, Ellis *et al.*, 2005) possesses a number of advantages over conventional methodologies:

- Slides are not required to be physically distributed to participants, speeding up the process, reducing costs and removing the inherent risk of damaged/lost slides.
- All participants examine the same slide. Identical copies of the same digitised slide images are transmitted. The requirement to only digitise one slide permits the use of needlecore biopsies.
- Fast and efficient data collection, providing the capability for real-time analysis and feedback from participants
- Capability for monitoring examinations and analysing the diagnostic process utilised.

A fundamental shortcoming of conventional QA studies is their difficulty in detecting and elucidating sources of error and the reasons for inter-observer variability. Inability to locate histological artefacts relevant to diagnosis, and inability to correctly identify pathological changes when successfully found are

often cited as the two primary factors involved in failure to conclude an accurate diagnosis (Ramsay, 1999). Conventional studies are unable to monitor microscopic examinations, prohibiting extensive interrogation of the diagnostic process and therefore inhibiting more accurate analysis of the cause of diagnostic discordance. Methods of identifying sources of error in diagnostic technique will be discussed in greater detail in later chapters.

1.4.3.2.4 Education

Of the various applications of telepathology, it is in pathology education that it has been most successfully utilised. Diagnostic pathology involves the classification of disease using tissues obtained during biopsies, operations or autopsies, based on a complex set of visual features identified with the aid of a microscope. To reach a confident diagnosis, pathologists are required to possess well-developed searching, perception and classification skills, developed during intensive specialist training (The Irish Committee on Higher Medical Training, 2004). During this period, the availability of reference material to help build up trainee's knowledge base and develop pattern-recognition skills is imperative. Online educational resources are capable of the rapid publication and update of extensive image and information reference material, and providing an 'always on' resource available from any Internet connected PC (College of Medicine University of Illinois at Urbana-Champaign, 2005, Nayar and Solomon, 2005, Johnston *et al.*, 2005, Lundin *et al.*, 2004, Jones *et al.*, 2002, University of Alabama at Birmingham, 2005, Kronz *et al.*, 2000, Klatt, 1997, Helin *et al.*, 2005, Melton and Swanson, 1997, Feit *et al.*, 2005, Dee and Leaven, 2005). This is of particular benefit to small institutes, as it provides a means of broadening the range of experience obtained during training.

Secondly, there has been a significant paradigm shift towards improving the level of support provided in the development of perceptual skills required during microscopic examination. As previously described, these skills are acquired through intensive microscopic examination sessions supervised by an expert pathologist. However, development of digital slides, and the utilisation of

Bayesian Belief Networks in training software programs (Diamond *et al.*, 2002, Crowley and Medvedeva, 2003), provides a means of developing these skills without the aid of experts, who's time is often severely restricted.

1.5 Computer-Assisted Learning in Pathology

By definition, any use of a computer by a student involved in learning may be considered computer-assisted learning (CAL). Pedantic observers might consider emailing a faculty member for information as computer assisted learning, however for the purpose of this evaluation, the term shall be used to refer to any computer-based system that supplements conventional teaching practices. CAL systems tend to be designed with specific objects and goals, and are often aligned with academic curricula. This differentiates them from more generalised online pathology resources such as Pathmax (Cowper, 2005), which acts as a general pathology Web portal, and information exchanges such as iPath (Brauchli *et al.*, 2005, Kayser *et al.*, 1999).

While CAL systems can be utilised in the teaching of most topics, they are particularly appropriate for use in pathology training, as the use and assessment of images is central to pathology practice. Computer Assisted Learning has been successfully incorporated into the teaching of gross anatomy (Nieder *et al.*, 2000, Nieder and Nagy, 2002), however, it is in supplementing the development of microscopic examination and diagnosis skills that their application may be of greatest use. CAL systems may be implemented with specific educational objectives; Image reference libraries provide an excellent resource for illustrating pathological changes, especially when accompanied by extensive descriptions, while digital slides or systems that utilise Artificial Intelligence, such as DSS or Intelligent Tutoring Systems (ITS) (Described later in this Chapter) can enhance searching and perceptual skills development. Use of a Web-based tutorial for the Gleason grading of prostate carcinoma on needle biopsies was observed to result in a significant improvement in the grading of 55% of images (Kronz *et al.*, 2000), illustrating that the use of CAL can have a quantifiable impact on student performance.

When compared with traditional teaching methodologies, CAL programs are often the preferred method of instruction. When the University of South Carolina incorporated CD-ROM based digital slides into their histology course it was embraced by the majority of faculty and students, and preferred to the use of glass slides (Blake *et al.*, 2003). The inherent flexibility of Web-based CAL systems allows existing content to be updated, and new content to be added quickly and easily. This enables the provision of up-to-date information that incorporates most recent publications and scientific opinion, extending the shelf life of a resource considerably longer than conventional resources such as textbooks. Finally, online CAL systems are available anytime, anywhere to students with an Internet connected PC, making them available for significantly longer periods than laboratory based facilities. In 18 months, users from 40 member institutions conducted more than 6000 individual searches on the GRIPE Digital Library (Jones *et al.*, 2002), viewing more than 37,500 images. This 'always on' availability allows students to learn in their own time, at their own pace, a benefit highlighted by students using CAL programs incorporated into the general pathology course at the University of Edinburgh (Reid *et al.*, 2000).

Virtual microscopy has been successfully integrated into the University of Iowa's (UI) histology course, supplementing light microscopy (Harris *et al.*, 2001, Heidger *et al.*, 2002). During and outside formal laboratory sessions, students are provided with access to labelled and unlabeled digital slides, and obtain a broad experience of diverse histology images from other participating institutions such as The Virtual Slide Box of Histology (Dee and Leaven, 2005). Assessment of student's perceptions at UI indicated that virtual microscopy was considered easier to use, and of higher educational value than traditional microscopy. While formal laboratory contact time was reduced with the incorporation of virtual microscopy, the authors considered this a positive development, as it encouraged self-directed learning, and enhanced efficiency when using light microscopy.

The incorporation of computer-based learning into medical curricula has required the procurement of computer amenities to facilitate their use, and also the development of computer-based assessment. The Online Testing Center (OTC) is a computer laboratory designed specifically for the assessment of health profession college students at the University of Florida (University of Florida College of Medicine, 2000). Comprising 75 permanent, and 55 temporary seats, it utilises XAM (Extensible Assessment Machine), a Web-based evaluation engine designed specifically for high stakes examinations (University of Florida College of Medicine, 2003). The system supports multiple choice, extended matching, short answer, and essay question types, and can incorporate multimedia content, such as HTML, images (JPEG, GIF), video (Quicktime) and audio files. During the 2002/03 academic year, the centre logged over 200 assessment events and 34,000 student hours for Medicine, Dentistry, and Nursing.

CAL systems differ greatly in their design and execution, often incorporating overlapping technologies, which make it difficult to identify distinctive system categories. Static images and digital slides are used both exclusively and simultaneously. Some systems act only as reference libraries, while others incorporate Multiple Choice Questions (MCQ) or self-assessment capabilities. Resources that utilise databases often incorporate search capabilities that allow images and information to be retrieved for a specific system, diagnosis or pathological feature, while AI based systems enable educational programs to simulate human tutoring environments. It is debatable whether CAL systems in pathology should be categorised according to the learning strategies used (Active Vs Passive), the skills they develop (Searching & Perceptual Vs Inference) or the imaging methodologies used (Static Vs Virtual). While some observers have attempted to classify CAL systems in two categories; tutorial and simulation (Abrioux, 1989), this classification has become outmoded. For the purposes of this work, CAL systems will be differentiated into three broad categories: Atlases, Virtual Laboratories and Simulations.

1.5.1 Pathology Atlases

Atlases are structured systems that act as a repository for reference material. Ideally they are online resources, significantly enhancing accessibility and rapid updating of material. Atlases encompass resources that offer rigid, structured learning environments, such as static image libraries and tutorials. Similar to conventional textbooks, students utilising atlases are involved in passive learning, as there is little active interaction with the material, and no incentive provided to encourage the student to direct the learning. Atlases are excellent resources for developing inference skills and facilitating the development of visual knowledgebases of pathological changes. As with conventional teaching, structured reference material is the most commonly available resource. Lesgold's (1988) observations that perceptual skills develop earlier than cognitive processes associated with inference may in part be related to the fact that such a heavy emphasis is placed on these skills in educational material.

Pathology resources that might be considered atlases abound on the Internet, as illustrated by a Google search for the keywords *histopathology websites,* which locates 11,400 results. It is beyond the scope of the this work to describe every online resource that can be categorised as an atlas, however there are a number of extensive and successful examples of pathology atlases that are of benefit to highlight.

1.5.1.1 WebPath

Florida State University's *WebPath* (Internet Pathology Laboratory for Medical Education) (Klatt, 1997, Klatt and Dennis, 1998) resource includes over 2,300 images along with text, tutorials, laboratory exercises, and examination items for self-assessment. Hosted by the University of Utah (U of U) Health Sciences Centre, it has both general and systemic image libraries, but lacks search capabilities. Text and image based self-assessment tests are available in general and organ system pathology, in addition to general review quizzes. The testing component provides instant feedback on performance, keeping a running score and providing a brief explanation when incorrect answers are provided. Tutorials

are provided in roughly 20 different areas, such as breast cancer, prostate pathology, firearms and urinalysis and are both text and image based.

Klatt (1997) observed that the incorporation of WebPath into the curriculum at the U of U was widely approved, with classroom hours being reduced by 30%. Students preferred CAL to traditional glass slides, and consistently rated WebPath the highest (3.8 on a scale of 0-4) of items under evaluation. The resource was heavily used, with 750,000 US and 150,000 non-US 'hits' per month. Importantly, the implementation of WebPath removed the need for a remedial program in pathology for students with poor examination scores, illustrating the cost benefit of successful implementation of CAL. It also noted that pathology examination scores in the University of Utah improved from a mean of 212 before the WebPath project to 230 (+18) after, at a time when national mean increased from 204 to 217 (+13).

1.5.1.2 GRIPE

The Group for Research in Pathology Education (GRIPE) began in 1971 as a forum for educators in medical school pathology to discuss, evaluate and share materials and ideas (Kent, 1977). One of GRIPE's objectives is to provide peer-reviewed instructional materials that are continuously updated and improved, such as medical student objectives, multiple choice and clinical case examination question databases, and a pathology teaching image collection. The GRIPE Digital Library (Jones *et al.*, 2002) contains approximately 3000 peer-reviewed gross and microscopic pathological images and textual descriptors, and is connected to a multiple-choice question bank. The system has proved popular amongst faculty and students alike, with users at 40 GRIPE member institutions using the library to perform more than 6000 individual searches and view more than 37,500 images.

1.5.1.3 PEIR

The University of Alabama at Birmingham's *Pathology Education Instructional Resource* (peir.net) is an excellent example of an extensive online reference library. The largest educational pathology resource online, it contains more than 40,000 multidisciplinary images for use in medical education (Hamza *et al.*, 2001). PEIR's search capabilities allow images to be retrieved from different sources (e.g. University of Alabama Pathology/Radiology), by image type (gross, micro, immunofluoresence) and by keyword. While of great benefit to medical students, PEIR is also an excellent resource for educators, as it allows images to be collected for use in presentations via its 'shopping cart'. When the necessary images have been obtained they are archived and compressed with WinZip (WinZip International LLC, 2005) before transmission.

1.5.1.4 Dermatlas

When their web-based cardiac and respiratory physiology application CV Sim proved a success (Lehmann *et al.*, 1997), John Hopkins University developed a Dermatology image atlas, *Dermatlas* (Lehmann and Cohen, 2002). An interactive multidisciplinary resource for physicians, fellows, residents, medical students and other health care professionals, Dermatlas contains almost 8000 images associated with over 1000 diagnoses. Developed as a learning and reference tool, it provides high quality dermatology clinical and histological images, as well as detailed descriptions of the patient and the associated disease processes. Users may search for images by body area, patient age, skin pigmentation, contributor, image name or any combination thereof, and all text entries are completely indexed, allowing keyword searches. Additionally, every diagnosis is connected via hypertext link to a query of the National Library of Medicine's Pubmed for relevant articles. The Quiz facility allows users to evaluate their ability to diagnose from dermatlas images and clinical data, with up to 20 multiple diagnosis options (extended multiple choice design) available.

1.5.1.5 The Bethesda System Atlas

The *Bethesda system website atlas* (Nayar and Solomon, 2004, Smith *et al.*, 2000) is, as the name suggests, a web-based atlas of 349 images representing a range of morphologic findings seen on both conventional cervical/vaginal smears and liquid based preparations. The atlas provides extensive searching capabilities, and provides a self-test MCQ section. The testing function provides instant feedback on performance, allowing the examiner to view the percentage distribution of answers for each question after it is answered. Interestingly, a study by Smith *et al* (2000) concluded that use of the Bethesda system atlas by itself does not appear to improve the reproducibility or accuracy of cytological diagnoses.

There are a number of other online pathology atlases that are worthy of mention. The University of Illinois at Urbana-Champaign (UIUC) *Atlas of Histology* (College of Medicine University of Illinois at Urbana-Champaign, 2005)contains more than 1000 labelled histological features with accompanying descriptions, which can be searched by keyword. The site links to the UIUC histology course website, which contains lectures presentations and notes, as well as self-assessment image based quizzes. Loyola University *Dermatology Medical Education* (Melton and Swanson, 1997), hosts a number of dermatological image atlases, both clinical & histological, with accompanying descriptions in Dermatology, Anatomy and Histology, Mohs Micrographic Surgery and Skin cancer and benign tumour.

1.5.2 Virtual Laboratories

Virtual laboratory CAL systems are also predominately online resources, but provided a more open learning environment in which the user is an active participant. Similar in architecture to atlases, the most significant difference is the incorporation of digital slides to provide an environment that facilitates the development of searching and perception skills. Many systems may be considered atlas- virtual laboratory hybrids, as they utilise both static and virtual imaging, and it is probably more accurate to categorise these systems as such.

However for the sake of simplicity, these systems will be classified as Virtual laboratories.

1.5.2.1 WebMicroscope

An excellent application developed at the University of Helsinki (with the University of Tampere) is the *WebMicroscope*. This extensive, multi-functional pathology resource utilises digital slides with annotation capabilities, and contains several components:

- Atlas of Breast Pathology
- Interactive Gleason Grading Site
- Slide Seminars
- Publication Supplements

The Atlas of Breast Pathology (Lundin *et al.*, 2004, Lundin *et al.*, 2004) features more than 150 digitised slides accompanied by concise descriptions, and follows published WHO classification of tumours of the breast (WHO, 2003). Using the virtual slide viewer, the user may navigate by zooming and panning, and annotate the slide by selecting regions and marking them with various shapes and text. Self-assessment is permissible through the quiz mode, which hides diagnoses until the user has examined the slide.

The Interactive Gleason Grading Site (Helin *et al.*, 2005) consists of 62 digitised prostate biopsy digital slides, graded by expert uropathologists. Developed as a tool for assessing observer variability of Gleason grading, it allows the examination and grading of slides, in order to obtain instant feedback to assess interobserver agreement with the experts. Additionally, it enables users to re-score a set of biopsies already scored once, to evaluate intra-observer variability. Developed also for teaching and standardising grading, it provides a Schematic diagram of the Gleason grading system by the UPMC Cancer Centers, University of Pittsburgh, Pennsylvania, USA.

The utilisation of Web-based virtual microscopy on WebMicroscope allows participants of clinical meetings; conferences, and slide seminars to view digital slides of the cases before the presentation, such as:

- Dept of Pathology, University of Helsinki - weekly slide seminars
- International Academy of Pathology (IAP) seminars

The Publication Supplements section allows readers of journal articles to view digital slides scanned from microscopy specimens presented in reports.

1.5.2.2 VMic

The Institute of Pathology in Basel's *vMic* (Glatz-Krieger *et al.*, 2005, 2003) site is an attractive web-based resource that enables users to search for digital slides according to organ system. Slides are annotated, allowing the examiner to jump to fields containing described pathological changes. However the system is not particularly intuitive, and features within the field are not marked. While this is not disadvantageous to intermediates, its absence may be a hindrance to novices that are not sufficiently knowledgeable. vMic also provides access to cases presented in seminars, with students able to review both digital slides and clinical data after seminars.

1.5.2.3 Atlas of Dermatology

The Masaryk University in Brno's *Atlas of Dermatology* (Feit et al., 2005) is an example of a hybrid system that utilises both static images and digital slides. The Atlas has two components: (1) a dermatopathology atlas of conditions with clinical, etiology and histology images and information (2) a library of annotated digital slides of melanocytic lesions. Students may freely examine slides, or select to view annotations, allowing them to select specific pathological feature present from a list and view the appropriate region on the slide.

1.5.3 Simulations

The application of Artificial Intelligence (AI) in educational tools allows for the simulation of scenarios that allow the user to develop and apply their knowledge in environments closer to those encountered in practice (real world). Applications such as Decision Support Systems (DSS) and Intelligent Tutoring Systems (ITS) provide an open but structured learning environment that facilitates the development of a wide range of skills: searching, perception, identification, inference and classification. Designed to supplement human tutoring rather than replace it, these systems attempt to replicate the benefits of human tutoring by using AI to provide support, encouragement and when needed, advice during problem solving.

1.5.3.1 Decision Support Systems

Decision support systems (DSS) were developed to reduce errors and improve patient care, however they are increasingly finding application in training and education (Diamond *et al.*, 2002). Their ability to guide users through the diagnostic process and assess performance make DSS ideal for use in developing diagnostic skills, effectively mimicking the performance of a human tutor. Importantly, DSS encourage active, self-directed learning; allowing students to learn at their own pace, identify important diagnostic parameters and evaluate performance.

CytoInform (Diamond *et al.*, 2002, Morrison *et al.*, 2002) is a Bayesian belief Decision Support System that allows evidence nodes for a case to be assessed sequentially, by selecting a position on a slider of possible values. This data is used to generate a cumulative probability graph that shows the belief in the outcome as evidence for diagnostic clues are entered. As a training tool, the system uses model tracing (Anderson *et al.*, 1990) to evaluate user-entered information against the dataset of an expert and identify discrepancies between the user and the expert. As previously mentioned, Morrison *et al* (2002) illustrated that use of CytoInform enabled inexperienced users to achieve acceptable diagnostic performance in the classification of endometrial

hyperplasia. When assessed as a training tool in the training of fine-needle aspiration cytology (Diamond *et al.*, 2002), two pathologists only misinterpreted one case, and assessed and interpreted 86%/88% of clues correctly. CytoInform provides visual prompts for accessing diagnostic clues, which is of particular benefit in training as it helps to reduce subjectivity in diagnosis. This work led to the development of a commercial DSS by i-Path diagnostics (i-Path Diagnostics, 2004). Inview© is a software-based expert system capable of supporting decisions in a wide variety of diagnostic applications, as diagnostic modules can be "plugged" into the core Bayesian belief network-based engine.

1.5.3.2 Intelligent Tutoring Systems

Bloom (1984) concluded that individual human tutoring can significantly improve student performance above normal classroom tuition. Intelligent Tutoring Systems (ITS) (Langer *et al.*, 1998, Martens *et al.*, 2001, Tang *et al.*, 1988) attempt to replicate the benefits of one-on-one instruction by replicating the human tutoring environment, using artificial intelligence to tailor multimedia learning. Designed to supplement rather than replace conventional tutoring, Intelligent Tutoring Systems facilitate student-centred active learning, and encourage deep learning through problem-based tuition. Studies have shown ITS can improve performance beyond classroom instruction, and even approach performance attained with human tutoring (Anderson *et al.*, 1995, Koedinger and Anderson, 1997). ITS's also possesses a number of inherent advantages from both financial and administrative perspectives, and can reduce time required to become proficient in the tutored topic. In a study using their LISP tutor, a computer programming ITS, Corbett and Anderson (1989) illustrated that rapid feedback of the system lead to significant reductions in learning time needed to reach post-training performance levels, with learning time 3 times longer in the most delayed feedback condition that in the most immediate.

Like training simulations, ITS enable participants to practice their skills by carrying out tasks within highly interactive learning environments. However, ITS goes beyond training simulations by answering user questions and providing

individualised guidance. Unlike other CAL technologies, ITS systems assess performance within these interactive environments and develop a model of users's knowledge, skills, and expertise (model tracing). Based on the learner model, ITS's tailor instructional strategies, in terms of both the content and style, and provide explanations, hints, examples, demonstrations, and practice problems as needed (Ong and Ramachandran, 2000).

ITS are usually organised into 4 separate software modules (Crowley and Medvedeva, 2003, Khuwaja and Patel, 1996):

- *Expert Model* - represents subject matter expertise and provides the ITS with knowledge of what it's teaching.
- *Student Model* - represents what the user does and doesn't know, and what he or she does and doesn't have. This knowledge lets the ITS know who it's teaching.
- *Pedagogic model* - enables the ITS to know how to teach, by encoding instructional strategies used via the tutoring system user interface.
- *Interface* – The interface between system and user.

While ITS have been developed for application in a number of fields, only a handful of medical ITS exist (Crowley and Medvedeva, 2003, Martens *et al.*, 2001, Langer *et al.*, 1998, Crowley *et al.*, 2003, Freedman *et al.*, 2001, Azevedo and Lajoie, 1998). Due to the complex nature of pathology, the number of ITS developed for teaching histological assessment is limited (Crowley and Medvedeva, 2003, Crowley *et al.*, 2003). SlideTutor (Crowley and Medvedeva, 2003, Crowley *et al.*, 2003) is a web-deployed, image-based, model tracing ITS for teaching microscopic diagnosis, which provides individualised coaching to students as they search and interpret digital slides. Implemented in the sub-domain of inflammatory diseases of the skin, its expert model encompasses 500 diseases and 3000 visual features, allowing it to guide students to diagnostic areas, offers progressive diagnostic hints in identifying specific visual features and provide a scaffolding structure for hypothesis generation. Crowley *et al* have also developed a natural language interface to SlideTutor, which enables students

to type in diagnostic reports as they inspect a virtual slide. The system monitors where they are, what they are doing, and the content of the report, providing real-time, individualised feedback to help students learn.

ITS's are designed to replicate one-to-one tutoring, but have also been utilised in collaborative, problem based learning environments. COMET (Suebnukarn and Haddawy, 2004) is an ITS for medical problem-based learning in head injury diagnosis, allowing small groups of students to discuss cases online and collectively form hypotheses. Tutoring in PBL intelligent tutoring systems is particularly challenging, as the tutors role is not to instruct but facilitate group discussion and problem solving. The tutor should provide as little guidance as possible while at the same time not allowing the students to get lost. Suebnukarn and Haddawy (2004) identified eight hint strategies commonly used by experienced human tutors and incorporated them into COMET: 1) focus group discussion using general hint, 2) focus group discussion using specific hint, 3) promote open discussion, 4) deflect uneducated guessing, 5) avoid jumping critical steps, 6) address incomplete information, 7) refer to experts in the group, and 8) promote collaborative discussion.

1.6 Defining the Need for New Pathology Training Tools

Atlases are excellent resources for developing inference skills and facilitating the development of visual knowledge bases of pathological changes. However, they do not support the development of searching and perception skills required to successfully locate visual features. It is one thing to be able to identify a pathological feature within a single, small image, but quite another to identify the same feature within an area as large as the examinable region of a tissue section. Digital slides provide the freedom to actively develop searching and perception skills, however unless they are annotated they do not provide a structured learning environment. Without a supportive structure, novices may have difficulty locating relevant pathological changes, somewhat negating the benefits of an open examination environment.

Simulators provide this structure, giving rapid feedback on diagnostic performance that more closely replicates the one-to-one experience of human tutoring. Intelligent tutoring systems (ITS) guide a user through the examination of a tissue section, while decision support systems (DSS) facilitate evaluation of trainee's assessment of diagnostic parameters. However, while these interactions enhance the learning experience, they still do not support the development of searching and perception skills. Additionally, DSS and ITS systems require significant development and validation before they can be used in an educational environment.

Much of trainee's diagnostic skills are acquired through intensive one-on-one sessions with expert pathologists using a multi-headed microscope. This allows the trainee to observe and learn from the diagnostic technique of the expert, while also allowing the expert to assess the abilities of the trainee. However the excessive time constraints placed on pathologists often restrict the frequency and duration of these sessions.

Computer-based examination environments possess the potential to track examinations, creating the possibility of recording examinations and reproducing

them at a later date. Such a 'virtual double-headed microscope' has obvious potential as a training resource, removing the temporal and spatial limitations of conventional double-headed microscopy.

This would also have tremendous potential for providing additional value to External Quality Assurance (EQA) schemes. Conventional schemes are limited to identifying incidences of diagnostic discordance. In contrast, a system that can 'replay' examinations allows those that resulted in discordant diagnoses to be evaluated by an expert pathologist/panel of experts, in order to identify the reasons for discordance.

The following project addressed this need for a virtual double-headed microscope by developing and validating a novel telepathology tool for use in pathology education. This was accomplished through the implementation of the following objectives:

- Development of a new telepathology tool (ReplaySuite) to exploit examination-tracking data collected by the VPS to provide examination 'replays.'

- Evaluation of ReplaySuite technology in a preliminary study and subsequent External Quality Assurance setting.

- Utilisation of ReplaySuite technology to identify sources of error in virtual slide examinations, and evaluate examination technique.

It is hoped that this would illustrate the significant potential of computer-based learning systems in pathology education, and help identify why pathologists diagnose discordant with 'gold standard' diagnoses.

Section 1: - Development Of Technologies For The Deployment Of The ReplaySuite

Abstract

It had previously been illustrated that pathologists were confident in making diagnoses using a specific digital slide technology, the Virtual Pathology Slide (VPS) (Costello, 2004), and that the tracking data obtained from this could be used to identify correlations between regions examined on a needlecore breast biopsy, and diagnoses concluded. The method of exploiting tracking data to make such observations however was cumbersome, and required manual intervention and the use of multiple applications. The result was that the benefit to end users, pathologists, was negligible.

The purpose of the work was to identify whether the use of electronic resources could be of benefit in pathology for educational purposes, and if examination tracking could be exploited to identify sources of error in histological assessment. In order to achieve these objectives, new applications were required that would facilitate the use of tracking data in a manner that was of direct benefit to pathologists.

The following section describes the technical work undertaken to develop multiple new technologies to a sufficient level that they could be utilised for this purpose. It describes the development of the ReplaySuite application to exploit Virtual Pathology Slide examination tracking data (Chapters 2 & 4), and the creation of novel slide scanning algorithms that allowed sufficient quantities of slides to be digitised (Chapter 3). The undertaking of this work was essential in order to achieve the objectives previously described.

Chapter 2: Development Of The ReplaySuite 1.0

2.1 Introduction

Much of a histopathologists training may be considered an apprenticeship, in which the trainee spends significant periods of time with an expert pathologist examining tissue sections with the aid of a double/multi-headed microscope. This piece of equipment allows both expert and trainee to observe the tissue at the same time, enabling the trainee to learn from the expert's diagnostic technique, and facilitating evaluation of the strengths and weaknesses of a trainee's examination technique. However, pathologists are subjected to considerable workloads; expert time is a precious commodity and as such the opportunity to participate in such training sessions is restricted.

In response to this, researchers have attempted to supplement traditional teaching methodologies with Computer-Assisted Learning (CAL). Kronz *et al (2000)* illustrated the potential benefits of tutoring pathologists with CAL, demonstrating that improvements in Gleason grading of prostate carcinoma in needlecore biopsies could be achieved between two grading sessions, if (in-between grading sessions) users were provided with access to reference material selected by an expert. However such reference material consisted of static images with associated descriptions of observed pathological changes. While enhancing inferencing skills used to associate observed pathology with diagnoses, the utilisation of static images does not facilitate the development of searching and perception skills required to locate pathological features.

Intelligent Tutoring Systems, on the other hand, provide a more accurate reflection of the double-headed microscope teaching environment (Crowley *et al.*, 2003). Users are not restricted to static images (examination environment uses digital slides), are provided with suggestions by the system when encountering difficulties, and may even ask questions. However ITS applications are time consuming to create and do not truly replicate the feeling of observing an experts examination, as they do not allow trainees to observe an expert searching for and identifying pathological features, as they would in practice.

Decision Support Systems (DSS) such as CytoInform (Diamond *et al.*, 2002, Morrison *et al.*, 2002) are an excellent means of providing insight into the diagnostic process, and provide a greater understanding of how the assessment of pathological features influences the diagnostic probability. CytoInform also facilitates assessment of trainee's diagnostic abilities by generating a cumulative probability graph that highlights discordant assessment of pathological features with respect to expert evaluation. However, while reference images are provided as a guide to assessing each pathological feature, users are still unable to relate these high magnification images to the examination process used to locate them at lower magnification. If users were able to review the diagnostic process used by the expert while examining the slide it would enhance the understanding of the process through which an expert locates visual features. Such capabilities would require the expert examination to be tracked and information describing the process recorded, however to date, very few computer-based system possess this capability.

Brauchli *et al* (2002) described a dynamic telemicroscopy system that could track examination sessions, saving viewed image fields on a database. However this was utilised primarily for teleconsultation, and the authors did not consider the potential application of this technology in training and education. Tiersma (2003) assessed the visual scanning patterns of pathologists using 'eye-tracking' hardware, however this was restricted to single images, and may be considered invasive monitoring. The utilisation of such technology in anything other than a research environment would be impractical, time consuming, and provide little benefit to trainee pathologists in the development of diagnostic skills.

The VPS was originally developed as a means of tracking the diagnostic trace of virtual slide examinations, for the purpose of elucidating reasons for inter-observer variability (Costello, 2004). The work identified two important issues that contributed to inter-observer variability in the classification of breast biopsies:

- Correlation between areas examined and classification concluded for some cases

- Discrepancies between textual (diagnostic comments) and classification-based diagnoses for some examinations

This suggested that some pathologists discordantly diagnosed as a result of not viewing the correct areas (missed pathology), while others had problems correlating observed pathology with an appropriate classification (misinterpretation/misclassification). However, the methods used to obtain this information were laborious, time consuming and inefficient. 'Heat maps' indicating areas examined were generated from the data by a separate application, while still images of fields viewed during examinations were required to be manually located and captured, then presented to an expert pathologist for review.

As a means of collating examination data, the VPS proved an effective tool for non-invasive observation of diagnostic processes. However, it was not capable of rapidly utilising the data collected to provide an intrinsic benefit to the end user, the pathologist. Data describing expert examinations of digital slides would be of immense benefit as a resource for pathology education. Additionally, the capability to review examination data would potentially have considerable applications in External Quality Assurance (EQA) and the identification of source of error in microscopic examination (Chapter 6).

The objective of this chapter was to develop a web-based application that exploited VPS tracking data to allow examinations to be replayed. In response to this inherent limitation of the VPS, the ReplaySuite was developed. The primary function of the ReplaySuite was to utilise VPS collected data to mimic the use of a double-headed microscope, allowing one pathologist to observe the examination of another, in a manner similar to watching a video of the examination. The following Chapter describes the technical development of the ReplaySuite (version 1.0), and the process through which it interacts with, and exploits examination data obtained by the VPS (version 1.0).

2.2 Technical Implementation

In order to deliver functionality to end users, the ReplaySuite utilised the same programming languages used to develop the VPS version 1.0 (Costello, 2004). A more detailed description of the technology used is provided in Appendix A. Three key architectural features of VPS 1.0 were maintained in development of the ReplaySuite:

Client-side Customised Browser

A customised browser was developed to display the ReplaySuite Graphical User Interface (GUI), control access for users and ensure a uniform experience for users who would otherwise experience subtle differences due to the variety of web browser versions available, such as Internet Explorer, Firefox, Opera and Netscape Navigator. Constructed using Visual C++ application wizard, the VPS Customised Browser was a Microsoft Foundation Class (MFC) application that utilised Internet Explorer file libraries in order to behave as a customised browser.

Web-based Delivery

The system utilised web-based delivery methodologies for transmitting images and data between the VPS server and the remote client application. PHP (PHP Hypertext PreProcessor), a server-side scripting language, was used to dynamically generate HTML (HyperText Markup Language) documents that comprised the Graphical User Interface (GUI), while JavaScript function libraries were used to provide client side interactivity (Appendix A).

Web-based delivery was used as it enabled remote access to VPS images and data from any Internet enabled PC with the client application. PHP was used for two reasons: (1) it is an open-source language that facilitates database communication (2) it enabled dynamic GUI generation.

Database Driven Architecture

Examination and study data was stored on a relational, database management system called Oracle 8i. PHP's Oracle Call Interface (OCI) was used to communicate with the database, submitting SQL (Structured Querying Language) queries that enabled information to be written to, and retrieved from the database. A relational database was selected as it enabled efficient management of the data and rapid access and retrieval of information.

2.2.1 Database Architecture

Data that described VPS examinations was stored in a relational database management system (Oracle).

Figure **2.1** is an entity relationship diagram that illustrates the architecture of the database. Tables within the database stored information describing examined fields (*sessionlist*, *tracking*), the examiner (*users, machinespec)* the slide examined (*slidespec)* and the concluded diagnosis (*sessionreport)*, and was accessed via SQL querying (Appendix A). The ReplaySuite used the same methods of communicating with the database as the VPS with one important distinction: The VPS wrote data to the database, the ReplaySuite read data from it.

2.2.2 Client-Server Interaction

Figure 2.2 illustrates the interaction between ReplaySuite application and server. When a user performed an action, the application requested a dynamic page from the web server. The web server located the page, embedded any variables passed from the application and sent the page to the application server, which parsed the file for PHP tags. The application server passed embedded queries to the database driver, where they were executed, retrieving a record set from the database. This record set was passed back to the application server, where it was inserted into the document as instructed. The static HTML document was then passed to the web server, and back to the ReplaySuite application where it was displayed.

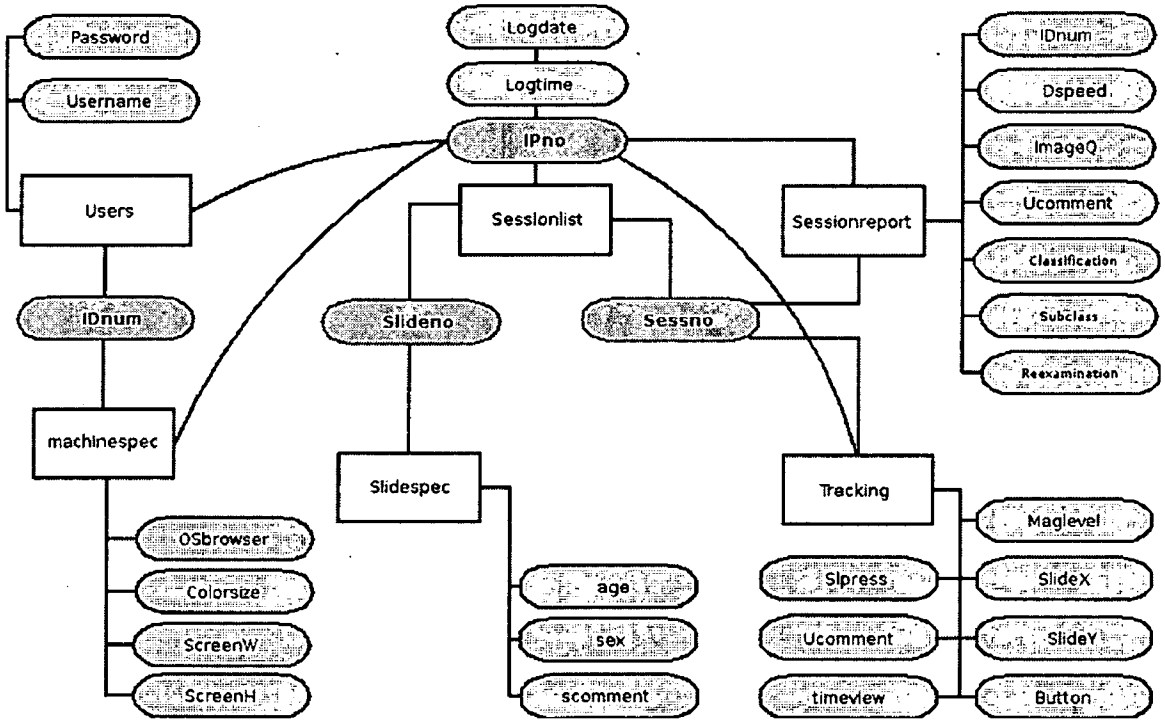


Figure 2.1 Entity relationship diagram of the VPS (version 1.0) database. Tables containing data are denoted by white boxes, data columns by light grey and key data columns by dark gray. Tables contain information describing the examination (sessionlist), the concluded diagnosis (sessionreport) and the fields examined (tracking).

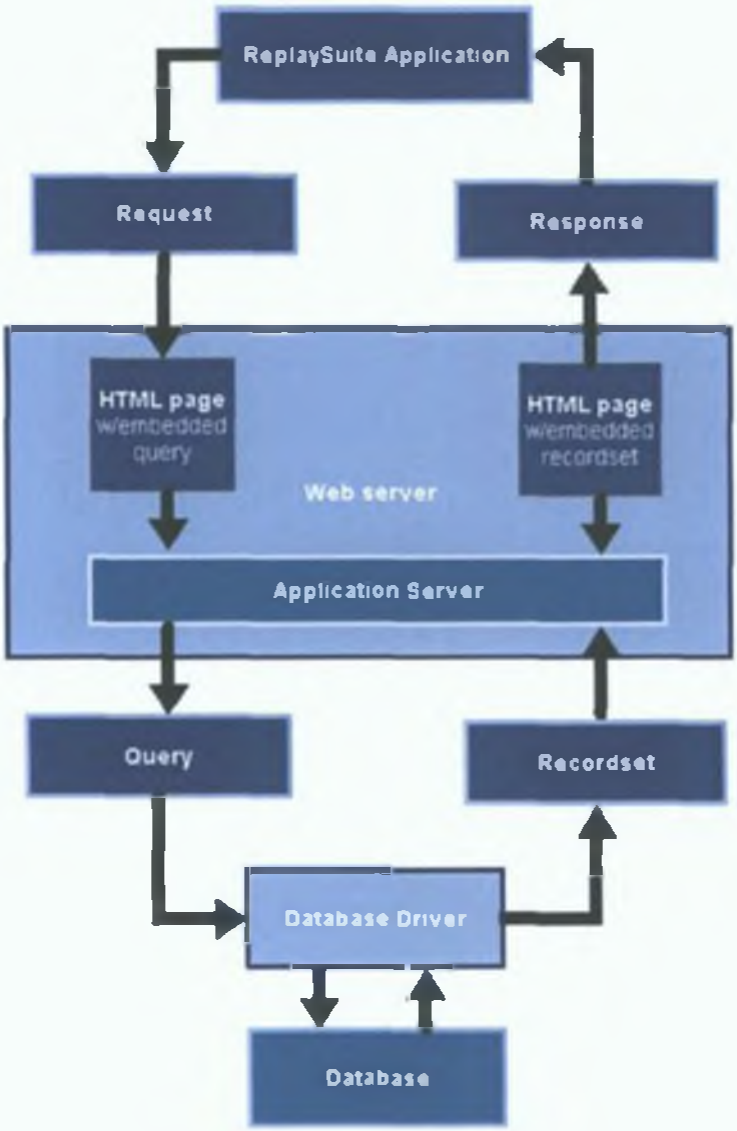


Figure 2.2 ReplaySuite 1.0 client application interaction with server side VPS oracle database when requesting a file.

2.3 Graphical User Interface

The ReplaySuite application performed as an Internet browser, displaying dynamically generated HTML documents, which constituted the ReplaySuite Graphical User Interface (GUI). When executed, the ReplaySuite application displayed a PHP file (*login.php*) enabling users to log in using their unique username and password. This information was submitted to the server via a HTML form and passed to a PHP file (*logon.php*), which interrogated the VPS database with the login information, via SQL query. If login was successful, the main ReplaySuite GUI was loaded, comprised of six HTML documents displayed on screen together. In order to display six documents at once, the ReplaySuite used multiple 'frames' within a single window. Each frame held a separate webpage, creating a flexible means of displaying information from a multitude of sources, in one webpage.

The ReplaySuite enabled users to view lists of VPS examinations, view diagnostic concordance data and replay previously performed examinations. Users could navigate between these ReplaySuite functions via onscreen hyperlinks. Figure 2.3 illustrates the system architecture of the ReplaySuite and how users could navigate between the different functionality.

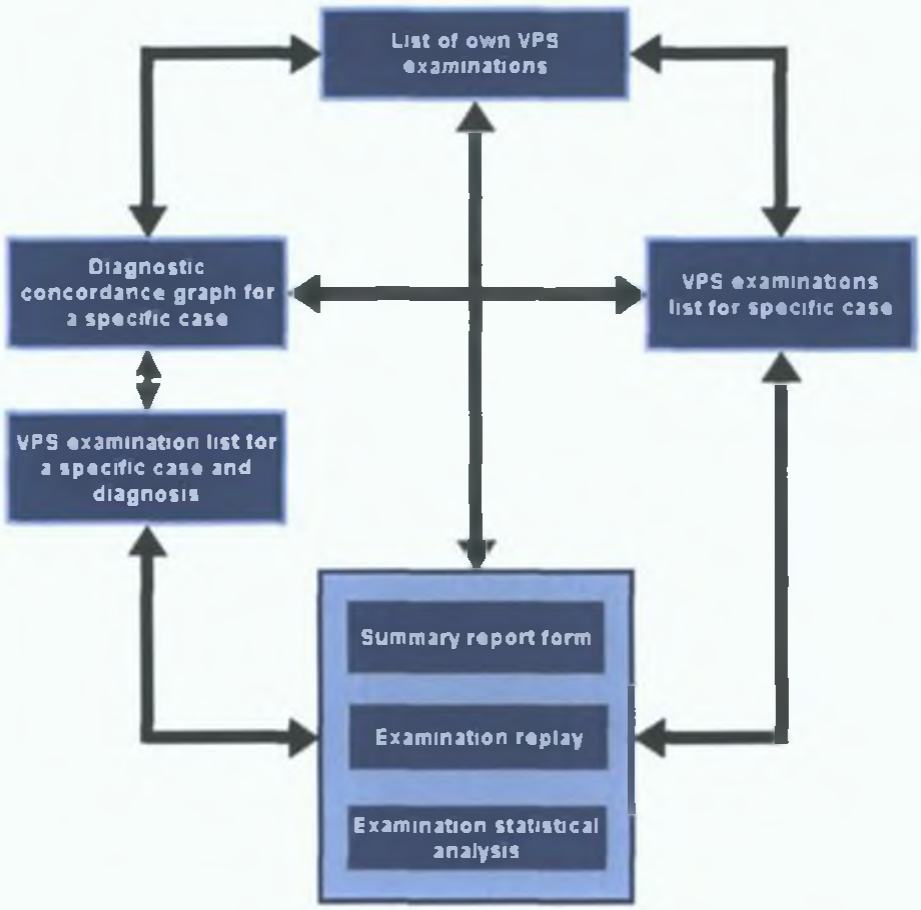


Figure 2.3 Entity relationship diagram illustrating ReplaySuite 1.0 functionality and methods of navigating between functions, via hyperlinks.

2.3.1 Examination Lists

VPS examination lists displayed summarised information about VPS examinations. Users could view a list of their own examinations, those for a specific case or for a specific case and diagnosis. Users with administrator access could also view lists of examinations performed by specific users and examinations with a specific final diagnosis. Entire list of all examinations of a specific slide could be viewed by clicking the relevant slide number in the *Slide* column. When first loaded, the ReplaySuite displayed the users own examinations. The two columns to the right of the list, labelled *Statistics* and *Replay,* contained hyperlinks that permitted navigation to diagnostic concordance graphs and replay examinations. Figure 2.1 presents a VPS examination list, as displayed by the ReplaySuite.

Case history information and slide overviews were displayed by moving the mouse over the corresponding list element. Examination lists displayed the following information:

Table 2.1 Data displayed by the ReplaySuite 1.0 on examinations lists for each examination.

User ID	**Examining pathologists unique VPS ID Number**
Slide:	Slide examined
Session:	Diagnosis submitted at the end of the examination
Group:	Most common diagnosis for this slide
%Consensus:	The percentage of the group who concur with this diagnosis

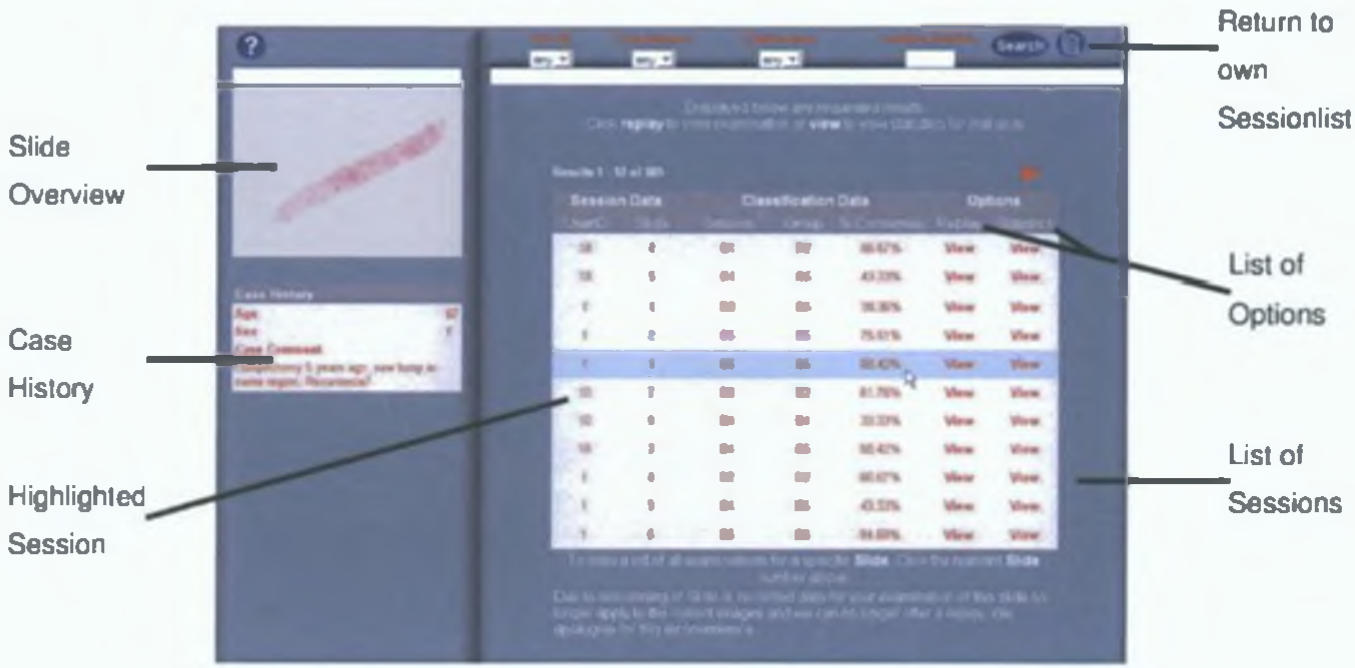


Figure 2.4 Reviewing lists of VPS examinations using the ReplaySuite 1.0.

2.3.2 Classification Concordance Graphs

The classification system displayed was an adaptation of the Core Biopsy Reporting Guidelines for Non-operative Diagnostic Procedures and Reporting in Breast Cancer screening as used by the British national coordinating committee for breast screening pathology. This was hard-coded into the ReplaySuite to display classification data from previous work (Costello, 2004), and these classifications are described in greater detail in Chapter 5.

Classification concordance graphs could be requested by clicking the appropriate hyperlink in the VPS examination list *Statistics* column. For the selected case, the ReplaySuite displayed, in graphical and numerical form, the range and frequency of classifications for all examinations. The top graph represented variation in classification by all examiners, while the second graph displayed the variation in sub-classification for diagnosis of B5. The percentage and number of examining pathologists who concluded each diagnosis was displayed to the right of the chart. Each bar in the graph also acted as a hyperlink to an examination list of all sessions for that slide that concluded the diagnosis represented by that bar (classification or sub-classification). Figure 2.5 below illustrates the graphical representation of classification variation displayed by the ReplaySuite.

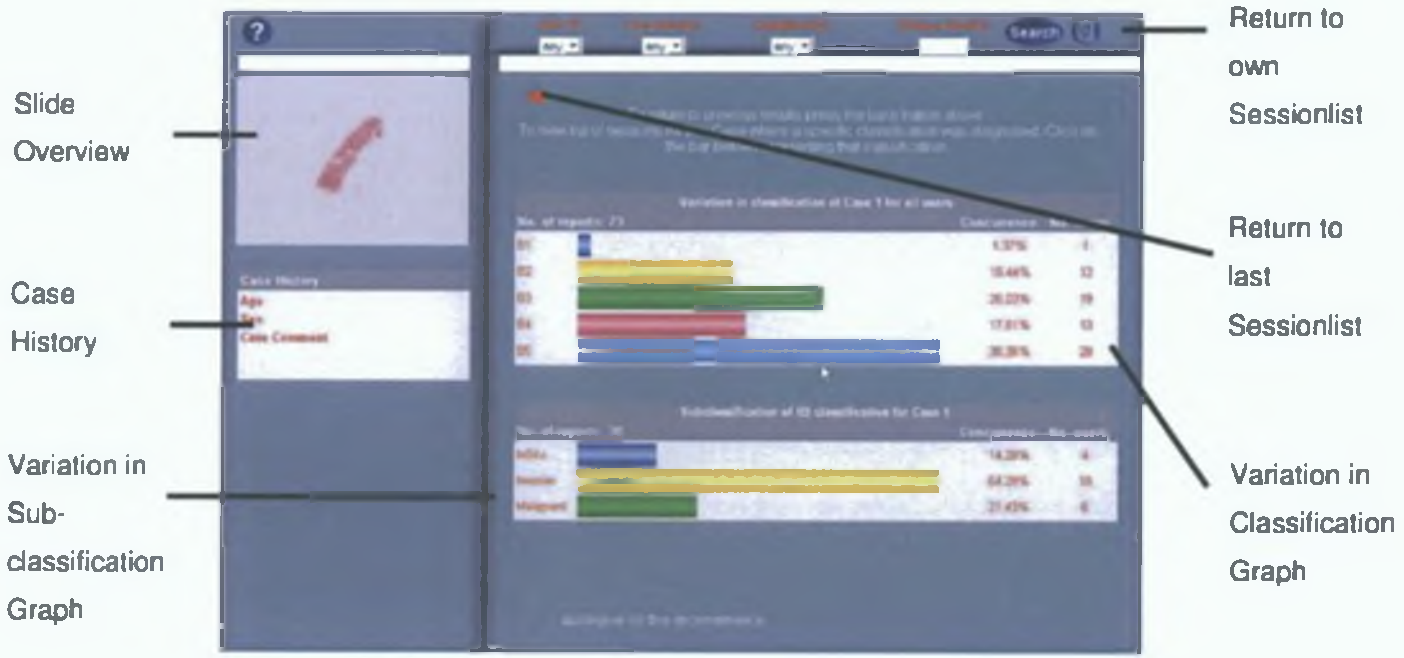


Figure 2.5 VPS group concordance graphs for Case 1, as displayed by the ReplaySuite 1.0.

2.3.3 Replaying an Examination

Examination replays displayed each field of view observed by the examining pathologist, for the same duration and in the same chronological order. Comments pertaining to fields of view were displayed in tandem, and examination replays could be paused, rewound and fast-forwarded. Summary Report forms completed subsequent to examination were displayed prior to replays, and statistical analysis of examinations displayed subsequently.

VPS Examinations could be replayed by clicking the appropriate hyperlink below *replay* on a VPS examination list. The user was requested to select the location of slide images: Web or CD. By selecting Web, the root directory from which to retrieve images would be assigned to the remote server. By selecting CD, images were obtained directly from the CD in the client PC's CD-ROM drive. Delivery via CD-ROM was considered faster for users with limited Internet bandwidth.

Examination replay functionality was delivered in the three frames that comprised the left-sided control panel. The information panel displayed the slide overview image, enabling the user to locate the main field of view on the slide. Current field of view magnification, location in the chronological order of the examination and remaining time onscreen were displayed in the replay information box. Additionally it indicated the next type of move (zoom in/out, lateral motion) in the replay. Comments pertaining to onscreen field of view by the examining pathologist were displayed in the comment box.

Control panel icons activated JavaScript functions that controlled presentation of the replay, enabling the user to play, pause, rewind and fast-forward through the chronological order of fields of view. Additional icons permitted the user to return to their own or the last viewed examination list. The right panel contained the navigation panel that provided information on how to utilise the replay, and the main field of view, as viewed by the examining pathologist. This consisted of 1, 4 or 16 tiled images, depending on the magnification used.

During a replay, the *seconds to next move* counter in the information panel counted down the duration of time the examining pathologist viewed the onscreen field of view, with the next field of view in the sequence displayed at the end of the countdown. Prior to an increase in magnification, a *'zoom in'* box would appear on the main field of view to indicate the area into which the examining pathologist zoomed. Figure 2.1 displays an example examination replay.

Once loaded, the replay would begin. Fields viewed by the examining pathologist were displayed; in chronological order and for the duration they were originally examined. At lowest magnification → 500x 16 tiled images were used to display the main field of view. At 1000x only 4 were used, and at 2000x (maximum magnification) only one image tile was used.

2.3.4 Summary Report Forms

Once loaded, the ReplaySuite displayed the summary report form, an adaptation of the Core Biopsy Reporting Guidelines used in previous work to evaluate the VPS (Costello, 2004). At the time of development of the ReplaySuite, the data from this study was the only examination data available, resulting in its inclusion in this work.

Submitted as an online report by the examining pathologist after an examination, it provided a summary of the conclusions made by the examining pathologist about the case and their experience using the VPS. In addition, it displayed information about the examination session (time and date), case history and slide overview. Figure 2.7 below displays a sample summary report form. Clicking the play button on the control panel removed the summary report form and began the replay.

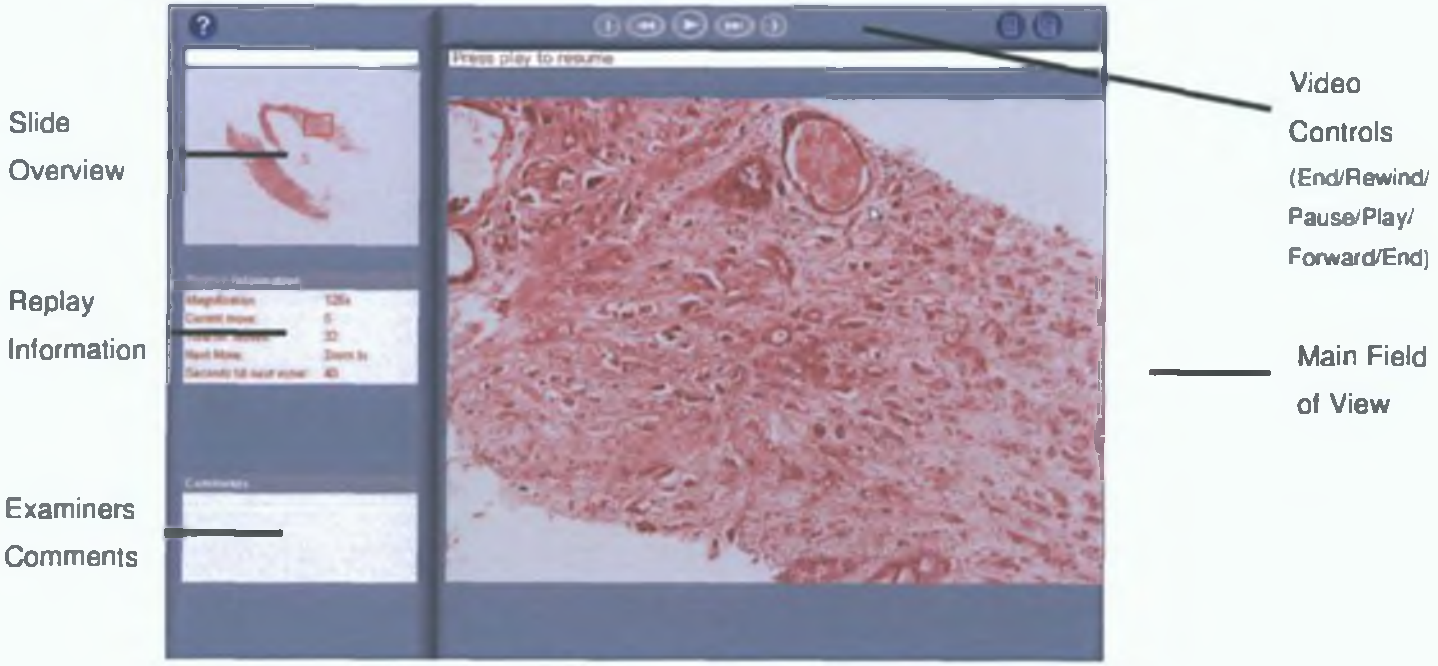


Figure 2.6 Replaying a VPS examination of Slide 2 using the ReplaySuite 1.0.

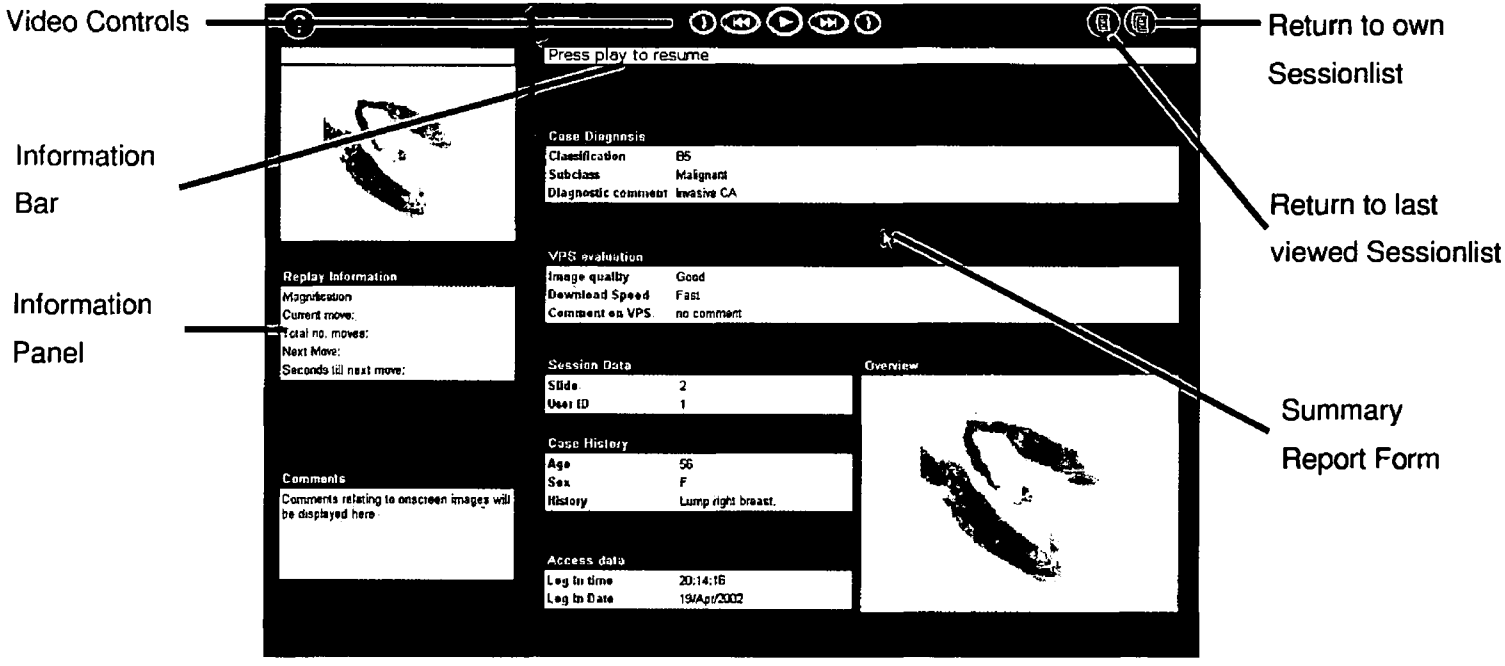


Figure 2.7 VPS Summary Report Form submitted for Case 2 by User 1, as displayed by the ReplaySuite 1.0.

2.3.5 Examination Statistical Analysis

Storing data that describes the diagnostic trace for a VPS examination in a structured manner permitted pertinent statistical information retrieval. Figure 2.8 displays the statistical analysis of an examination. Analysis was confined the following:

- Type and number of zoom and lateral moves performed during the examination.
- Magnification level information
- Area viewed (mm^2)
- Number of fields of view examined
- Total time spent
- Average time spent
- PC specifications of the examining pathologist, specifically the Operating System (OS), Browser, screen configuration (pixels) used during examination and colour size (bits).

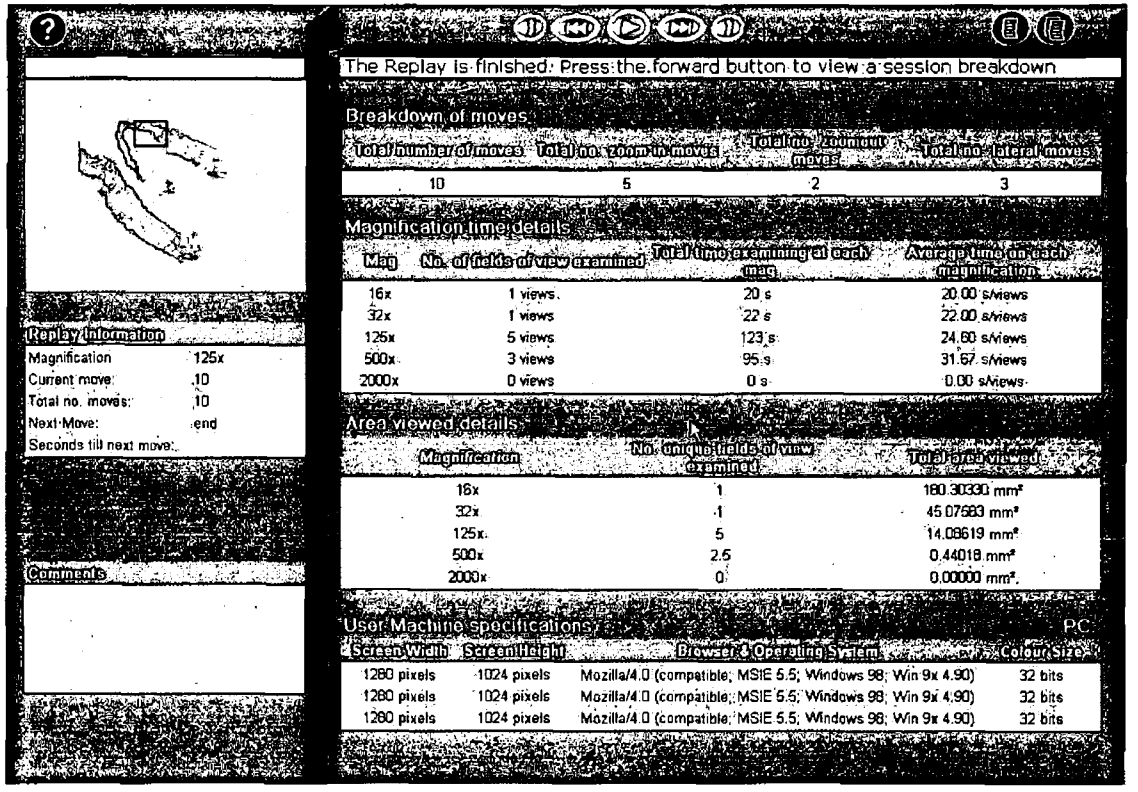


Figure 2.8 Statistical analysis of an examination of Slide 2, as displayed ReplaySuite 1.0

2.4 Conclusion

The ReplaySuite (version 1.0) illustrated that it was technically feasible to utilise VPS tracking data to create a virtual double-headed microscope that removed the temporal and spatial issues associated with conventional double-headed microscopy. This built on previous work (Costello, 2004), facilitating rapid evaluation of diagnostic technique in a manner that was easy to use, intuitive and reproducible. Fields examined were no longer required to be manually located from tracking coordinates, and were presented in the order examined, creating a more accurate representation of a double-headed microscope environment.

Obtaining images for static image tutorials (Kronz *et al.*, 2000, Jones *et al.*, 2002, Hamza *et al.*, 2001) is time consuming; the pathologists is required to have a camera attached to his/her microscope. For each field that must be obtained the following protocol must be adhered to:

- Acquire image from the camera-mounted microscope
- Save image file to the local hard drive
- Use a naming convention that describes the image
- Transfer the image file to a location (remote server, CD-ROM, E-mail account) from where it can be accessed by the recipient

This protracted process is not very conducive to the rapid creation of extensive tutorials, meaning that only small numbers of fields are usually acquired. In contrast, the time required to create examination replays was short (time required to examine a digital slide), and the acquisition process un-intrusive. In order to access the images, the recipient was only required to use the ReplaySuite to search for the referring pathologists User ID number and examined Slide ID to locate the examination.

Once recorded, the examining pathologist's presence was no longer required. Archived examinations could be replayed anywhere, anytime from an internet-enabled PC without placing any additional time participation requirements on the examining pathologist. The use of web-based delivery methodologies was key to this; the use of open-source PHP of pivotal importance, as it facilitated both the retrieval of pertinent data and images from the database, but also dynamically generated HTML to construct the Graphical User Interface of the ReplaySuite.

The most significant limitation of the technology pertained to the amount of time required to load an examination replay. Internet bandwidth can fluctuate, even over short periods of time. Streamed images to the ReplaySuite during replays could be delayed, interrupting the replay process and creating an unfavourable experience. It was concluded that, given the size of image data required to display an examination replay, it would be more advantageous to load all images into the PC cache, prior to commencement. This facilitated smooth replay presentation without interruptions, but required a significant amount of time to 'preload' all the required images. However, as broadband becomes more widely embraced such issues may be resolved, and the utilisation of streaming ultimately incorporated.

Regardless of this issue, the ReplaySuite accurately replicates the double-headed microscope environment, allowing a remote observer to view a virtual slide examination performed in a different location, and at a different time. It uses a novel approach that has not been previously exploited to facilitate this, using open-source software that may be accessed by anyone with an Internet-enabled PC.

Chapter 3: Redevelopment Of The Slide Scanning System

3.1 Introduction

The concepts of static and dynamic telepathology are now firmly established within the consciousness of pathology. However hybrids or digital slides are a relatively new concept who's viability as a practical solution for remote pathology has improved substantially, mainly due to developments in web-based delivery systems, improved bandwidth, the evolution of increasingly specialised hardware and the reduction in computer storage space. Digital slides are quickly becoming a more viable alternative in a range of applications to both static systems, as they are free of sampling bias, and dynamic telemicroscopy, as they are simpler and easier to use.

Many slide scanning systems are constructed from 'off the shelf' components, with software also developed in-house specifically for slide digitisation. Typically, these systems will acquire a large number of image 'tiles' that cover the entire tissue section, and then use image pyramid algorithms to resize images whilst maintaining spatial resolution. This creates multiple resolution layers for a single slide, which allows tissue to be examined at a range of magnifications. Saeger *et al* (2003) developed such a system that used an application written in Visual C++ to control a motorised stage. Designed to digitise slides for examination with their DVM (Digital Virtual Microscope) slide viewer system, it scanned tissue in meandering lines, subsequently de-fragmenting and compressing images into Enhanced Compressed Wavelet (ECW) format using commercial 'ER Mapper' software (Earth Resource Mapping Incorporated©). This system created digital slides that averaged 380Mb (megabytes) in size with a 1:15 compression ratio (5.5Gb uncompressed), small enough to store a single slide on a CD-ROM, or multiple slides on a DVD.

Glatz-Krieger *et al* (2005, 2003) suggested that it required 9 hours to digitise a 4 x 2.5cm ($10cm^2$) slide using a 40x objective, based on an acquisition speed of 1 image per second. Demichelis *et al* (2002) at the University of Udine, Italy reported even longer digitisation times, with a significantly smaller section of tissue of 1.5 x 1.5cm ($2.25cm^2$) requiring just under 7 hours to digitise. Leong and McGee (2001) claimed digitisation times of 70 minutes for a small 1cm x

1cm tissue section, however this was utilised a 20x objective, requiring fewer fields to be digitised. Such digitisation times make systems of limited use for anything other than low-throughput scanning applications, as some observers claim that an estimated scanning time of 1min/slide is required to handle the daily throughput of an average pathology institute (Saeger *et al.*, 2003).

Many systems utilise pre-scanning capabilities, which involve digitising the section at low magnification to identify the location of tissue on the slide. This significantly reduces the time required to digitise slides, especially when dealing with small amounts of tissue, such as needlecore biopsies. While digitisation speed is an important parameter to consider when assessing scanning systems, image quality is generally considered of greater significance, and a number of issues are pivotal in ensuring images of optimised quality are obtained.

The inherent variability of tissue thickness means the image can quickly go out of focus without adjustment, even over small distances. To compensate for this, scanning systems incorporate autofocus capabilities to ensure that they are in the correct focal plane while scanning. Two approaches have been identified; (1) auto-focusing at every field, which ensures that each image is in focus but significantly increases the time required to digitise a slide, and (2) incorporating a map of reference focal points. This involves randomly or intentionally selecting a series of points on the tissue section at which to obtain a focal value. These values are then used to generate a map of reference points, which are used during scanning to interpolate a focal value for the current position. Demichelis *et al* (2002) commented that when using approach 1, focusing was the most time-consuming task during the digitisation process. They reported acquisition times of 3 seconds per image (stage movement, auto-focus, acquiring and image storage) when using their system, a considerable length of time when it is considered that even relatively small sections require in the order of >10,000 fields to digitise.

The quality of objective used is often overlooked as a factor that affects image quality in slide digitisation. The resolving power of an objective is the most critical issue and influences the ability to distinguish between fine details of a

particular specimen. The primary factor in determining resolution is the Numerical Aperture (NA), which is a measure of an objectives ability to gather light and resolve fine specimen detail at a fixed object distance. NA is calculated by multiplying the angular aperture (which dictates the size and shape of the illumination cone) and the refractive index (measure of how much light is changed when travelling through a material) (Spring and Davidson, 2005). Oil has a higher refractive index than air meaning oil immersion lenses are of higher refractive index that conventional objectives. Higher NA objectives are often better corrected for chromatic aberration (white light splitting into its constituent colours along the optical axis of a lens) and spherical aberration (deviation of light from the focal point along the axis of a lens), but have lower depth of field (z-range in which objects are in focus).

The capability of the image-capturing device to adequately obtain a suitable image is an important consideration in slide digitisation, as a system that is able to produce excellent quality images but not acquire them at sufficient quality is of little benefit. A Charge-Coupled Device (CCD) is a light-sensitive chip or image sensor used in digital cameras to convert light into proportional (analogue) electrical current. When a picture is taken, the CCD is struck by light coming through the camera's lens. Each of the thousands or millions of tiny pixels that make up the CCD converts this light into electrons. The number of electrons, usually described as the pixel's accumulated charge, is measured, and then converted to a digital value via an analogue-to-digital converter. The resolution of a CCD (measured in megapixels) determines the size of image that can be captured, and is a function of the number of photodiodes, and their size relative to the image projected onto the chip's surface by the microscope optics (Viosport, 2004, Flynn and Davidson, 2005). Glatz-Krieger *et al* (2003) stated that a 40x objective with 0.95 NA requires a 1.3 megapixel CCD resolution to acquire the maximum optical data. She stated that a common misconception is that by increasing the CCD resolution beyond this leads to more information and better quality images, but that this contradicts physical laws and results in larger (redundant) datasets without any improvement in image quality.

The size of an image file is dependant on two factors, the size of the image in pixels (resolution), and the amount of information used to describe each pixel (colour depth). The continuing reduction in the cost of computer memory has made the size of digital slides a less contentious issue. However many virtual slide systems are web-based and require images to be transmitted over the Internet. Bandwidth varies, and larger files take longer to transmit, so any reduction in image size that does not impair quality is of benefit. This can be achieved by image compression, a process that reduces the amount of memory required to store an image. There are two distinct categories of image compression: Lossless and Lossy. Lossless compression methods are fully reversible, but are typically confined to compression ratios of the order of 1:3. With lossy compression, the retrieved file can be quite different to the original at the bit level, while being indistinguishable to the human eye for most practical purposes (of significance in image analysis). Higher Lossy compression rates are also possible but result in visibly degraded images, with level of acceptable degradation usually the limiting factor. In attempting to establish a set of guidelines on achievable compression ratios in telepathology systems, Foran *et al* (1997) invited 5 pathologists to assess a series of digital images compressed at increasing rates using the JPEG (Joint Photographic Experts Group) technique. They concluded that images could be compressed 20-35% without compromising clinical usefulness, depending on the type of image. Another interesting study attempted to establish whether pathologists could differentiate between images of different colour depth. While it is generally regarded that 24-bit or 'true colour' images are required in pathology images, Doolittle (1997) concluded that 8-bit images were of sufficient colour depth for diagnosis, and observed that some participating pathologists actually preferring the crispness of 8-bit images.

Digital slides may be saved in a number of formats, depending on the scanning system used and how the images are displayed. FlashPix (EastMan Kodax) is an image file format that also allows multiple resolutions of the same image to be stored within the same file. The advantages of such a format in telepathology are obvious, as it allows layers representing tissue at different magnifications to be stored together. Dee el al (2003) used FlashPix to provide digital slides in the

annual pathobiology of cancer workshop laboratory in the University of Iowa, while other systems such as vMic (Glatz-Krieger *et al.*, 2003, 2005) and SlideTutor (Crowley *et al.*, 2003) also use this format. Aperio scanning systems create multi-layered tiff (Tagged Image File Format) files that can examined both with Aperio's own slide viewer, ImageScope, or Zoomify, a file format with similar functionality to FlashPix (Zoomify Inc., 2004). Other tile based scanning systems often save images as JPEG's, an industry standard file format, and develop customised viewers to display images in a format that allows navigation between images representing tissue at different magnification.

3.2 Evaluation of Previous Work

The objective of the following chapter was to redevelop the scanning algorithms that were used to digitise glass slides. The primary scanning algorithms developed in previous work involved 2 phases; a raster scan phase in which a 180mm^2 area was scanned, irrespective of the tissue size, and a layer building phase (Costello, 2004). These algorithms were simple and effective, but not designed for high throughput use. The original scanning process required up to 36 hours for digitising and publishing of a single virtual slide. In order to digitise the quantity of slides required for the work described in this thesis, new scanning algorithms were required to enable higher scanning throughput (1 slide/4 hours). These new algorithms addressed design flaws of the originals algorithms, improving efficiency, reliability and flexibility.

3.3 Hardware and Software

The VPS imaging workstation consisted of an Olympus BX-40 microscope incorporating a 40x Plan Apochromat lens with a 0.95 numerical aperture. The microscope was fitted with a Prior robotic stage and a JVC 3-CCD (3-chip charge-coupled device) video camera. VPS slides were created using in-house developed scanning algorithms, developed with Optimas 6.5 image acquisition and analysis software (Media Cybernetics). Optimas was capable of interfacing with, and acquiring images from a framegrabber, which can subsequently be

manipulated using Optimas's programmable platform. Control of RS232 interfaced devices, such as a robotic stage via ALI (Analytics Language for Images) enables algorithms to perform a programmed raster scan of a tissue section. Additional algorithms were then used to build subsequent layers of lesser magnification, using these images.

3.4 Redevelopment of Scanning Algorithms

The primary consideration in redevelopment of scanning algorithms was to streamline the process and remove unnecessary components. The area of glass slide containing the tissue section varies considerably, depending on the type of specimen. Large biopsies and cytological smears can be several centimetres in length, while needlecore biopsies can be as small as a few millimetres. As a result, the bulk of a slides area is often extraneous space that does not contain viable tissue. Scanning empty sections is time consuming and inefficient, however the size and geometric positioning of tissue can vary greatly, preventing the differentiation of tissue and non-tissue areas by a standardised scanning procedure.

In order to enable high throughput scanning, redeveloped scanning algorithms required a number of key improvements:

- Improved scanning efficiency (and reduction in scanning times) by incorporation of a PreScan
- Capability to scan up to 4 slides in one session.
- Reduced publication times. Incorporation of parallel FTP (File Transfer Protocol) capabilities to enable automated publication of images (transfer them onto the remote server) during scanning.
- Increase the potential area that can be scanned from 180 mm^2 (128x128 fields at 40x) to 720mm^2 (256x256 fields at 40x).
- Publish additional magnification layers. This was in response to VPS evaluation study participants (Costello *et al.*, 2003) commenting that

additional magnification layers between existing layers would be beneficial.

- Incorporation of restart capabilities. Due to the extensive workload, the Imaging workstation was predisposed to crash periodically. Any PreScan data pertaining to the location of tissue on the slide needed to be written to the hard drive in case of a system crash. Additional functionality would then be able to utilise this data to continue scanning from the point it crashed.
- Improved automation. Reduction of human interaction required to digitise and publish images to 2 events; Initiation of the digitisation process, and changing the objective when required.
- Slide validation capabilities. Development of algorithms for checking image quality and allowing areas to be rescanned.
- Flexibility to scan at either 20x or 40x

A viable solution is to incorporate the ability to discriminate between tissue and non-tissue fields by providing a 'map' of X,Y coordinates indicating the locations of fields containing tissue. This is feasible by conducting a PreScan of the slide at low magnification. A single 4x field of view corresponds to a 10x10 matrix of 40x fields of view. The presence/absence of tissue in each of these 40x fields is determined via image analysis and the X,Y coordinates of those fields containing tissue stored. These coordinates are then used during scanning at higher magnification (e.g. 40x). Only fields containing tissue are digitised as a result, reducing the total time spent scanning.

3.4.1 Algorithm Architecture

The Scanning algorithms were redeveloped to provide a more robust, flexible scanning system with increased throughput, reduced scanning times, larger slide areas scanned and a greater number of magnification layers published. The redeveloped process may be separated into 6 distinct modules:

- Data Initialisation

- PreScan
- MainScan
- Layer Building
- Image validation
- Parallel File Transfer Protocol

From the remainder of this section on slide scanning algorithms, any utilisation of the following format: *Acquire(), GreytoBinary(),* refer to Optimas defined ALI (Analytical Language for Images) functions. A more detailed description of ALI is provided in the Appendices.

3.4.1.1 Data Initialisation

The Scanning algorithm architecture is designed to be flexible, enabling 1-4 slides to be scanned at once, at either 20x or 40x. When the Scan macro is run, the user is required to input a number of parameters, requested sequentially. These are:

- Magnification at which to scan.
- Number of slides to be scanned
- Identification numbers of slides

Once these parameters have been inputted, a number of operations are performed. First, the amount of free space on the Hard Drive is checked to ensure there is enough space to store all the images. If there is enough space, file directories are created for each slide as illustrated in Figure 3.1.

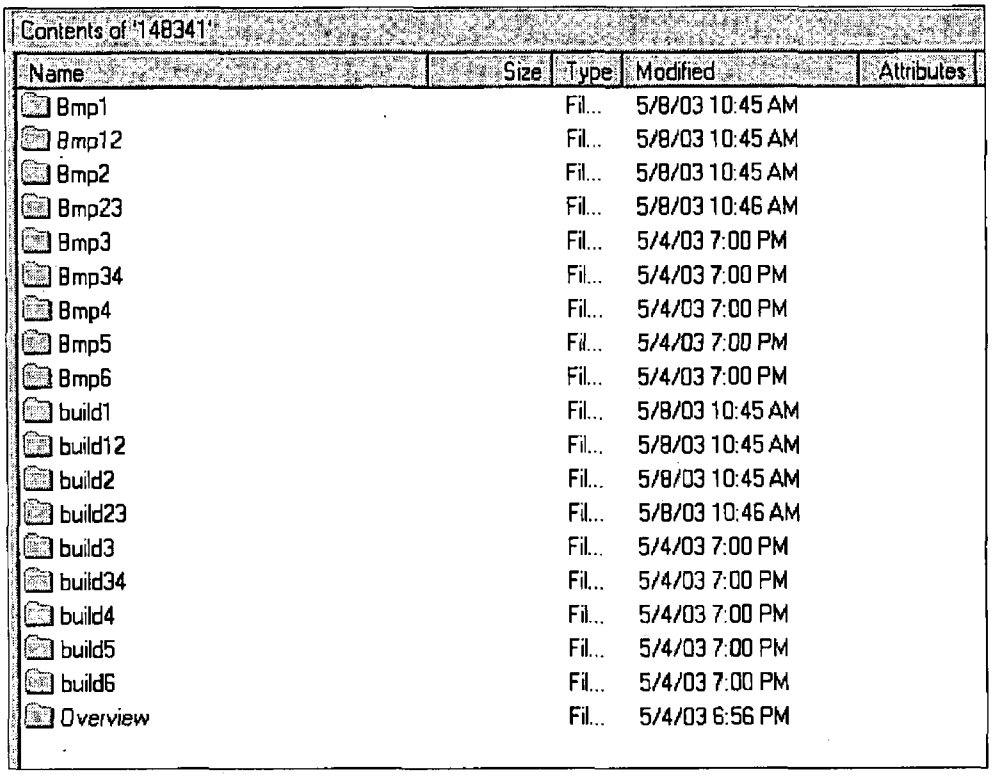


Figure 3.1 Creation of a slide file directory on the local drive to store bitmap (bmp) and JPEG (build) image files.

When scanning at 40x, 19 (9 bitmap, 9 jpeg & 1 overview) subdirectories are created. JPEG images for each magnification layer are stored in *'Build'* folders. These image tiles are transferred to the remote server and used by the VPS to represent the digitised slide. Bitmap images, used to create these JPEG's are stored in *'Bmp'* folders. The overview folder stores the slide overview bitmap.

Two data files (.dat) are created on the hard drive for each slide. The first file is used to record the stage X,Y coordinates of the centre of the area to be scanned, and the second to store the X, Y coordinates of all slide fields containing tissue. If the workstation crashes, this data is used continue the scan from the point it crashed.

3.4.1.2 Stage and PreScan Calibration

In order to access images taken by the framegrabber, a new image file is opened. Two Optimas functions are used by the program during scanning; *Acquire()* to cause the framegrabber to begin acquiring frames, *Freeze()* to stop image acquisition.

The stage is calibrated by providing a 'zero' value. The user moves the objective over the furthest top left area of the stage. The X,Y stage coordinates of this point are set as 0,0 and all subsequent X,Y coordinates are referenced from this point.

The depth/height of tissue can differ significantly between slides, requiring a focal depth reference point for each slide. The user is asked to focus on each slide sequentially and the focal height recorded. Prior to scanning each slide, the focal height is adjusted to the appropriate reference height. As biopsies may be fixed anywhere on the slide, the user additionally marks the centre of the area containing the tissue of interest on each slide.

The objective is automatically moved over an empty area of slide and the user asked to adjust the lighting level. The lighting level is checked until within an optimum range for image acquisition. Once obtained, an image of the blank area is acquired and a series of convolve filters applied to remove any visible debris. A convolution is performed within a region by taking the summation of the kernel weighted grayscale values of each pixel and its neighbours and applying it to all pixels within the region. The 'Blank' image is subsequently stored in the Optimas buffer. An inherent property of light microscopy images is an uneven distribution of illumination between the centre of the image and the outer edges, a phenomenon known as vignetting (Leong, 2002). This presents difficulty when the intention is to tile and stitch a number of images together as the uneven distribution of light across each image presents a non-seamless effect. In order to neutralise this effect, the variation in illumination across an image may be normalised by dividing acquired images by the 'Blank' image.

3.4.1.3 PreScan

Once the scan data has been initialised and the stage calibrated, a new image object is created. This is the Raster Image, to which images taken during the PreScan are added to construct an overview of the Slide. Using the centre point of the scan area recorded during data initialisation as a reference, the stage moves the objective to the centre of the top left field of view. This is the starting point for the PreScan. A grid of 4x fields of view containing the area to be scanned at higher magnification will be acquired. Each time the stage moves to the next field of view, the previous image is added to the Raster Image and the original image processed to extract the X,Y coordinates of any higher magnification fields containing tissue. The area was scanned via a snake pattern as illustrated in Figure 3.2.

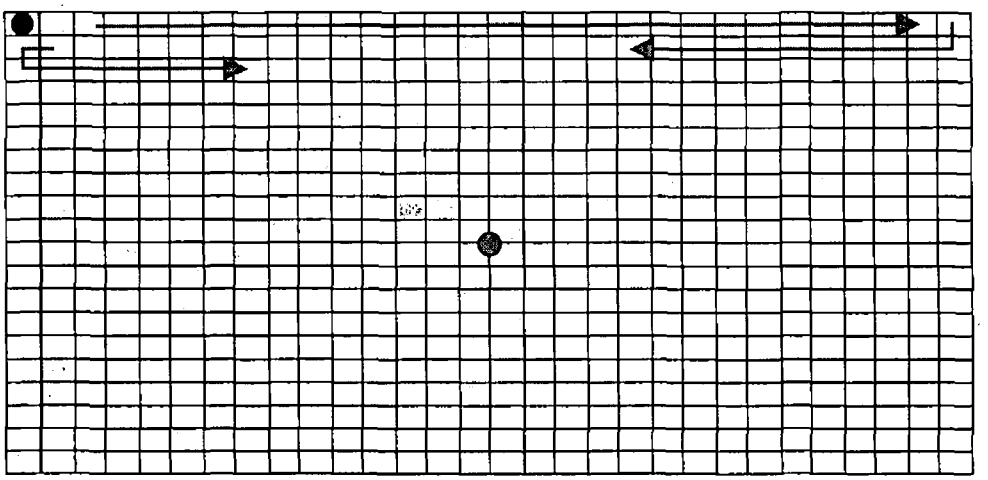


Figure 3.2 Snake pattern used to create PreScan image at 4x during slide scanning.

For each field, a framegrabber image is acquired and frozen, then checked to ensure no framegrabber errors have resulted in an unusable image. Parallel processing capabilities of the workstation enable the image to be processed and added to the Raster image while the stage is moving to the next field of view,

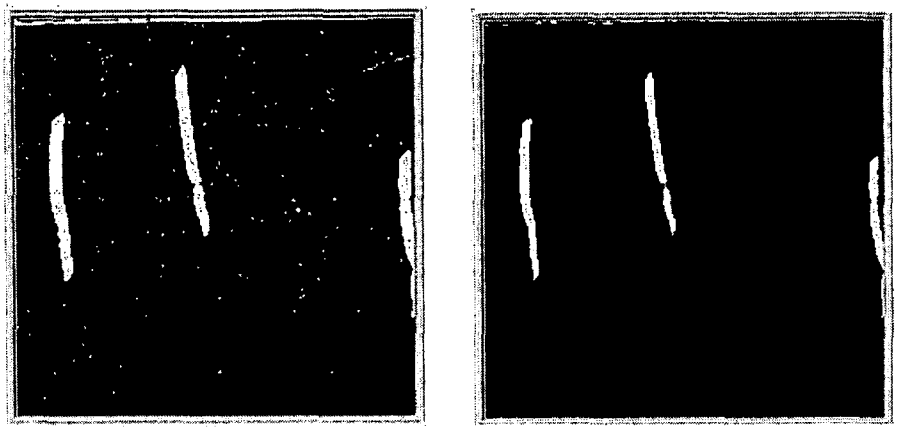
improving efficiency. Movement is achieved by sending stage coordinates and a stage command to the open RS232 communications port.

In parallel, a copy of the 4x image is added to the appropriate region on the Raster Image, creating the Overview image tile by tile. Green is then selected as the band of interest to process the original image as this provides better contrast between tissue and background than other bands. Band of interest refers to the colour plane selected; only green plane pixel data will be affected. The Optimas ALI function *GreyToBinary()* segments the grayscale image into foreground and background components based on a preset threshold range. The original image is now represented as a binarized image, tissue as white foreground and non-tissue as black background. Regions of the 4x image corresponding to 40x/20x fields of view are then processed. One 4x field represents a 10x10 matrix of 40x fields of view. The percentage of tissue within each region is calculated. If above a predefined limit the X,Y coordinates of the higher magnification field are recorded in a 2 dimensional array. Once the PreScan has been completed, this array is written to data (.dat) file.

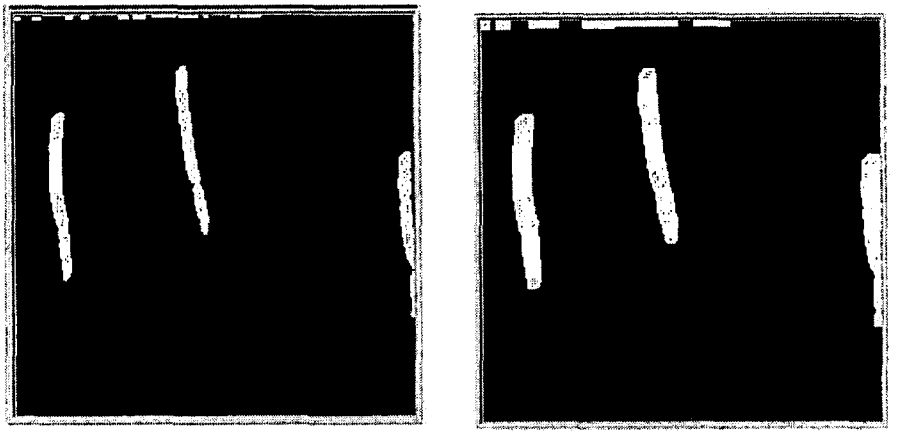
The data file contains an array of all fields on the slide containing tissue. As well as containing large sections of tissue, the array also contains debris, small pieces of dirt and extraneous artefacts that are not to be digitised. Up to 50% of all fields considered to contain viable tissue may actually be debris. A means of differentiating between debris and viable tissue is essential in order to maintain efficiency. For this the array is converted into a 2 dimensional bitmap. The bitmap can then be processed using image analysis to automatically filter out any unwanted areas. Additionally it permits the array to be edited by the user, allowing areas to be added or removed before the main scan commences.

This process is illustrated in Figure 3.3. A bitmap is created to represent the scanned area, in which each higher magnification field of view is represented by a single pixel. When the array has been added to the bitmap any field containing tissue/debris is indicated by a white pixel. Black pixels indicate areas with no tissue/debris. Figure 3.3 (a) shows an unedited array bitmap, illustrating the large number of small areas of debris picked up by the PreScan. An Erode filter

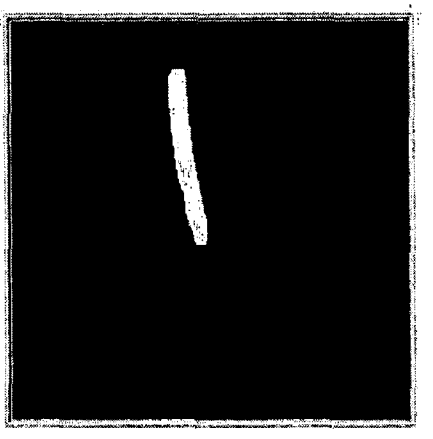
(a common filter in imaging applications) is applied to the bitmap to remove these regions, by eliminating foreground pixels (white) that are 8-connected to a background pixel (black). The Erode filter is applied several times to result in an image as shown in (b) and (c). In addition to removing debris, the erode filter removes fields along the edge of tissue sections, which must be replaced. This is achieved using a Dilate filter (also a common image filter), which effectively is the opposite of Erode. The dilate operation identifies background pixels (black) that are 8-connected to a foreground pixel (white) on the binary image, and changes them to foreground. The result is an image similar to that shown in (d). The Dilate filter is applied several times to ensure that all required tissue fields that may have been removed by the Erode filter are replaced. Finally, manual intervention is allowed to remove any sections that were not removed by the Erode filter application, but which are not required. This allows the user to select areas with the mouse to add/delete. Figure 3.3 (e) shows the same array after the user has added and removed regions.



(a) (b)



(c) (d)



(e)

Figure 3.3 Filtering a PreScan tissue array using image analysis to remove debris and tissue from a visual a representation of the array (a) Pre-filtered tissue array (b) Eroded array (c) Dilated array (d) Second iteration of a dilation filter (e) Final tissue array after manual removal of redundant tissue areas.

3.4.1.4 MainScan

Prior to the MainScan, a number of adjustments must be made to the system. Changing the objective from a 4x aperture to a 40x aperture results in a reduction in the amount of light reaching the camera. Lighting levels are re-adjusted to reach the required range. Variations in focal height have a more pronounced effect on focus at 40x than at 4x, therefore a new reference focal height is taken for each slide.

A new 'Blank' image is obtained to normalise lighting variations. The blank also serves a secondary purpose. Only areas containing tissue will be scanned during the MainScan, however when displaying images via the VPS, non-tissue areas must be represented by a background image. Accordingly, the blank is saved to each *'Build'* folder and will be displayed when an area without tissue is requested.

Once the MainScan has begun and the focal height adjusted to the slides reference height, the stage moves from field to field containing tissue acquiring images and saving them to file. Stage coordinates for each point are provided by the PreScan tissue array. To reduce downtime associated with stage movement, each new point to be acquired is the nearest point to the previous point. In order to do this, the tissue array must be sorted, and the nearest geometric point determined. However performing such an operation with such a large array directly in Optimas is not feasible, as ALI is an interpreted language rather than a compiled one. Compiled program code is reduced to a set of machine-specific instructions prior to execution, allowing it to run faster than interpreted ones, which must be reduced to machine instructions at runtime. To circumnavigate this, the array is passed to an in-house developed Dynamic-Linked Library (DLL) which performs the calculation, as illustrated in Figure 3.4. A DLL file allows programs to share code and other resources necessary to perform particular tasks. Written in C++, the DLL selects the nearest point to the current field that contains tissue. The fields X,Y coordinates are returned to Optimas, along with the tissue array (with the previous point removed).

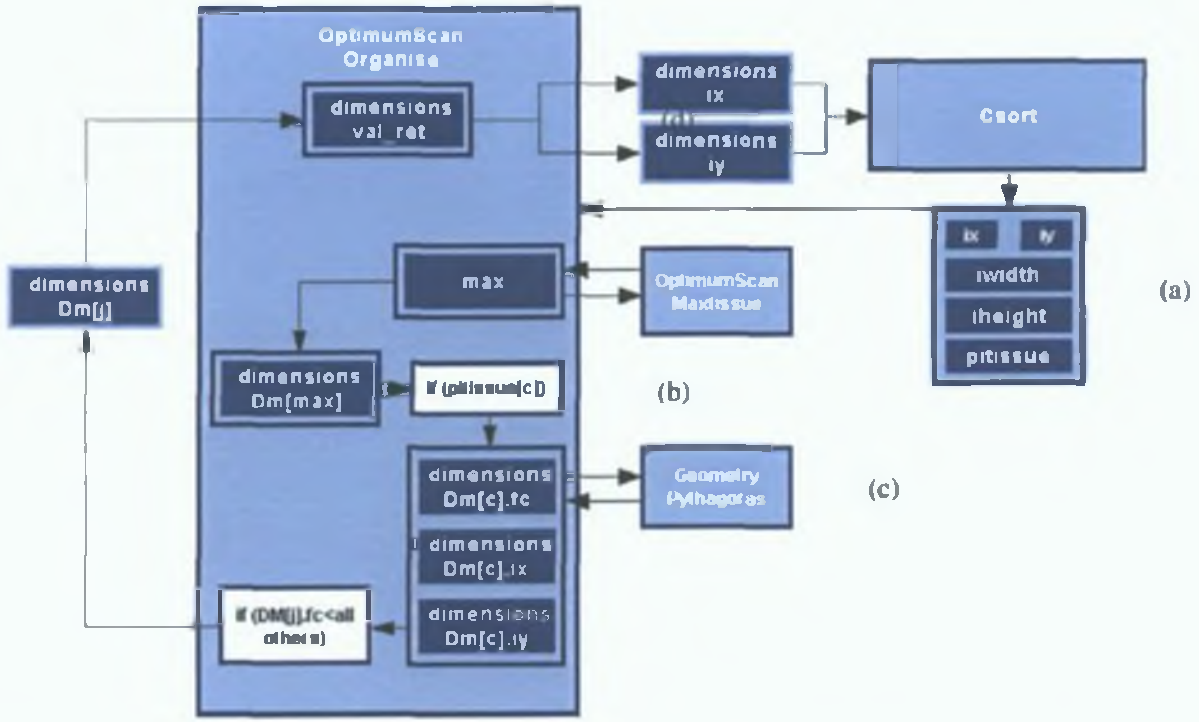


Figure 3.4 Calling a DLL in Optimas to return the nearest unscanned field with tissue. (a) Optimas passes current x,y coordinates to the DLL. If (b) any tissue fields remain unscanned, the Geometry class (c) uses Pythagoras theorem to determine nearest point containing tissue, which is (d) passed back to Optimas.

For each field, a number of operations are performed. These ensure the image is in focus, normalised and that no framegrabber errors are present. An autofocus is performed if it is determined that there is enough tissue onscreen to focus on. This is achieved by passing an Autofocus command to the open communications port via the Optimas *CommWrite()* command to the Prior stage controller. The autofocus range is also set via this command, determining the distance from the current focal height the focus will search to obtain a focused image. The further away, the greater the probability that the view is out of focus and the greater the range of autofocus used.

In order to gauge the progress of the scan, the tissue array bitmap is displayed during the scan. As each field is acquired, the corresponding pixels luminance is modified, resulting in acquired images displayed as grey; fields still to be acquired are indicated in white, as illustrated in Figure 3.5.

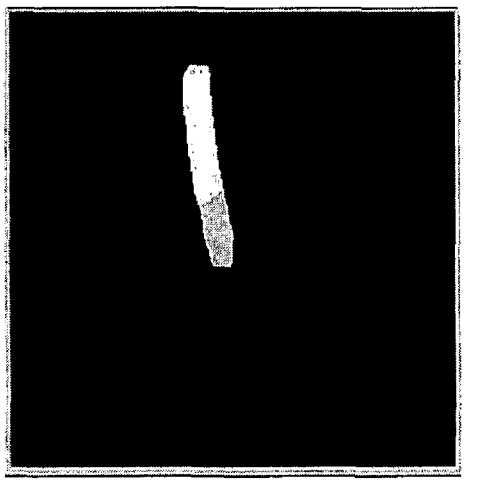


Figure 3.5 Progress of a slide scan, as indicated by the tissue array map.

The Cjpeg.exe application is used to compresses each acquire image to a progressive JPEG file (1998). Images are initially saved as bitmaps, then converted to progressive Jpegs. The **-progressive** switch creates a "progressive JPEG" file. In this type of JPEG file, the data is stored in multiple scans of

increasing quality. The following syntax converts a bitmap (1.bmp) to a progressive jpeg (1.jpg):

cjpeg -progressive 1.bmp 1.jpg

If the file is being transmitted over a slow communications link, the decoder can use the first scan to display a low-quality image very quickly, and can then improve the display with each subsequent scan. The final image is exactly equivalent to a standard JPEG file of the same quality setting, and the total file size is about the same, often a little smaller.

3.4.1.5 Building Layers

When all images have been acquired for each slide, the final module of the slide scanning is activated. This generates layers of image from the acquired images that correspond to geometrically related areas at lower magnification. Four adjacent images (40x) resized and tiled into a new image represent that area at 20x. This process is repeated for the entire base layer (acquired image layer) and then subsequent layers, creating a pyramid of geometrically related images that represent the digitised tissue section at a range of magnifications. The number of images contained within each layer is reduced proportionally with each increase in layer. The scanned area is represented by a potential maximum of 65536 acquired images (displayed onscreen as 1000x) to a potential maximum of 16 images (displayed onscreen as 40x).

The existence of images required to create each new image in the layer above are first checked. If none exist, the new image is not created, as it does not contain tissue. If images do exist, they are opened, resized and copied into the selected region of interest in the new image. A 'blank' jpeg image is created for each layer using the base layer blank. At the end of the creation of each layer, the layer building parameters are adjusted to reflect the reduction in images to be created and images to be used to create them. Each new image is first created as a bitmap then compressed into a progressive JPEG by CJPEG.exe.

3.4.1.6 Image Validation

Even with all the previously mentioned safeguards in place, images are acquired which are not of the required quality, most commonly due to being out of focus during acquisition. A simplified virtual slide viewer is used to review newly created slides, in order to identify sections that require rescanning. Strips, points and rectangular sections may be rescanned, by inputting the appropriate x,y coordinates. Once inputted, MainScan is run, with the original PreScan array replacing the coordinates of sections to be rescanned. The scan process is repeated, new images replacing inadequate field images. Once the scan is finished, the subsequent layers are rebuilt with the new images incorporated.

3.4.1.7 Parallel File Transfer Protocol

Transfer of slide images from the imaging workstation to the server in series is time consuming and an inefficient use of the resources available. The parallel processing capabilities of the imaging workstation enable image publication (transfer to server) to run concurrently with scanning. This is achieved using a DLL developed to transfer images to the server. The DLL is called by Optimas each time a new image is created.

3.5 Conclusion

In-house developed systems adapt equipment for a purpose that they were not specifically designed to do. While they perform admirably, they are incapable of competing with custom-built hardware systems designed for high-throughput. The development of ultra-fast slide scanners by vendors such as Aperio (Aperio Technologies, 2005), Trestle (Trestle Corporation, 2005), Dmetrix (Dmetrix Inc, 2005) and 3Dhistech (3DHistech Ltd, 2002) has altered the landscape of virtual slide creation, as complete automation of the slide digitisation process is quickly becoming a reality. The operational design differs from system to system, but incorporate technology such as nanomotors, linear CCD array sensors and slide loaders. Aperio promote the 'one touch scanning' capabilities of their ScanScope products, although in practice this approach does not produce optimally scanned slides (O'Shea, 2005). Purpose-built systems are capable of rapid digitisation and publication of images, with digitisation times of 5-30 minutes, depending on tissue size and system. DMetrix promise even more rapid digitisation, with Weinstein proposing a Minaturized Microscopy Array (MMA) based system capable of digitising a slide in roughly 1 minute (Weinstein *et al.*, 2001, Weinstein *et al.*, 2004).

While the system described here is unable to compete with such rapid digitisation, it represented a significant improvement on the previous incarnation. The incorporation of a PreScan prevented the digitisation of redundant fields, a feature that has also been incorporated into ultra-fast slide scanners (Aperio Technologies, 2005), somewhat validating the approach. Other features, such as the use of a focal map and the use of a 'blank' to represent all empty fields, are also rapidly becoming standard approaches to address issues in scanning.

While somewhat obsolete due to the development of ultra-fast scanners, it is important to note that this system outperforms similar customised scanning systems, achieving shorter scanning times and superior image quality (Demichelis et al., 2002, Leong and McGee, 2001). Given the limitations of the technology available, it may therefore be considered one of the more optimised motorised stage based systems available.

Chapter 4: Redevelopment Of The VPS And ReplaySuite

4.1 Introduction

Eventually, even the most successful computer application becomes obsolete. Technologies improve, preferences change and once popular software is ultimately replaced or superseded. Real world use also identifies technical deficiencies that may have been overlooked during testing. In response to the ever-changing demands of software users, all major software developers such as Microsoft ((Microsoft corporation, 2004), Apple (Apple Computer Inc, 2005) and Macromedia (Macromedia Inc., 2005) continually adapt and upgrade software.

To avoid releasing multiple copies of revised software to address several minor changes over a short time, developers launch iterative software versions. Each version exhibits a number of modifications that differentiate it from its predecessor, enticing new users whilst encouraging existing users to 'upgrade'. These modifications may improve existing functionality, or add new capabilities, depending on the focus of redevelopment. However, redevelopment of all software technology is conducted to achieve one objective: to better satisfy the needs of users.

Telepathology and virtual slide software are no exceptions to this process. Early non-dynamic telepathology systems were restricted to limited numbers of static images per slide. However the preference of pathologists to examine entire tissue sections led first to static image libraries that displayed multiple fields at different magnifications, then to the availability of entire sections as virtual slides. New versions of virtual slide viewers such as Aperio's ImageScope (Aperio Technologies, 2005) are regularly released, incorporating features such as annotation, image property adjustment (such as lighting) and simultaneous multi-slide viewing.

During this work, new versions of the VPS and ReplaySuite developed to incorporate new functionality and address limitations of the preceding versions. The following Chapter describes the limitations of existing VPS and ReplaySuite

technology, and their subsequent redevelopment in response to these issues. The objective of this work has to create more robust and flexible applications that could be utilised in new studies.

4.2 Critical Appraisal of VPS 1.0

The VPS was originally conceived and developed in late 2000 at a time when digital slides were still at an early stage of evolution. It was designed to facilitate the examination of slides digitised with the in-house scanning system originally developed during the same period, and track data describing the examination process. A more comprehensive description of the architecture of the original system is available in the thesis work describing its development (Costello, 2004). The original VPS possessed a number of important architectural features that were fundamental to its design:

Client-side Customised Browser

The Graphical User Interface (GUI) was delivered via a customised web browser, which controlled access for users during VPS dedicated studies and optimised the integrity of recorded data.

Web-based Delivery

The system utilised web-based delivery methodologies (PHP, HTML, JavaScript) for transmitting and displaying VPS images and data between the VPS server and the remote client application.

Database Driven Architecture

Examination and study data was stored on a relational, database management system, and SQL queries were used to write to and retrieved data from the database.

However, proposed utilisation of the VPS in the EQUALIS EQA scheme necessitated redevelopment of the technology. VPS 1.0 was designed to enable users to examine ten breast biopsies and submit diagnoses using a classification

system with a specific architecture (as described in Chapter 5). All filepaths and scoring data were hard-coded into the system, meaning that implementation in the EQA scheme required the code to be restructured. Secondly VPS 1.0 only allowed the user to examine 1 slide per case. The EQA scheme required two (differently stained) slides to be presented per case, which required additional redevelopment of the VPS code.

Finally, it was determined that a number of additional modifications were required in order to address deficiencies identified by the authors and during previous work (Costello, 2004). These pertained primarily to the GUI and methods of navigating slides. Figure 4.1 illustrates the stepwise process required to navigate to a field at highest magnification (2000x corresponding to a 40x objective field). Three additional factors were identified that were considered necessary to address in any subsequent redevelopment of the technology:

- A number of participants in the VPS evaluation study (Costello *et al.*, 2003), stated a preference for additional magnification layers. This request was addressed in the subsequent revision of scanning algorithms, however these image layers were not accessible with the original VPS, and required modification of the viewer.

- Navigational methodologies utilised by VPS version 1.0 were restrictive, and precluded rapid navigation from low to high magnification; users were required to step through all magnifications incrementally to reach higher magnification layer, as illustrated in Figure 4.1.

- Annotation capabilities to allow users to mark regions within a field.

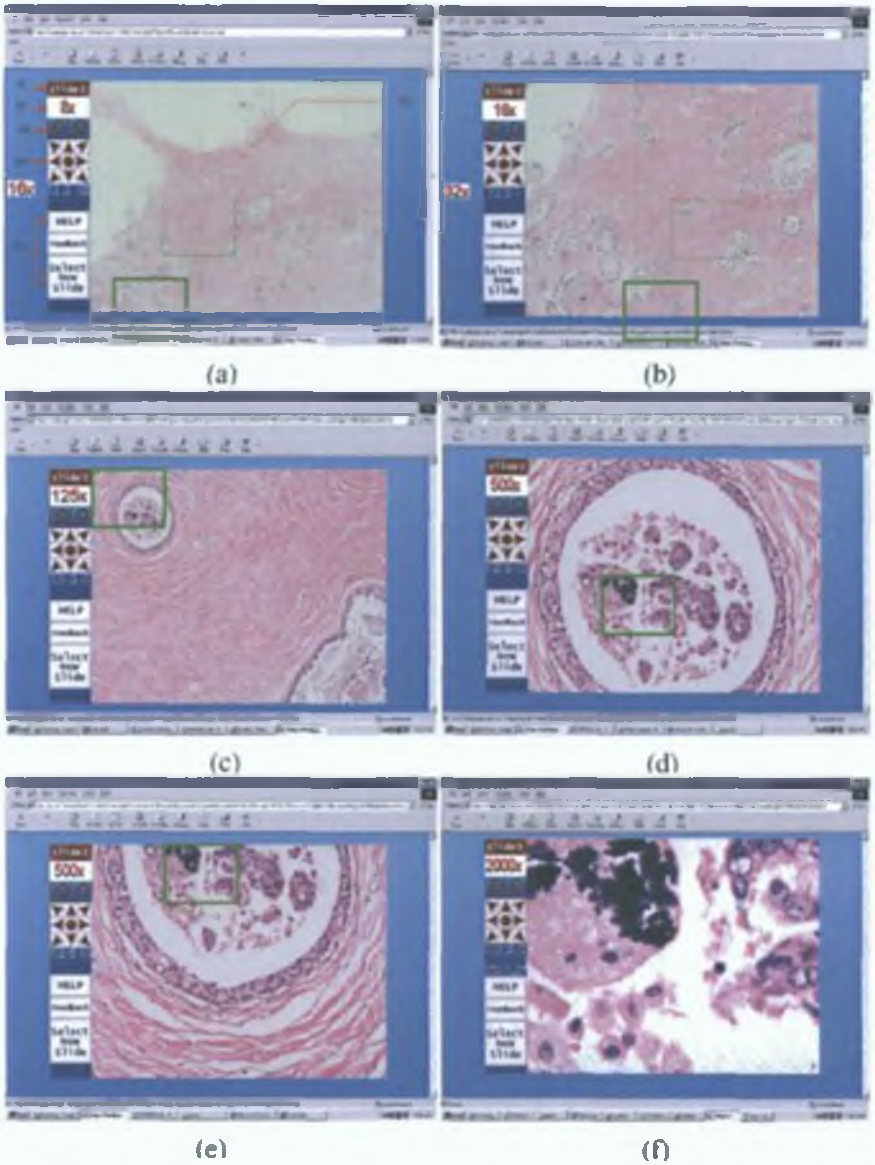


Figure 4.1 Migrating through VPS slide from low to high magnification. (a) Initial view of VPS at 16x (b) Zoom to 32x on area of interest. (c) Zoom to 125x on area of interest. (d) Move laterally around area of interest.(e)Zoom to 500x on area of interest.(f) Zoom to 2000x on area of interest. (i) Slide selected for viewing. (ii) Magnification of current view. (iii) Zoom and Zoom out buttons. A user may also zoom in by clicking on an image in the VPS field of view.(iv) Lateral navigation buttons for traversing within a given magnification.(v) Auxiliary user information buttons. (iv) The VPS field of view. Reproduced from Costello (2004).

4.3 Development of VPS 2.0

The existing VPS architectural design was considered a viable means of providing virtual slide viewing functionality and covert examination tracking, therefore it was concluded that the 3 fundamental architectural features of the VPS (customised browser, web-based delivery & database-driven) would continue to provide the structural basis of a redeveloped VPS. The customised browser was retained as the conduit for delivery of images and functionality, as it provided a means of controlling interaction with the WebPages that composed the VPS Graphical User Interface (GUI). A limited amount of redevelopment of the browser occurred, allowing it to store limited user session data and facilitate database interaction via the MFC CHTTP class (described in greater detail in the Appendix), separating content retrieval and database communication.

Web-based delivery of content was maintained, with continued utilisation of PHP for both dynamic HTML generation and Oracle database interaction (examination tracking). However, the incorporation of additional navigation methodologies resulted in the generation of additional data pertaining to the examination process, which necessitated modifications to the existing database structure. The architecture of the revised database is described in detail later in this Chapter.

4.3.1 Redeveloped Database Architecture

Changes to the VPS required new fields to be stored, to describe new navigational methodologies and annotations, as previously described. Additionally, the scoring regimen used during the EQA study differed to that used by the original VPS, with two parameters (stage and grade (Batts and Ludwig, 1995)) replacing classification and sub-classification. As previously mentioned, this scoring system is described in greater detail in Chapter 6. Figure 4.2 shows an entity relationship diagram that describes the structure of the database, with new tables and data columns within existing tables indicated by red boxes. Three new tables were created:

- *Masterequalis.* Set access to the active study phase for both the VPS and ReplaySuite. For example, if Phase I was the active scoring phase, cases 1-10 rather than 11-20, were available for examination with the VPS.
- *Zoomboxequalis.* Recorded pixel coordinates of areas selected via ZoomBox during examination.
- *Annotateequalis.* Recorded pixel coordinates of areas marked during annotation.

Three important new data columns within *trackingequalis* were optionally recorded: *annotated, MouseX* and *MouseY.* A value was recorded in *annotated* if the field was marked, or had a diagnostic comment appended to it, while *MouseX* and *MouseY* recorded the pixel coordinates of the mouse cursor for mag-list navigation.

4.3.2 Graphical User Interface

Figure 4.3 illustrates the architecture of the redeveloped VPS system and the channels through which to navigate between different VPS functionality, delivered via multiple Webpages. Discreet communication between the client side VPS browser and server-side Oracle database was facilitated through CHTTP, enabling VPS use to be monitored and recorded independently of content delivery. This reduced the number of events in series required to display images, improving system speed.

Navigation between webpages that constituted the GUI did not vary significantly between the original and redeveloped version. As Figure 4.4 illustrates, users accessing VPS 2.0 were able to (a) log in, (b) select a case, (c) examine and (c) submit a diagnosis as version 1.0 had allowed. The revised visual appearance of the interface was controlled with the use of Cascading StyleSheets (CSS), which enabled the separation of content (information displayed onscreen) from presentation (colours, fonts and layouts used). This allowed parameters such as background colour and default text style to be set for all webpages by one file (denoted by the .css suffix). A more detailed description of StyleSheets is provided in Appendix A.

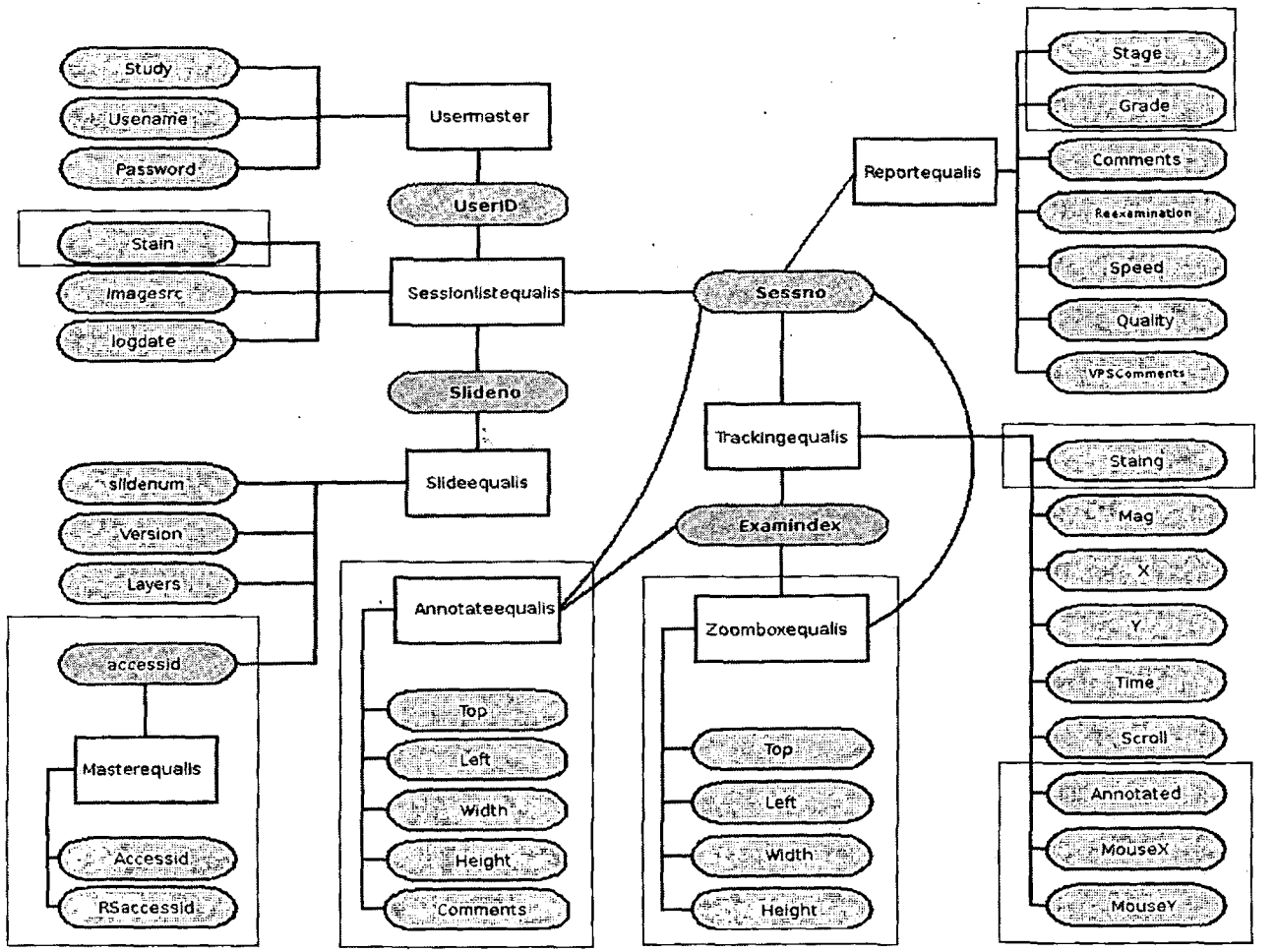


Figure 4.2 Entity Relationship diagram describing the architecture of the VPS relational database.

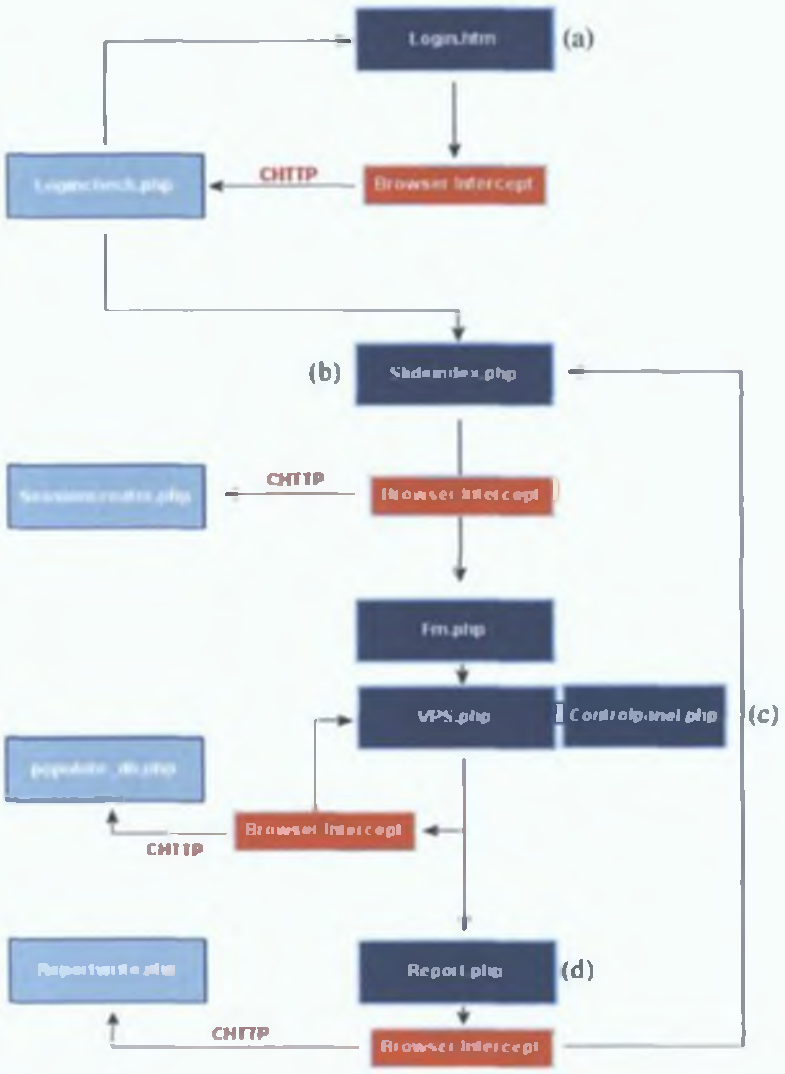


Figure 4.3 User navigation between VPS version 2.0 GUI webpages. (a) Log in using a username and password (b) Selecting a case to examine (c) Examining a virtual slide (d). Submitting a diagnosis using the summary report form.

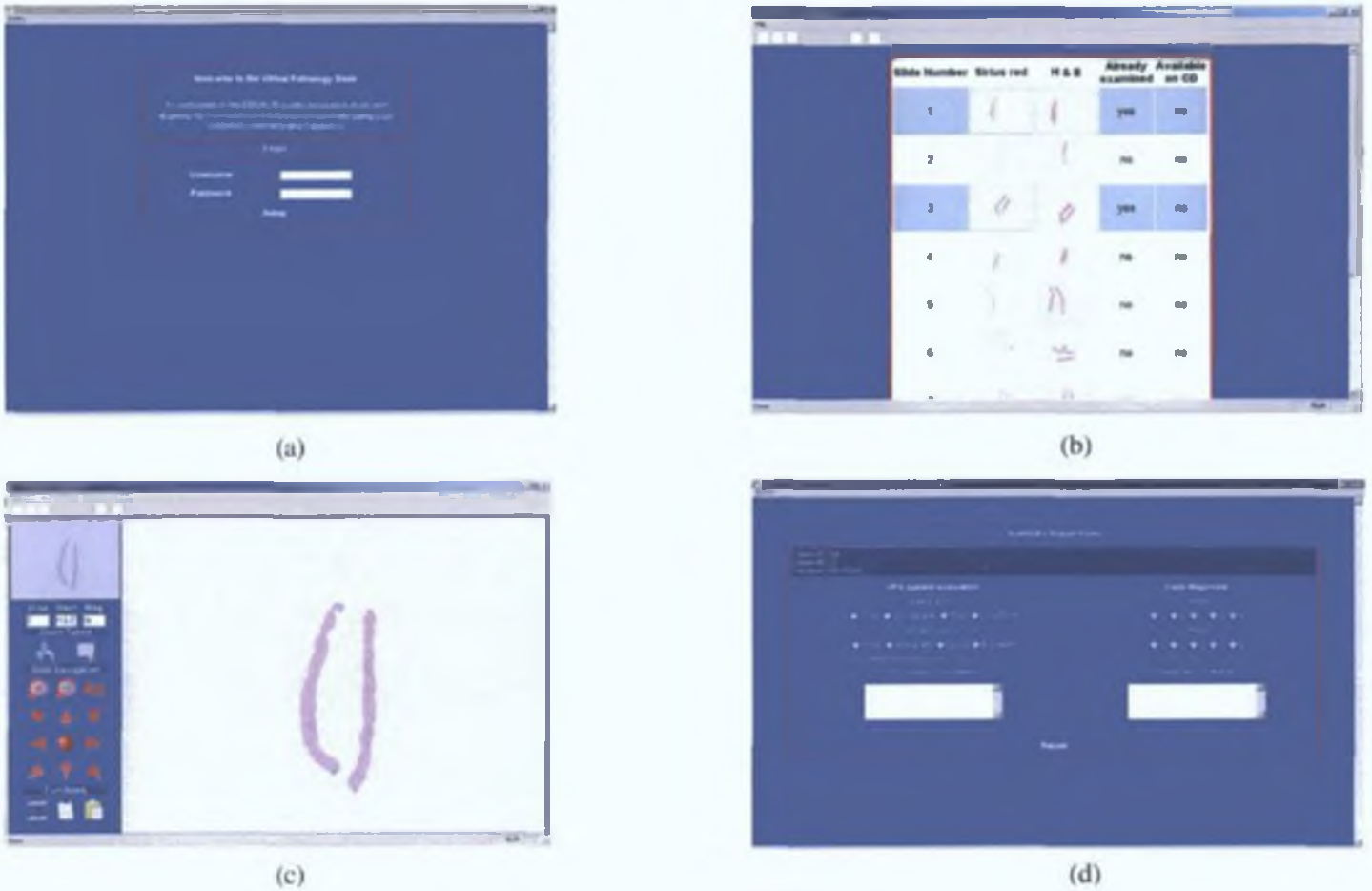


Figure 4.4 The VPS 2.0 Graphical User Interface. (a) Logging in using a username and password (b) Selecting a case to examine (c) Examining a virtual slide (d) Submitting a diagnosis using the summary report form.

Executing the browser application displayed the login page (Figure 4.4 (a)), which allowed the user to access the system and information pertaining to their participation in the EQA study. Login requests were intercepted by the browser, and transmitted via CHTTP to a PHP page (*logincheck.php*) that validated their login data. If the user's account does not exist, the user is redirected to the Login page; if authenticated, the user's VPS Identification number (*UserId*) is retrieved from the database, transmitted back to the client where it is stored as a variable within the browser. The requested Webpage (*slideindex.php*) is retrieved from the server and displayed within the browser. This webpage displays a table of cases available for examination in the study, each row relating to a study case. As Figure 4.5 illustrates, the VPS slide identification number and overviews for both case slides (Sirius red and H & E stains) were shown for each case. Once a user had examined a case and submitted a diagnostic report, they were no longer able to examine that case. If the user has already examined a case, it is indicated in the 'already examined' column, and also by a blue bar through the row.

Slide Number	Sirius red	H & E	Already examined	Available on CD
1			yes	no
2			yes	no
3			yes	no
4			no	no

Figure 4.5 Displaying cases available for examination using the VPS 2.0.

The 'Available on CD' column displayed whether a VPS CD in their CD/DVD drive, and which slides were contained on the CD. Selecting to examine a case 'available on CD' served slide images from that location during examination. If not, images were retrieved from the remote server. Clicking an overview for a specific stain directed the user to the examination suite (Figure 4.4 (c)), and displayed the requested stain.

Selecting a case to examine requested a frameset file (*fm.php*) that managed the presentation of the examination suite GUI into frames (previously described in Chapter 2). As Figure 4.6 illustrates, the frameset creates two columns, into which two files are loaded:

- Left column: The controlpanel (*controlpanel.php*) displayed examination information and icons for access to VPS functionality
- Right column: the page (*vps.php*) that displayed the virtual slide main field of view.

The advantage of this is that when the VPS page was refreshed to display a new field of view, the controlpanel did not refresh. This improved the appearance of the interface, as apposed to a single webpage, and reduced the amount of data that must be transmitted from the server, improving speed.

Figure 4.6 Frame layout of the VPS 2.0 examination interface.

When the virtual slide was first loaded into the viewer, images representing the lowest magnification for the selected case and stain were displayed. In order to examine fields at higher magnification, the user was able to navigate via 3 methods:

- Icons on the control panel
- Keyboard keys
- Mouse interaction with the main field of view

4.3.2.1 Control Panel

The slide overview, which provides the location of the current field of view was displayed at the top of the control panel. As Figure 4.7 illustrates, the slide ID, stain and current magnification level were displayed below the slide overview. In addition to displaying information, the control panel provided the user with access to navigational functionality, annotation and diagnosis submission capabilities. As with the original VPS , '+' and '-' magnification icons enabled stepwise changes in magnification of the current field of view. The 8-way navigation panel allowed lateral motion. The newly incorporated '4x' icon permited the examiner to return to the lowest magnification image. The 'Zoom Type' buttons are used to select the style of zoom navigation, which is used while interacting with the main field of view with the mouse.

As with the original VPS, a 'submit diagnosis' icon displayed a summary report form, with which the examiner was able to submit a diagnosis. The 'change stain' icon enabled users to switch between different slides for the same case (e.g. Sirius red and Haematoxylin & Eosin). Clicking the icon saved the location on the current slide and changed the main field of view to the other slide. If the new slide was previously examined during the same session, the last viewed field was displayed. Beside the change stain icon was the annotation icon. This enabled users to not only add diagnostic comments to fields, but mark regions within the current field of view to associate with the comments. This will be described in greater detail later in the chapter.

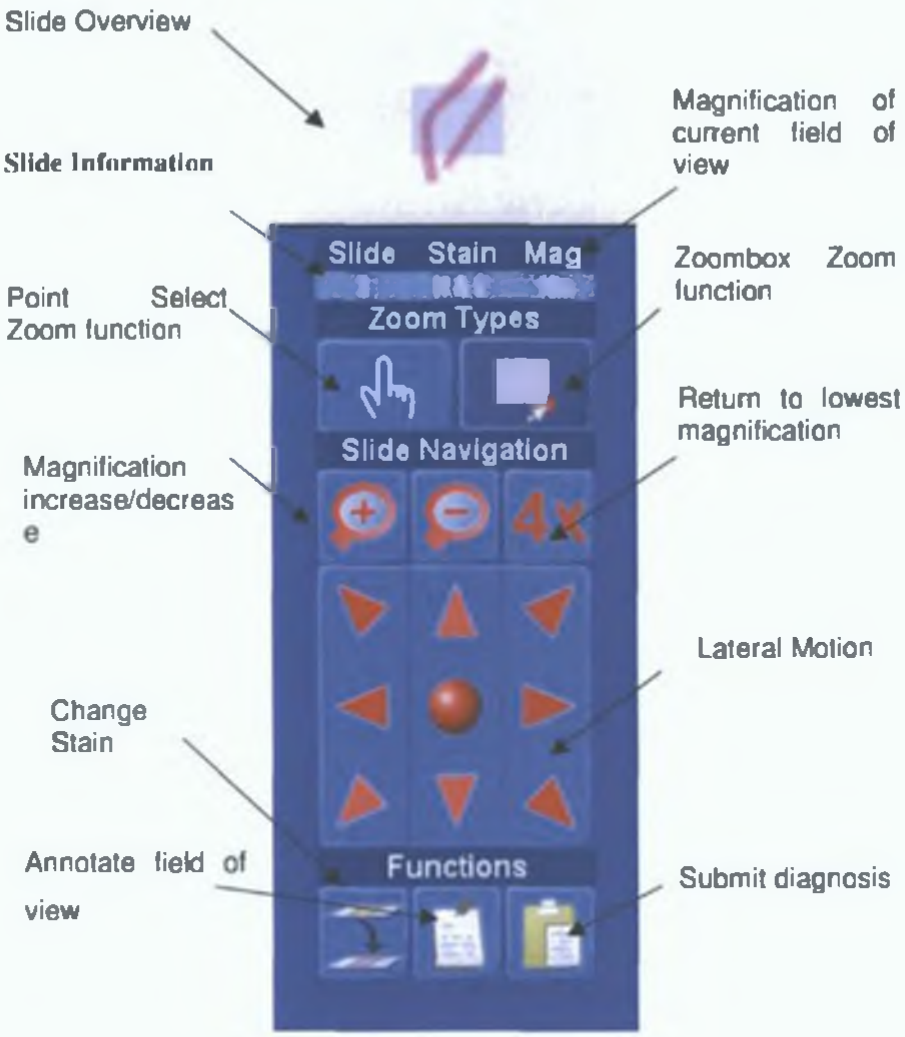


Figure 4.7 The VPS 2.0 Control Panel.

4.3.2.2 Navigation Methodologies

Version 2.0 of the VPS provided a number of methods of navigation while examining a slide, via mouse interaction with the GUI or digital slide, or via the keyboard.

4.3.2.2.1 Keyboard Navigation

Keyboard navigation functionality (an additional navigational methodology incorporated into the revised VPS) enabled 4-way lateral motion with the keyboard cursors, 8-way lateral motion using the number pad (1-9), and incremental increase/decrease in magnification using '+', '–' keys and 'Page Up', 'Page down' keys. Table 4.1 below denotes the function of each key.

Table 4.1 VPS 2.0 navigation using the keyboard.

Function		
		Numeric keypad
Zoom into centre of image	Page up	+
Zoom out of image	Page down	-
Move up laterally	Up arrow	8
Navigate laterally	Down arrow	2
Navigate left laterally	Left arrow	4
Navigate right laterally	Right arrow	6
Navigate left & up laterally		7
Navigate left & down laterally		9
Navigate right & up laterally		1
Navigate right & down laterally		3

4.3.2.2.2 Mouse Navigation

The user was able to select a region of interest in the main field of view to examine at higher magnification in 3 ways:

- By clicking the 'point select' zoom function, users were enabled to left mouse-button click on a region of interest to double the magnification and centre the view at that point. If the users mouse had a wheel, clicking this would zoom out (half the magnification), centring around the area clicked (Figure 4.8)

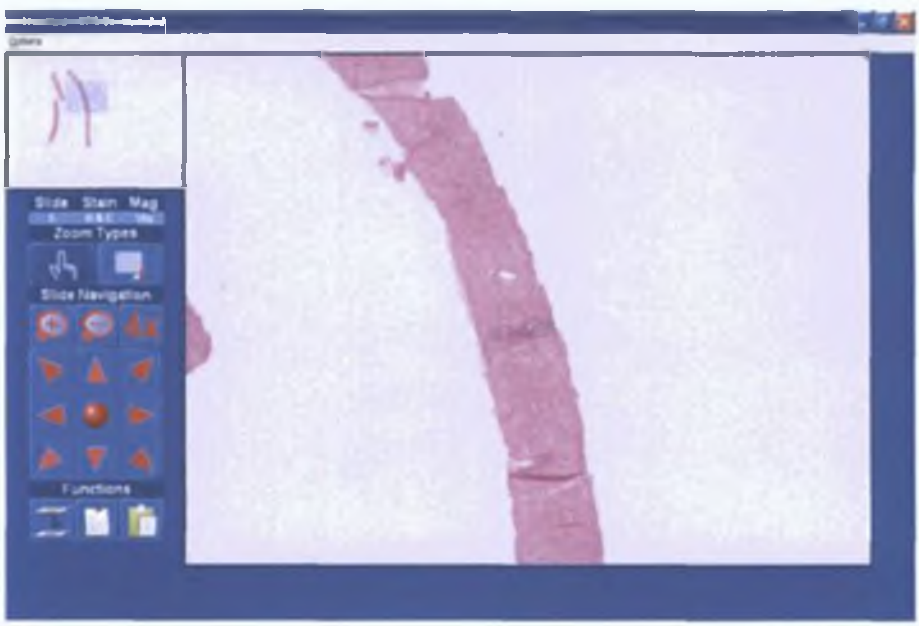
- By right-mouse clicking on a point, the user was presented with a range of options. Highlighting the 'mag list' option presented the user with the full range of magnifications available. This allowed users to skip unrequired magnification layers, and view a region at any magnification quickly, centring on the point where the right mouse button was pressed (Figure 4.9).

- By selecting the 'ZoomBox' function from the controlpanel, users were able to manually select an area to view at higher magnification. To select an area, the left mouse button was clicked once to begin creating a ZoomBox, then when the area had been created, clicking the left mouse button again navigated to the selected area. The VPS determined the appropriate magnification and centred on the middle of the ZoomBox (Figure 4.10).

4.3.2.3 Annotation

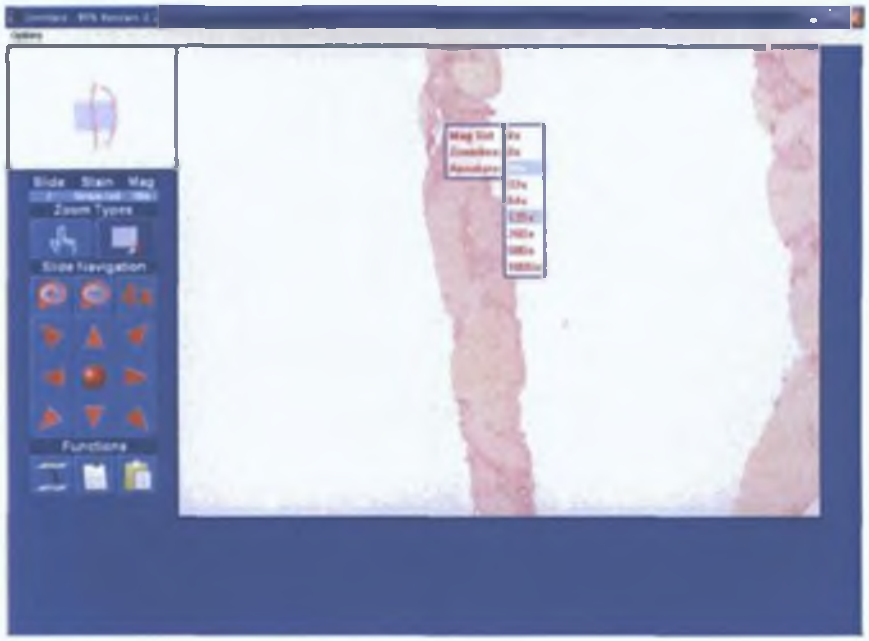
In addition to allowing users to attach diagnostic comments to examined features, the revised VPS enabled users to mark visual features within the current field of view. By clicking the 'Annotate' icon at the bottom of the controlpanel, the user was able to mark a rectangular region using the same approach in creating a ZoomBox. A comment box then appeared to allow the user to record any diagnostic comments. When finished annotating, clicking the 'add note' icon to switched off annotations. Figure 4.11 illustrates the creation of an annotation.

(a)

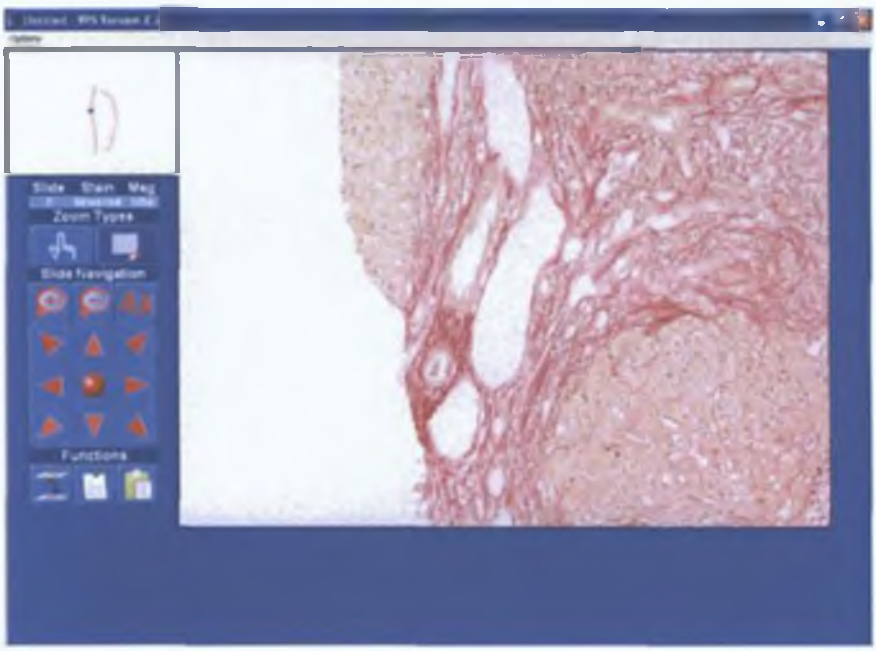


(b)

Figure 4.8 Selecting an area to view at higher magnification via point-select navigation in VPS 2.0. (a) Select an area with the mouse cursor to (b) centre on the region and double magnification.

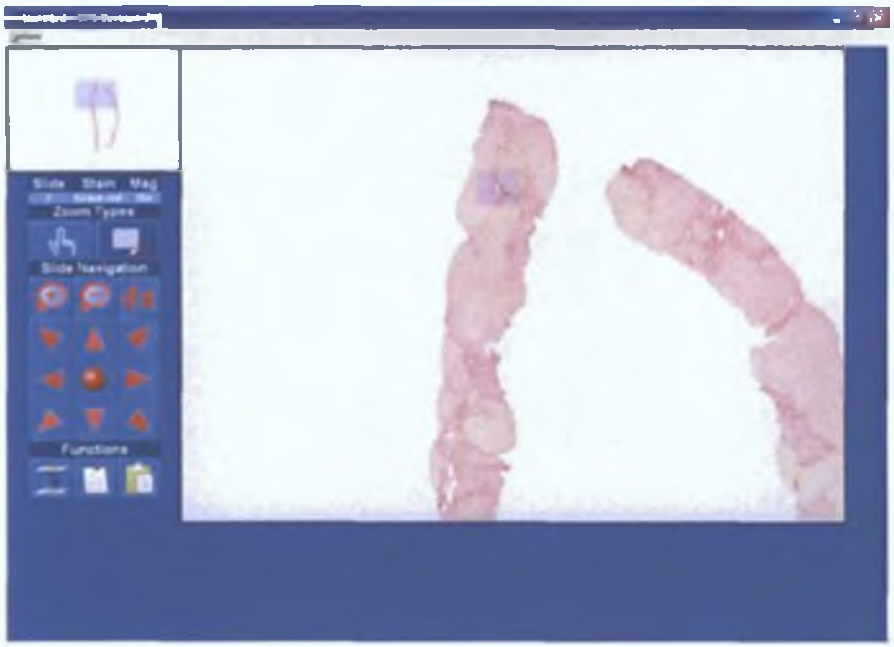


(a)

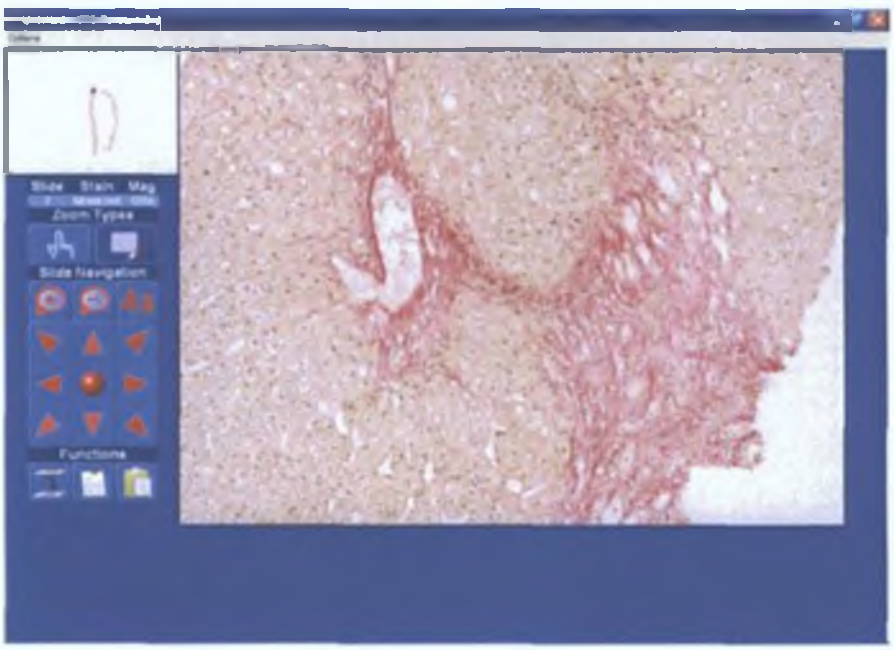


(b)

Figure 4.9 Selecting a region of interest using the mag-list menu option in VPS 2.0. (a) Clicking on a point and selecting a magnification from the drop down menu and (b) centering on the chosen point at the selected magnification.

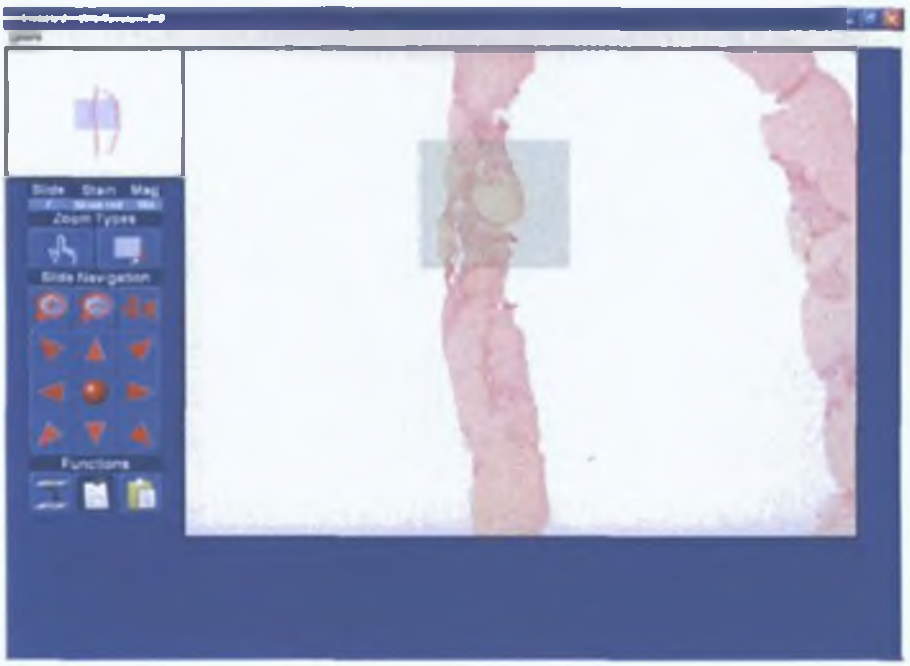


(a)

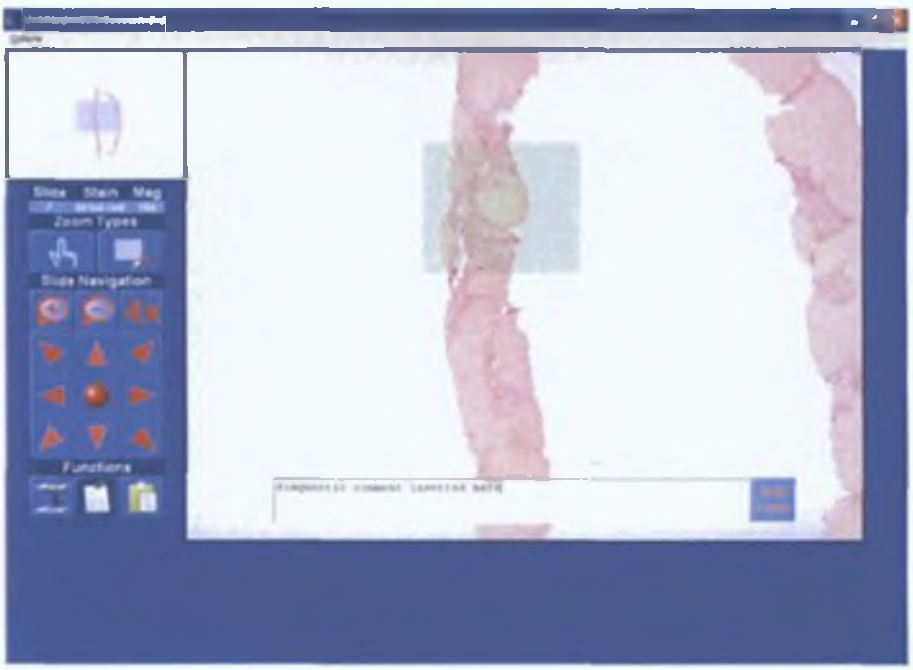


(b)

Figure 4.10 Navigating from 16x to 125x via the ZoomBox function in VPS 2.0. (a) Selecting a region with the Zoombox and (b) viewing the selected region.



(a)



(b)

Figure 4.11 Annotating a field with VPS 2.0. (a) Marking a region with an annotation Box and (b) adding a comment.

4.3.2.4 Submitting a Diagnosis

When the user was ready to make a diagnosis, clicking the 'report' icon at the bottom of the controlpanel directed the user to the summary report form (*report.php*). Here they could submit scoring information, diagnostic comments and evaluations of the VPS system. The classification system utilised by the EQA scheme contained to criteria that were evaluated, staging (fibrosis) and grading (necroinflammation). This scoring system is described in greater detail in Chapter 6.

The report provided the examiner with a means to submit case diagnosis data (Stage, Grade) via radio buttons, and evaluate the VPS system (speed, image quality) via free text submission, as Figure 4.12 illustrates. Once the report was submitted, the user was returned to the list of remaining slides to be examined (*slideindex.php*).

Figure 4.12 Submitting a summary report form after examining a case with VPS 2.0.

4.4 Critical Appraisal of ReplaySuite 1.0

The ReplaySuite was originally developed in 2001 to exploit data from previous work (Costello, 2004) to ascertain whether examination replays were of benefit as a pathology resource. Participation in the EQUALIS EQA study (Chapter 6) provided an opportunity to elucidate whether ReplaySuite use could impact on diagnostic performance in chronic hepatitis scoring. However, in its original format the ReplaySuite could not be appropriately utilised in the study. A number of deficiencies in the existing system were identified:

Hard-coded architecture

Version 1.0 had been developed to display data from examinations of 10 VPS breast needlecore biopsies; filepaths and classification regimens were hard-coded into the system. In order to provide access to liver biopsy images and analyse diagnostic data from the new scoring system, the system had to be rewritten.

Revised VPS navigational approaches

The redeveloped VPS (version 2.0) incorporated new navigational methodologies, which resulted in additional data stored on the relational database. Version 1.0 was unable to utilise this data, requiring new methodologies of data exploitation.

Additional analytical capabilities

New novel analytical approaches had been identified that could be incorporated into a new ReplaySuite version, providing additional functionality to end-users.

In order to address these issues and facilitate utilisation of the ReplaySuite in the EQUALIS EQA scheme, it was concluded that a new version of the ReplaySuite was required.

4.5 Development of ReplaySuite 2.0

As previously described in Chapter 2, the original ReplaySuite possessed three key architectural features: (1) *Client-side Customised Browser* (2) *Web-based Delivery* (3) *Database-driven Architecture.* In keeping with the successful redevelopment of the VPS (described previously in this Chapter), these features were maintained. The customised browser was retained as the conduit for delivery of images and functionality, as it provided a means of controlling interaction with the WebPages that composed the Graphical User Interface (GUI). It was delivered via the revised VPS browser, allowing separating content retrieval and database communication, as previously described. Web-based delivery of content was maintained, with continued utilisation of PHP for both dynamic HTML generation and Oracle database interaction.

The redeveloped ReplaySuite continued to provide core ReplaySuite functionality; VPS examinations could be replayed, diagnostic concordance graphs viewed and examinations searched for according to a number of parameters (Case, User, Diagnosis). Modifications pertained to changing the system to enable VPS version 2.0 examinations and study data to be accessed, and to provide new diagnostic trace analysis capabilities (SnapShots), described later in this Chapter.

4.5.1 ReplaySuite 2.0 Database Architecture

In order to monitor application use, the ReplaySuite tracked users logging onto the system, and recorded the type and frequency of functionality accessed. This is achieved in a similar manner to VPS tracking, with PHP files utilised to write data describing ReplaySuite use to the database. Figure 4.13 illustrates the architecture of the database table structure.

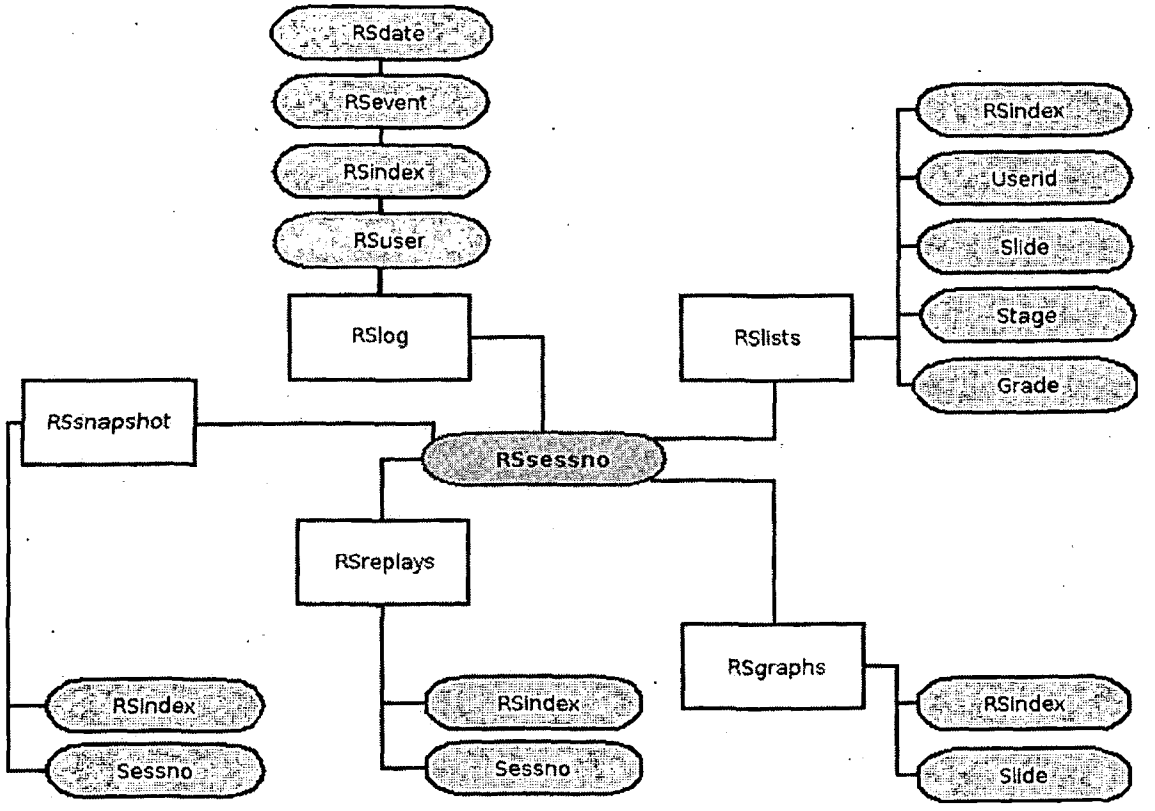


Figure 4.13 Database tables for tracking ReplaySuite 2.0 functionality use. Tables containing data are denoted by white boxes, data columns by light grey and key data columns by dark gray. Tables contain information on user log in (RSlog) and functionality accessed while logged in.

4.5.2 Graphical User Interface

As with the revised VPS, StyleSheets were used to control the appearance of the GUI, enabling separation of content (information displayed onscreen) and presentation (colours, fonts and layouts used). Frames were not used, with onscreen elements controlled via StyleSheets using absolute positioning (see Appendix A for more details).

The redeveloped ReplaySuite enabled users to view lists of VPS examinations, view diagnostic concordance data, replay previously performed examinations and view visual representations of entire examinations, known as SnapShots (described later in this Chapter). Users could navigate between these ReplaySuite functions via onscreen hyperlinks, as illustrated by Figure 4.14.

4.5.2.1 Study Data

Figure 4.15 illustrates the Graphical User Interface displayed when the user successfully logged on. In contrast to ReplaySuite 1.0, study data (describing gold standard and group concordance) was displayed on the right panel, while search parameters and results were displayed on the left, allowing the user to cross-reference the two. Users were still able to search for examinations according to a number of defined parameters, where Stage and Grade were the diagnostic parameters evaluated (described in greater detail in Chapter 6):

- User
- Case
- Stage
- Grade

The study overview panel provided the user with a range of pertinent data relating to cases previously examined. As Table 4.2 illustrates, the study overview panel was sub-categorised into 5 columns.

Table 4.2 The ReplaySuite 2.0 Study Overview Panel

My Examination options	**Replay:** Replay the logged in users examination of this case **Stats:** View diagnostic concordance graphs for this case **SS/SnapS:** View a snapshot of the logged in users examination of this case
My diagnosis	Logged in users case scores
KVAST* diagnosis	KVAST* case scores
Group consensus	Group consensus scores and highest % consensus

* KVAST group (described in greater detail in Chapter 6) provided 'gold standard' staging and grading.

In order to make scoring data more accessible and easier to evaluate, scoring data was colour coded to indicate agreement/disagreement with KVAST scoring. KVAST concordant diagnoses are identified by green, with over and under scoring examinations indicated by purple and orange respectively. Additionally, the shade of orange/purple indicated of the degree of discordance: the darker the shading, the greater the degree of discordance with KVAST.

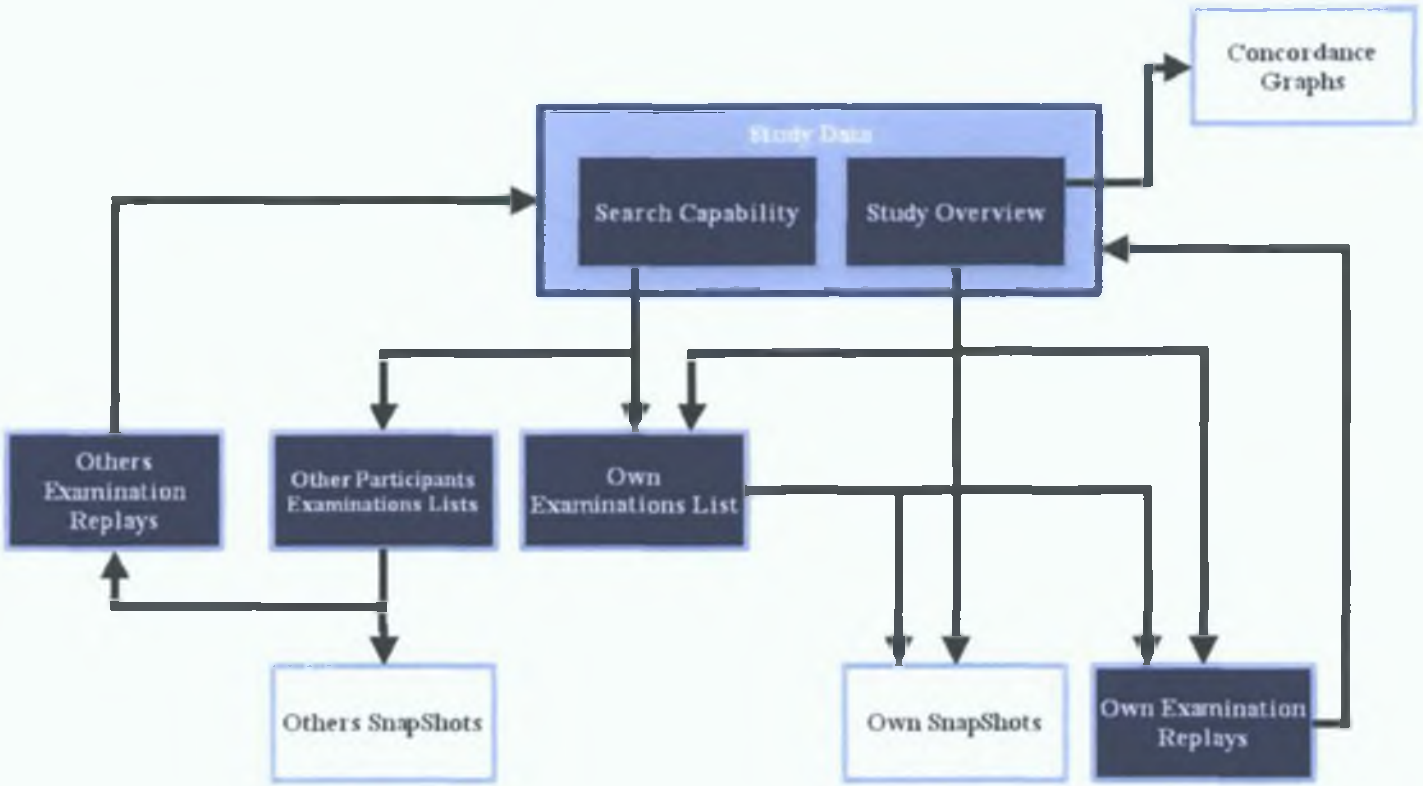


Figure 4.14 Navigation between ReplaySuite 2.0 functionality, where white boxes indicate the use of pop-up windows to display functionality.

Figure 4.15 Searhing for examinations, and viewing 'gold standard' and group concordance data for a scoring phase using the ReplaySuite 2.0.

4.5.2.2 Replaying an Examination

Clicking the movie camera icon within a row replayed the examination. As Figure 2.1 illustrates, the basic layout of the replay GUI was maintained. The principal focus of the GUI was the main field of view, which displayed fields viewed during the examination. To the left was the information panel, which displayed the slide overview, examination and slide information. While examinations could still be fast forwarded, rewound and paused, the video controls were now located in the left sided information panel. This provided more space for the main field of view. As with version 1.0, users could navigate to the last viewed examination list, or a list of their own examinations by clicking the appropriate icons.

A number of navigational methodologies were available for use with the VPS 2.0. JavaScript was used to indicate which strategies were utilised during replays. PHP variables recorded by the VPS 2.0 were fed into JavaScript functions, prior to commencing a replay. These variables indicated the type of navigation that was utilised, and where it occurred onscreen. Before each new field was displayed, the appropriate function was called, and an icon displayed within the main field of view to indicate the type of upcoming navigation. Figure 4.17 illustrates the eight different types of icon that may be displayed.

All icons were displayed in the centre of the main field, with the exception of indications of upcoming changes in magnification using point selection and ZoomBox. For these types of navigation additional data recorded was recorded by the VPS 2.0, indicating the location of the mouse cursor at the time. This data is subsequently used to place the relevant icon at the x,y coordinates of the mouse cursor at the time of the event during the examination. ZoomBox navigation is indicated using a transparent box, as displayed in Figure 4.17(f). A semi-transparent div was positioned according to the coordinates recorded by the VPS, with semi-transparency achieved using Microsoft visual filters, which allows the application of multimedia style effects to webpage elements (Microsoft corporation, 2004).

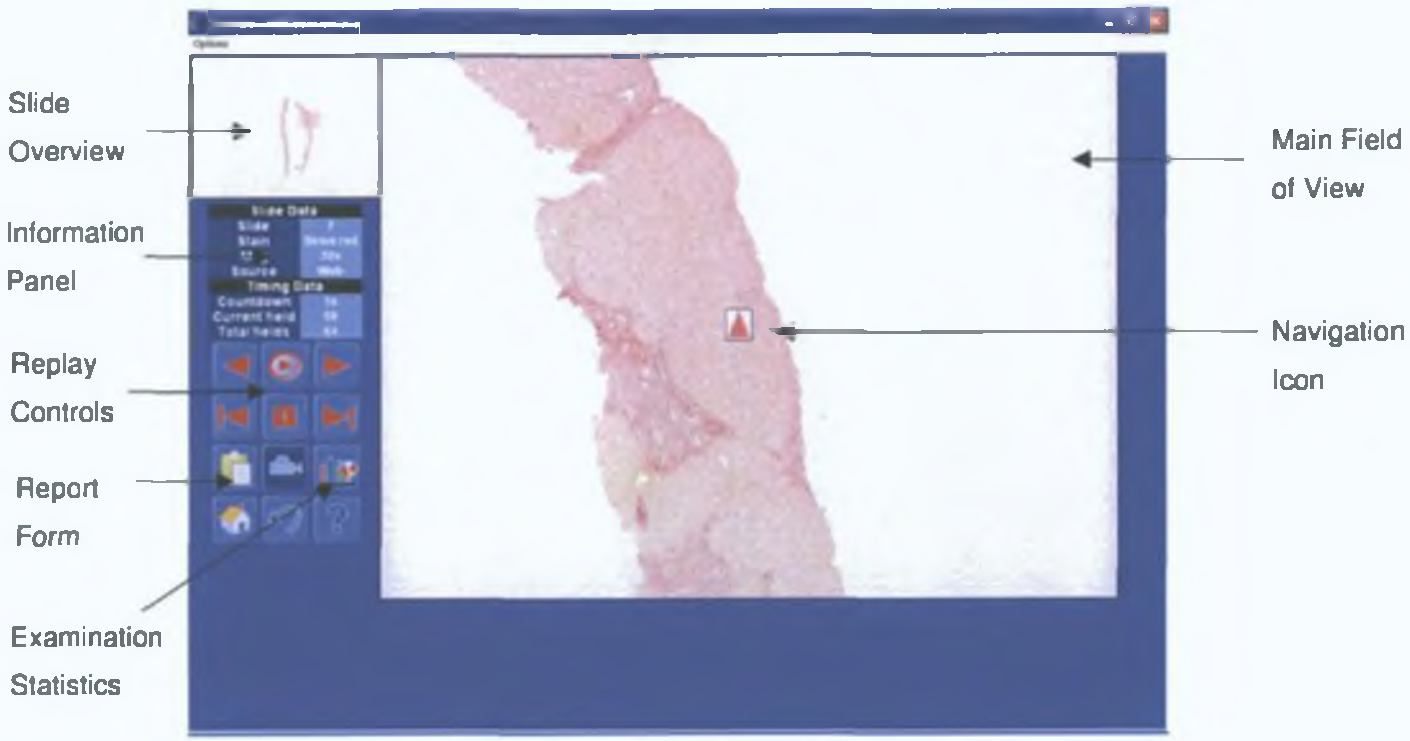


Figure 4.16 Replaying an examination of Case 7 using the ReplaySuite 2.0.

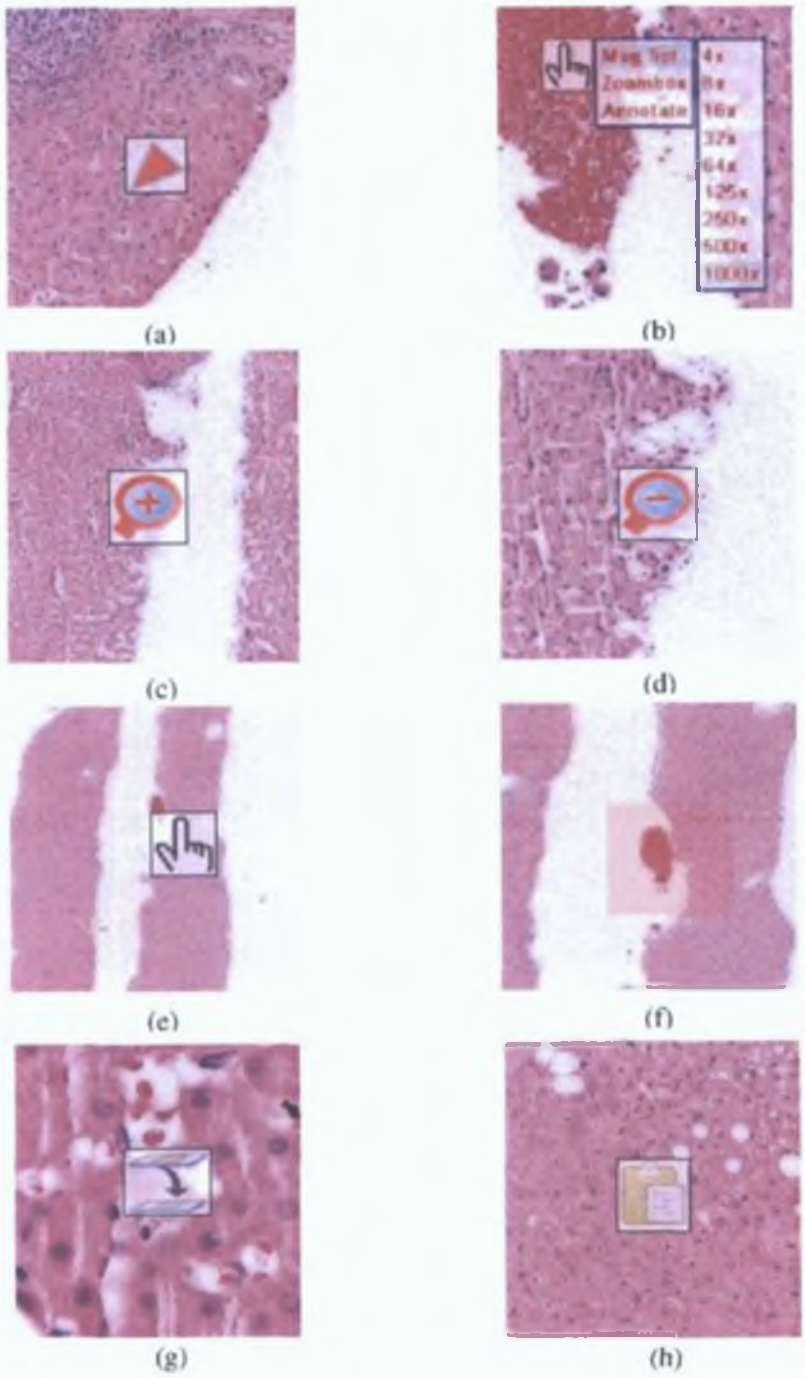


Figure 4.17 Indication of upcoming navigation during an examination replay: (a) lateral motion (b) Maglist navigation (c) doubling of magnification (d) magnification halving (e) Magnification increased by point selection (f) magnification increased by ZoomBox (g) slide switch (h) examination end.

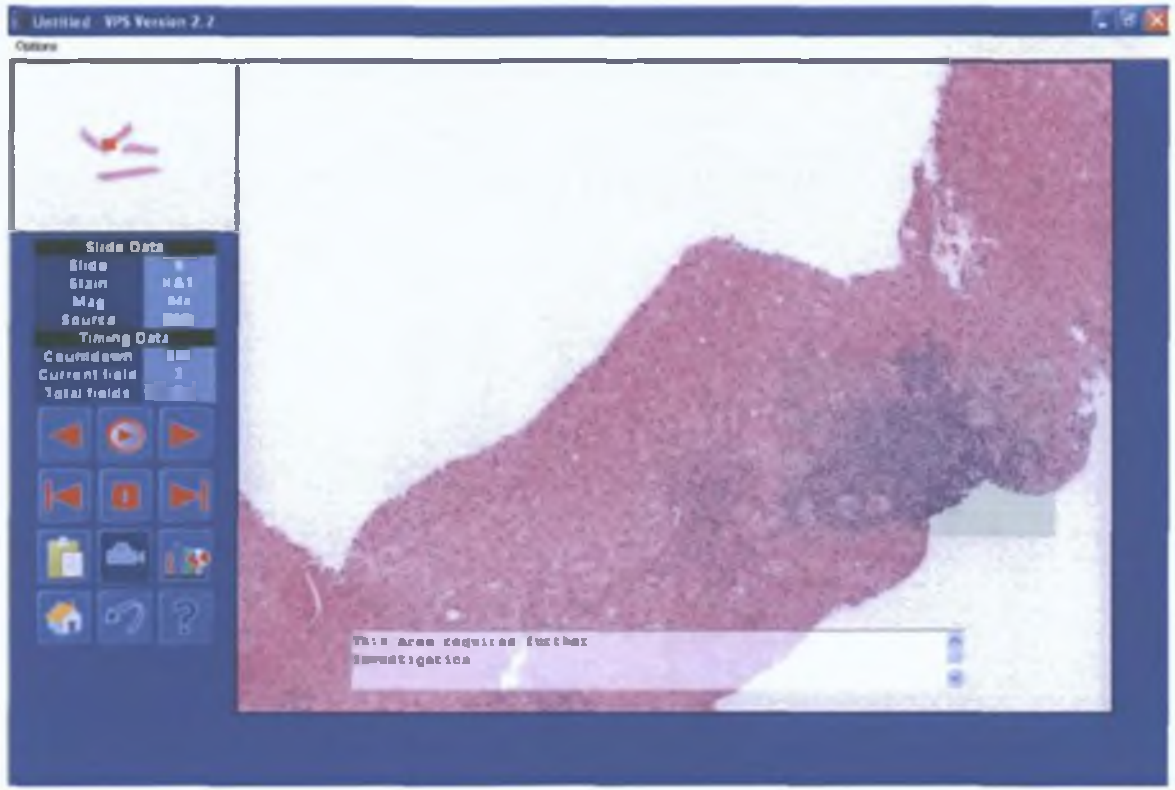


Figure 4.18 Reviewing annotated and marked fields during a ReplaySuite 2.0 replay.

4.5.2.3 Summary Report Forms

To improve ease of use, users were no longer required to rewind to the beginning of the examination to view the summary report form, or the end to view statistical analysis of the examination. Icons on the information panel provided access to these functions. By clicking either icon, the replay was automatically paused, and the position of the replay recorded. Clicking the Replay Icon returned the user to the examination replay at the previously paused position and restarted the replay. Figure 4.19 displays a summary report form, as viewed using the ReplaySuite 2.0. As the scoring system utilised by the two versions differ, this is reflected in the summary report form. All other displayed parameters remain the same.

4.5.2.4 Examination Statistics

In contrast to version 1.0, examination statistical analysis is restricted to evaluating time spent examining each slide at the available magnifications. As Figure 4.20 illustrates, the number of fields viewed, and the total and mean time spent examining at each magnification is indicated.

4.5.2.5 SnapShots

The most significant advancement of the ReplaySuite 2.0 over its predecessor is the incorporation of SnapShot functionality. A SnapShot is a visual representation of an examination, a slide overview indicating the areas examined, and the magnifications used to view each area. Figure 4.21 displays a SnapShot as displayed by the ReplaySuite. Colour coding is used to indicate the magnifications used, generating a heat map that indicates 'hot spots'; areas viewed at high magnification. SnapShots for both stains are available by clicking the icon to the top left of the SnapShot panel.

A JavaScript graphics library is used to convert tracking coordinates into transparent div elements, which are overlaid onto the slide overview to generate SnapShots, shown in appendix. Each magnification layer is assigned a colour, and layered sequentially (lowest → highest magnification) over the slide overview image. Absolute positioning is used to correlate the position of the layers with the appropriate area of the slide overview. Finally, a mask is overlaid to fit the SnapShot layers to the tissue.

Figure 4.19 Viewing a Summary Report Form using the ReplaySuite 2.0.

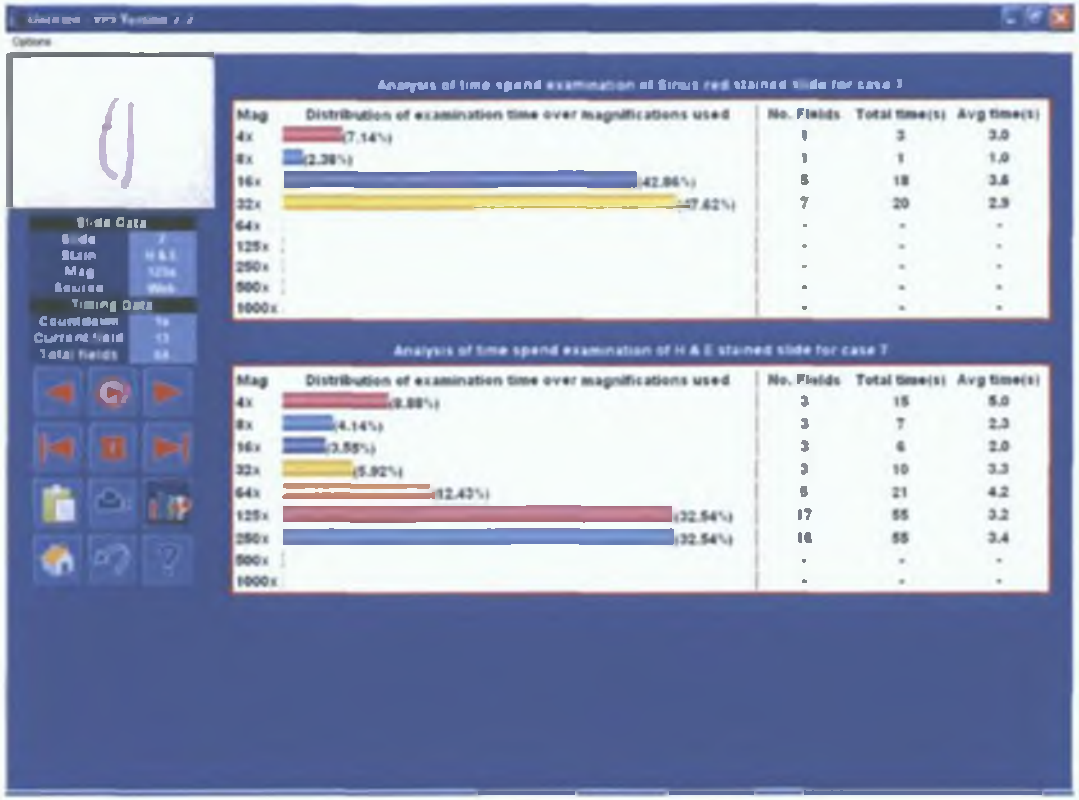


Figure 4.20 Viewing examination statistical analysis with ReplaySuite 2.0.

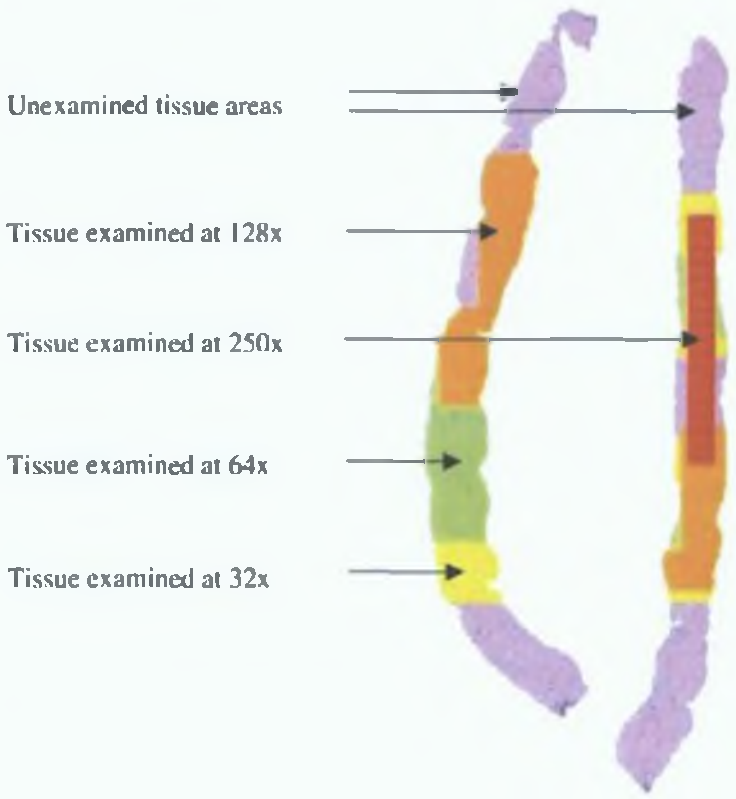


Figure 4.21 SnapShot of a VPS 2.0 examination of a H&E stained slide for Case 7.

4.6 Conclusion

Redevelopment of the VPS viewer was primarily to facilitate participation in the EQUALIS EQA scheme, but it also successfully addressed pathologist's recommendations from previous work (Costello *et al.*, 2003). Many of the features incorporated into the revised viewer have recently been implemented in other slide viewing technologies, such as ZoomBox navigation (Aperio Technologies, 2005) and annotation capabilities (Lundin *et al.*, 2004).

Glatz-Krieger *et al* stated as recently as 2003 that HTML based viewers were the most common form of virtual slide viewer, citing eSlide (University of Udine, 2004) as a prime example. However the use of formats such as FlashPix (EastMan Kodax), Zoomify (Zoomify Inc., 2004) and Java-based viewers is rapidly superseding HTML-based viewers. This is illustrated by the fact that the Telemedicine Research Group at the University of Udine have developed a Java-based viewer for viewing eSlides (University of Udine, 2004). Formats such as FlashPix and Zoomify permit multiple resolutions of an image to be stored in a single file and facilitate smooth scrolling; a feature almost universally preferred by pathologists to stop/start navigation.

While virtual slide viewers allow entire slides to be examined by panning and zooming, they do not provide control over every parameter available with light microscopy. Images are captured prior to virtual slide examination; meaning real-time adjustment of lighting and focusing parameters is not feasible, within the context of what is permissible with light and dynamic microscopy. While virtual slide viewers, such as Aperio's ImageScope, allow the digital adjustment of lighting parameters, this is performed on the static image, post-acquisition, rather than real-time. Aperio's 3D revisit option proposes to provide through-focus inspections with z-stack capture, however this effectively equates to dynamic microscopy, as the slide is required to be in scanner during the examination session.

The most significant technical failure in the redevelopment of the VPS related to the inability to adequately incorporate smooth scrolling. As previously mentioned, this is quickly becoming standard practice with slide viewers, however its incorporation raised issues regarding tracking, using the current incarnation. Technical difficulties were also encountered as a result of using a customised browser, illustrating the importance of extensive, long term testing regimens. The necessity to redevelop the VPS to facilitate use of an alternate diagnostic classification system illustrated the limitations of hard-coded systems, and the importance of dynamic scoring system generation. This observation have been applied to the development of the Digital SlideBox© (SlidePath Ltd, 2004), highlighting the necessity for flexible systems when employed in studies that use multiple scoring regimens.

Much of the ReplaySuite's redevelopment related to changes that enabled it to access and exploit VPS data obtained from the EQA scheme. However, the incorporation of SnapShot functionality that built on previous work generating 'heat maps,' may be considered the most significant contribution to the ReplaySuite's redevelopment. Previous work provided similar functionality via an MFC application, Bitmapper (Costello, 2004). However to generate heat maps, it required the user to physically search for the examination ID number of the requested examination by directly querying the VPS database, then input the data into the application. In contrast, the ReplaySuite incorporated SnapShot functionality into the GUI, allowing the user to search for SnapShots based on diagnosis, user or case, while using the ReplaySuite.

The use of JavaScript to generate SnapShots demonstrates the flexibility and capabilities of the scripting language. Often perceived as only useful for providing trivial features, such as rollover effects to webpages, its use in this work illustrates its potential for more practical purposes. However, while JavaScript proved a useful method of generating SnapShot, more powerful programming languages such as client-side Java or server-side PHP will ultimately supersede it for this purpose. The database-driven architecture of the system creates the potential for extensive data interrogation, in order to analyse examination techniques amongst large sample groups. Previous work conducted

preliminary evaluation of such concepts, identifying a correlation between classification concluded and areas viewed in some cases (Costello, 2004). Future work may expand on this to develop methods of correlating specific visual cues with classifications concluded.

Section 2: - Evaluation of VPS, ReplaySuite and Commercial Telepathology Technology

Abstract

This section describes the following undertakings:

- Preliminary evaluation of the ReplaySuite 1.0 (Chapter 5). This was conducted with three objectives in mind. (1) Evaluate the potential value of the ReplaySuite application in pathology EQA and training. (2) Assess its potential to influence diagnostic decision-making. (3) Identify any application limitations/deficiencies prior to validation in a larger study.

- Utilisation of the VPS 2.0 and ReplaySuite 2.0 technology in the EQUALIS External Quality Assurance scheme in chronic hepatitis (Chapter 6). This was undertaken with 3 key objectives in mind. (1) Assess uniformity of staging and grading of chronic hepatitis in Sweden. (2) Evaluate the impact of supplementary electronic resources on performance. (3) Ascertain whether the ReplaySuite could be used to identify sources of error in histological assessment, by allowing an expert pathologist to review examination replays.

- Assessment of Irish trainee pathologist's perceptions of computer-assisted learning (Chapter 7) using commercial telepathology applications.

The ReplaySuite was initially developed (Chapter 2) to exploit the dataset obtained by previous work in breast pathology (Costello, 2004). This data was utilised during the preliminary evaluation of the ReplaySuite (Chapter 5), during which pathologists assessed the technology and its possible applications.

It was subsequently determined that it would be beneficial to use the technology in a study designed specifically for its evaluation. However, workload demands placed on pathologists often restricted their participation in research studies, especially when participation is purely on a voluntary basis.

It was therefore concluded that in order to evaluate whether use of the ReplaySuite could elicit an improvement in diagnostic performance or aid identification of sources of error in diagnostic technique, that the technology must be incorporated into a formal setting, such as a training scheme or External Quality Assurance (EQA) study. Involvement in such programs however would be opportunistic, and study design would be dictated by the scheme organisers, rather than specifically for the purpose of evaluating the ReplaySuite.

Such an opportunity arose with an invitation to participate in the EQUALIS chronic hepatitis EQA scheme in Sweden, a study concerned with evaluating levels of uniformity in scoring chronic hepatitis in liver biopsies. This work elucidated whether the use of supplementary electronic resources, such as the ReplaySuite, could improve diagnostic agreement with 'gold standard' diagnoses. The technology was also utilised to identify possible sources of error that contributed to discordant diagnoses (Chapter 6).

Finally, the perceptions of Irish trainee pathologists towards the use of an online pathology resource were assessed, to elucidate whether they concurred with international findings (Chapter 7).

Chapter 5: Preliminary Evaluation Of The ReplaySuite

5.1 Introduction

While telepathology has yet to be incorporated into most pathology labs in a clinical role (Mairinger *et al.*, 1998, Della Mea *et al.*, 2000, Dennis *et al.*, 2005), its application in an educational context has garnered greater support. This has resulted in the development of a range of quality online training tools. Designed to supplement rather than replace human tutoring, such tools enable training pathologists to gain a wider range of educational experiences. Studies have shown the use of online training tools can improve diagnostic performance, beyond what is found with human tutoring alone (Dee *et al.*, 2003, Heidger *et al.*, 2002).

Accurate identification of visual features associated with specific diagnoses is a fundamental skill that pathologists are required to develop. Reference material is a vital resource for trainee pathologists developing these skills, and the provision of material via online tools delivers extensive 'add on' capabilities unavailable though conventional means. The majority of training tools comprise interactive tutorials that display a limited number of pre-selected images per case, often with accompanying notation (Kronz et al., 2000, Lehmann and Cohen, 2002, Smith et al., 2000, College of Medicine University of Illinois at Urbana-Champaign, 2005). However, as has been observed through the use of static telepathology systems for remote diagnosis, diagnostic accuracy is often dependent on appropriate field selection (Weinstein *et al.*, 1997, Weinberg *et al.*, 1996, Halliday *et al.*, 1997). Novices often make errors when searching the slide (Crowley *et al.*, 2003), so the development of the skills required to locate relevant visual features are as important as those required for identifying them.

With the cost and time required to digitise entire slides decreasing, digital slides are finding greater application in pathology education (Glatz-Krieger *et al.*, 2005, 2003). Providing unrestricted examination of entire digitised tissue sections, they give a more accurate representation of the microscopic environment used by pathologists for diagnosis. When incorporated into training

tools, they can provide unrestricted but supportive examining environments, allowing trainees to develop searching and identification skills.

The VPS provides an unrestricted examination environment, but also discreetly records examination data on a remote, relational database, where it is available post-examination for interrogation. This potentially enables the diagnostic technique of different pathologists to be analysed and studied, however the VPS does not utilise this data constructively to provide an intrinsic benefit to the end user, the pathologist. For this purpose, the ReplaySuite was developed.

The ReplaySuite is a web-based, user-friendly software tool that enables pathologists to replay virtual slide examinations, performed using the Virtual Pathology Slide (VPS). Unlike interactive tutorials and annotated digital slides, pathologists using the ReplaySuite are able to observe the diagnostic trace of examiners, in a manner similar to the use of a double-headed microscope. Previous work with the VPS (Costello, 2004) enabled an expert pathologist to review fields examined during previous examinations, in order to attempt to identify sources of error. However this process was laborious, time consuming and inefficient, as it required the physical location, capture and presentation of individual fields. In contrast, users of the ReplaySuite may view all examined fields in chronological order, without expending any more effort than identifying an examination to review. This possesses significant potential for use in both pathology training, where trainee pathologists may learn from the diagnostic techniques of experts, and quality assurance, for the detection and elucidation of sources of error in participants diagnostic technique.

The following Chapter describes the implementation of a preliminary ReplaySuite evaluation study, designed to elucidate experienced pathologists perceptions of the technology and its potential applications in Pathology External Quality Assurance (EQA) and Training. Also subsequent references to the 'ReplaySuite' in this Chapter refer to Version 1.0, as described in chapter 2.

5.2 Study Procedure

The preliminary study was conducted to demonstrate the capabilities of the ReplaySuite to pathologists, and evaluate their opinions on its use and potential application in training and quality assurance. In order to use the ReplaySuite, participants were required to have first examined VPS slides. Participants were provided with open access to the 10 needlecore biopsies examined during the VPS validation study (Costello *et al.*, 2003), which they were required to examine using the VPS. Once a participant had submitted a diagnosis for a case, that participant was permitted to review their own examination data and that of other pathologists, for that slide, using the ReplaySuite. Participants who had previously participated in the VPS validation study (by examining cases) were not required to re-examine cases, and could review any examination.

Using the ReplaySuite, participants were permitted to view group concordance graphs, replay their own and others examinations, view summary report forms and view statistical breakdowns of individual examinations. Use of the ReplaySuite was monitored and recorded to the database, allowing the identification of light/heavy users and highlighting the most frequently used functions. Cross-referencing VPS use (who examined which slides) with ReplaySuite use (who replayed which examinations) enabled the identification of which participants examinations were replayed and for which slides.

5.2.1 Slides

The ten needle core biopsies examined during the VPS validation study (Costello, 2004) were obtained by selecting the first breast biopsy generated each month for a ten month period from the Department of Pathology, Mater Misericordiae Hospital, Dublin, Ireland. Glass slide diagnoses (clinical diagnosis) were provided by a pathologist with a special interest in breast pathology, with consultation required to finalise glass slide diagnoses on some of the cases selected.

Diagnoses were concluded using the Core Biopsy Reporting Guidelines for Non-operative Diagnostic Procedures and Reporting in Breast Cancer screening (NHS Cancer Screening Programmes, 2001), as used by the British National Co-ordinating Committee for Breast Screening Pathology. This contains 5 classifications, with sub-classification for one category (B5), as illustrated in Table 5.1:

Table 5.1 Categories used to diagnose VPS needlecore breast biopsies, from the Core Biopsy Reporting Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening.

B1	Normal Tissue	Used to describe a core including normal breast ducts and lobules or mature adipose tissue or stroma only
B2	Benign Lesion	Appropriate for a range of benign lesions such as Fibroadenomas, Fibrocystic changes, Sclerosing adenosis and Duct ectasia
B3	Lesion of Uncertain Malignant Potential	B3 or B4 result in either diagnostic excision of the area or repeat of core biopsy sampling to obtain a definitive diagnosis
B4	Suspicious	
B5	Malignant	Indicates unequivocal malignancy on core biopsy. Further sub-classification into in-situ or invasive

Table 5.2 shows the glass slide diagnoses and VPS group concordance diagnoses for the ten cases. The most difficult cases were considered to be: Case 1 (presence of Apocrine changes), Case 8 (Lack of discreet division between B3 & B4), Case 9 and Case 10.

Slides were digitised with the customised VPS imaging workstation and transferred to the VPS server. A more detailed explanation of the slide scanning process used to digitise study slides is provided in dissertation work on the subject (Costello, 2004).

Table 5.2 Comparison of glass slide needle core surgical biopsy diagnosis and most-common VPS 1.0 diagnosis, in order of level of agreement (concordance) for each slide. Reproduced from Costello (2004).

	Slide									
	6	2	3	4	7	9	1	10	5	8
Glass Diagnosis	B5	B5	B5	B2	B2	B2	B5	B2	B5	B3
VPS Group consensus	B5	B5	B5	B2	B2	B2	B5	B2	B5	B4
Concordance (%)	100	94.1	82.4	76.5	64.7	58.8	52.9	52.9	47.1	35.3

5.2.2 Participants

70 pathologists were invited to participate in the study. 38 Irish participants were invited with the assistance of Professor Peter Dervan, Mater Misericordiae Hospital, Dublin, with the remaining invitees comprising the 32 members of the European Working Group of Breast Screening Pathology (EWGBSP). EWGBSP members are recognised as expert breast pathologists in their native countries and all member states of the European Union (EU) are represented within the group. 17 of the 70 invited participants previously participated in the VPS validation study, 4 of which were members of the EWGBSP.

5.3 ReplaySuite Study Exit Survey

Subsequent to the study, participants completed an electronic questionnaire on their use and impressions of the software and its potential applications. Participants were asked to submit their level of agreement/disagreement with 19 statements using a Likert scale, and answer Yes or No to a further 5 questions. 2 text boxes were provided to allow submission of additional open-ended commentary.

Invented by Renis Likert (1932), the Likert scale is a commonly used means of measuring attitudes and opinions. Originally implemented with 7 ordinal categories (e.g. Strongly Disagree/ Disagree/ Slightly Disagree/ Undecided/ Slightly Agree/ Agree/ Strongly Agree), it included a 'neutral' or 'undecided' middle point. The number of intervals to use and inclusion/exclusion of a neutral option is often debated. Advocates of the neutral point argue that it ensures respondents do not manufacture opinions instantaneously, while opponents argue that in reality people are never neutral on issues and always have an opinion (Jamieson, 2004). The ReplaySuite exit study survey questions used 5 categories including a neutral point, with a number of questions inverted to mitigate against the well-known bias of positively phrased questions.

5.3.1 Post Study Survey

Participants who used the ReplaySuite were resurveyed post-study in order to determine if they regularly participated in teaching. 3 questions were asked:

- Are you, or have you been involved in providing undergraduate medical training on a regular basis? (*Yes/No*)
- Are you, or have you been involved in providing postgraduate pathology training on a regular basis? (*Yes/No*)
- If you are not currently involved but have previously been in either activity, please state how long ago you were involved. (*Open-ended*)

5.4 Study Participation

Of the 70 pathologists invited to participate, 9 used the ReplaySuite to review examination data. All 9 completed the electronic questionnaire, and all but one had more than five years experience in pathology practice, the exception (User 36) possessing three years experience. As Table 5.3 illustrates, 4 of the 9 participants were members of the European Working Group on Breast Screening Pathology. Slides were examined either during the VPS validation study, or during the ReplaySuite preliminary evaluation study All 9 participants completed the post-study survey.

Table 5.4 summarises ReplaySuite use, based on functionality utilised. The average number of ReplaySuite functions performed (where 1 function relates to viewing an examination list, concordance graph or examination replay) was approximately 30 (269/9) per user. However one participant (User 55) used the ReplaySuite extensively, corresponding to 40.14% of all use (108/269). The reasons for this users extensive use of the system are unknown. By excluding User 55 mean use is adjusted to approximately 18 functions per user. ReplaySuite use was distributed between viewing examination lists (58.36%), diagnostic concordance graphs (25.27%) and examination replays (16.35%), with most frequently accessed concordance graphs and replays for cases with lower group concordance (1,5,8-10). Heaviest ReplaySuite users were also the heaviest users of examination replay capabilities (Users 10, 74, 55).

7 participants (77.7%) replayed at least one examination and of these, 3 participants replayed more than 10. Participants replayed their own examinations over other pathologist's examinations by a ratio of 3:2. Only 2 participants (Users 10 & 74) reviewed other's examinations extensively, as Table 5.4 illustrates.

Table 5.3 Origins of preliminary ReplaySuite 1.0 evaluation study participants and previous participation in VPS evaluation study.

User	Participant origin		Occurrence of slide examinations		Concordance with Glass diagnosis
	EWGBSP*	Irish	VPS evaluation study	ReplaySuite evaluation study	
74	√			√	7/8
102		√		√	8/10
55	√		√		8/10
10		√	√		7/10
57	√			√	6/10
7		√	√		6/10
39		√	√		6/10
36		√	√		4/10
60	√			√	0/1

* European Working Group of Breast Screening Pathology

Table 5.4 Participant use of the VPS 1.0 and different functionality of the ReplaySuite.

	VPS	ReplaySuite use					
User	Examinations performed	Total examinations replayed	Own examinations replayed	Other participants' examinations replayed	Diagnostic concordance graphs viewed	Examination lists viewed	Total number of functions performed
55	10	11	10	1	28	69	108
74	10	12	4	8	11	35	58
10	10	13	7	6	3	19	35
39	10	2	2	0	13	9	24
7	10	3	2	1	4	12	19
57	10	0	0	0	9	4	13
36	10	2	1	1	0	4	6
102	10	1	0	1	0	4	5
60	2	0	0	0	0	1	1
Total	92	44	26	18	68	157	269

Table 5.5 Distribution of own examinations replayed by participants using the ReplaySuite 2.0.

User	Slide										Total
	1	2	3	4	5	6	7	8	9	10	Replayed
55	6				4						10
10	1	1	1	1	1	1	1				7
74						1	1	1	1		4
7	1				1						2
39	1			1							2
36	1										1
57											0
60											0
102											0

Table 5.6 Distribution of other pathologists examinations replayed by participants using the ReplaySuite 2.0.

User	Slide										Total
	1	2	3	4	5	6	7	8	9	10	Replayed
74								3	2	3	8
10	1						1	3		1	6
7					1						1
36							1				1
55	1										1
102		1									1
39											0
57											0
60											0

As Table 5.7 illustrates, participants reviewed 68 diagnostic concordance graphs. 66.6% (6/9) of all participants viewed diagnostic concordance graphs for at least 2 slides and 4 participants (44.4%) viewed graphs for at least 7 slides. The two heaviest ReplaySuite users (55, 74) were also two of the most frequent graph viewers.

Table 5.7 Frequency of diagnostic concordance graphs viewed by participants using the ReplaySuite 2.0.

User	Slide										Total
	1	2	3	4	5	6	7	8	9	10	
55	6	3	2	2	4	2	2	3	2	2	28
39	3	2		2	1		2	2	1		13
74	2	1	1			1	1	1	2	2	11
57	1	1	1	1	1		1	1	1	1	9
7	1				3						4
10	1							2			3
36											0
60											0
102											0
Total	14	7	4	5	9	3	6	9	6	5	68

5.5 Survey Results

Table 5.8 Percentage distribution of 9 participant responses to ReplaySuite preliminary evaluation study survey questions 1-13.

Q	Please state how much you agree with the following statements:	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
1	The ReplaySuite is user-friendly	0	0	25	63	13
2	The ReplaySuite is slow	0	22	44	11	22
3	The ReplaySuite displays information clearly	0	11	0	89	0
4	The ReplaySuite has an innovative *help* file	0	0	67	33	0
5	The diagnostic pathway of the examining pathologist was apparent	0	0	11	78	11
6	Being able to pause, fast-forward and rewind was useful	0	0	33	44	22
7	The information panel was not useful	11	78	0	11	0
8	Images were of good quality	0	22	22	56	0
	Please state the importance of the following ReplaySuite features:	**Not at all**	**Of little importance**	**Neutral**	**A little important**	**Of great importance**
9	Examinations list	0	11	33	22	33
10	Diagnostic concordance graphs	0	0	11	11	78
11	Examination replays	0	11	22	22	44
12	Summary report forms	0	0	11	33	56
13	Individual examination statistical breakdown	0	0	11	22	67

Table 5.9 Percentage distribution of 9 participant responses to ReplaySuite preliminary evaluation study survey questions 14 & 15-19.

		Yes	No	N/A
	Where applicable answer yes or no to the following:			
14	Did you encounter technical difficulties whilst using the ReplaySuite	33	67	0
16	Did replaying an examination by another pathologist provide you with an insight into why a diagnosis was concluded?	44	22	33
17	Did replaying an examination by another pathologist with a diagnosis different to your own provide you with an insight into why that diagnosis was concluded?	56	11	33
18	Did replaying examinations performed by other pathologists make you reconsider your diagnosis for any slides?	22	44	33
19	If Yes, did your new diagnosis concur with group consensus?	22	0	77

Table 5.10 Percentage distribution of 9 participant responses to ReplaySuite preliminary evaluation study survey questions 20-25.

Q	Rate the benefit of using the ReplaySuite in for the following applications where:	Of no benefit	Of little benefit	Neutral	Of some benefit	Of great benefit
20	Training	0	11	11	56	22
21	Quality Assurance	0	11	11	56	22
	Rate the benefit of the following to telepathology:	**Not at all important**	**Of little importance**	**Neutral**	**A little important**	**Of great importance**
22	Digitised digital slides	0	0	22	22	56
23	Downloading virtual slide images from the Internet	0	11	11	33	44
24	Replaying your own examinations	0	0	11	33	56
25	Replaying other pathologists examinations	0	0	11	22	67

All participants who replayed at least one examination agreed or strongly agreed that the ReplaySuite was user-friendly (Figure 5:1) and 85.71% (6/7) of these agreed or strongly agreed that being able to pause, fast-forward and rewind an examination replay was useful. 6 of the 7 of participants (85.71%) who replayed an examination considered replays to be of some or of great importance as a ReplaySuite feature.

When asked to rate the potential benefit of the ReplaySuite in pathology training, 6 out of 7 of participants who had replayed at least one examination (85.71%) and all of those who had replayed more than 10 examinations rated it as of some or of great benefit (Figure 5.2).

When asked to rate the potential benefit of the ReplaySuite in quality assurance, 5 out of 7 of participants who had replayed at least one examination (71.4%) and all participants who had replayed more than 10 examinations considered it of some or of great benefit (Figure 5.3).

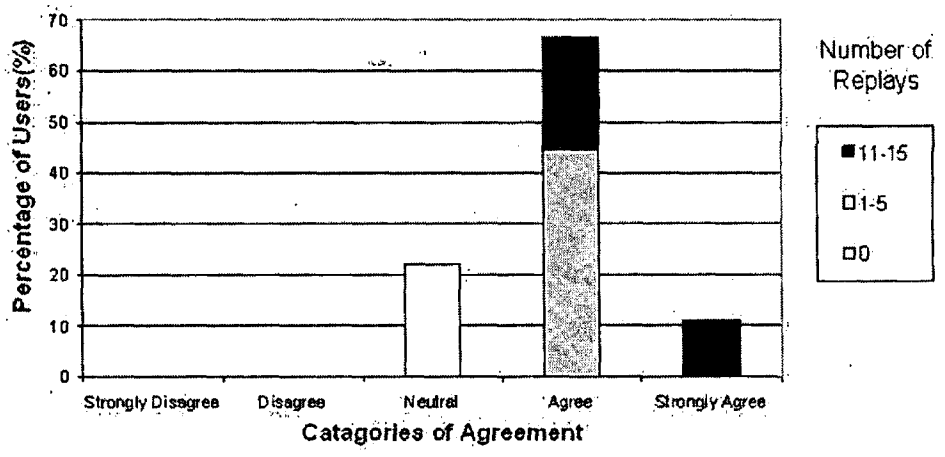


Figure 5.1 Levels of agreement with the statement 'The ReplaySuite is User-Friendly'

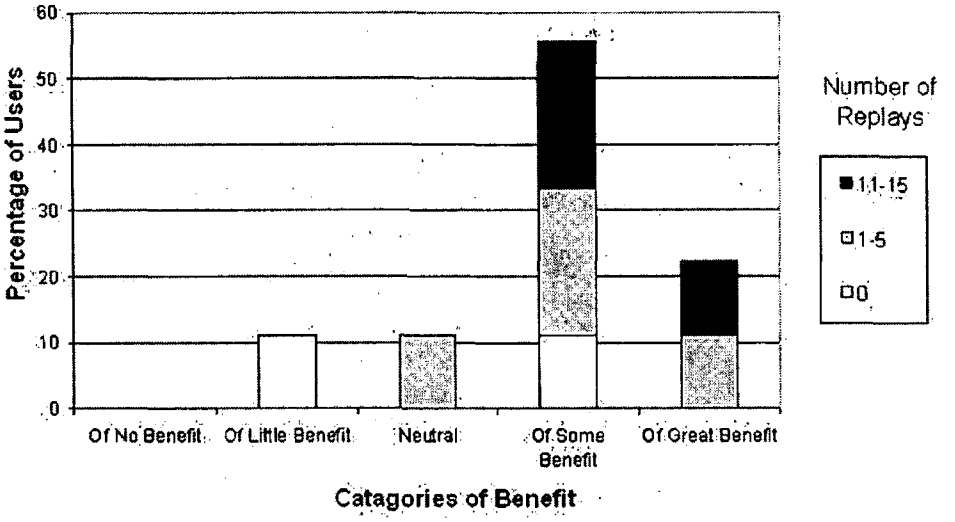


Figure 5.2 Levels of perceived potential benefit of the ReplaySuite in Pathology Training

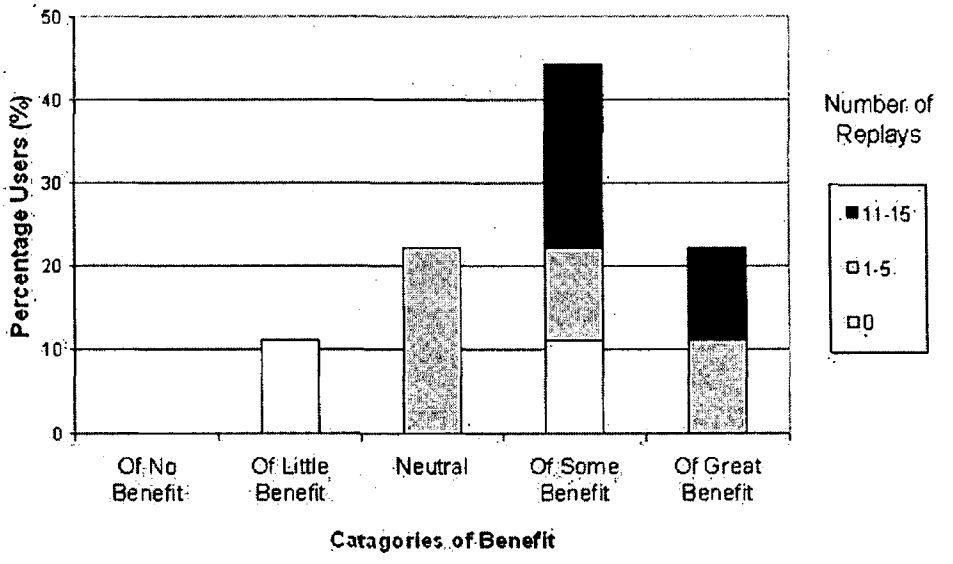


Figure 5.3 Levels of perceived potential benefit of the ReplaySuite in Quality Assurance

5.6 Diagnostic Re-Evaluation

All participants who replayed at least one examination agreed or strongly agreed that the diagnostic pathway of the examining pathologist was apparent during a replay. Of the 7 participants who replayed their own or others examinations, 6 (85.71%) replayed another pathologists examination. Of these 6 participants, 5 (83.3%) agreed that it provided an insight into the examining pathologists diagnosis. Of those who replayed another's examination, 2 (33.3%) reconsidered their original diagnosis, 1 re-diagnosing concordant with group consensus and 1 re-diagnosing concordant with original glass slide diagnosis.

After originally concluding an original diagnosis of B2 (Benign) for Slide 5, User 7 replayed his/her own examination, and then replayed an examination of Slide 5 by User 6. This examination (User 6) concluded a diagnosis of B5 (Malignant), concordant with group consensus.

User 10 concluded an original diagnosis for Slide 8 of B4 (Suspicion of malignancy), which concurred with group consensus but not glass slide diagnosis, B3 (Benign but of uncertain malignant potential). User 10 replayed their own examination, then an examination of the same slide that concluded the same diagnosis by User 5. Two examinations concluding a B3 diagnosis were then replayed (Users 87 and 55), concordant with glass slide diagnosis.

5.7 Post-Study Results

Of the 9 that completed the post-study survey, 7 (excluding Users 39 & 74) stated they were currently, or had been involved in providing undergraduate medical training on a regular basis, while 8 (excluding 74) stated they were currently, or had been involved in providing postgraduate pathology training on a regular basis.

5.8 Conclusions

Previous work had shown pathologists to be confident using the VPS to diagnose cases. 81.25% of participants in the VPS evaluation study indicated confidence using the VPS to make a diagnostic decision, with 56.25% describing themselves as "reasonably confident," 18.75% as "confident," and 6.25% as "very confident" (Costello, 2004). While a plethora of virtual slide viewers are available, the VPS examination tracking capabilities are unique. However, in isolation this data does not directly benefit the end user. The ReplaySuite was developed to exploit the VPS's examination tracking capabilities and utilise the data in a manner that would have tangible benefits for pathologists. Utilising the same technology as the VPS, the ReplaySuite does not require the installation of any additional software other than the customised browser, and Internet Explorer, which is the default web browser used by 95% of the web browser market (WebSideStory, 2004).

Use of the VPS for diagnosis in the VPS evaluation study had already demonstrated "substantial" diagnostic agreement between users, 88.23% of examining pathologists obtaining a Kappa of between 0.97 and 0.65 (Costello, 2004). In determining whether pathologists could benefit from using the ReplaySuite, the primary consideration was whether participants might reassess their diagnosis as a result of observing VPS examinations of the same slides by others. In the case of 2 users, observing another's examination caused them to reconsider their original diagnosis. In contrast to the re-diagnosis of Slide 8 by User 10: B4 (suspicion of malignancy) to B3 (Benign but of uncertain malignant potential), the difference between original and secondary diagnosis of Slide 5 by User 7: B2 (benign) to B5 (malignant) is a significant re-evaluation, one which would result in significantly different courses of treatment in a clinical environment. While the degree of discordance between original diagnosis and group consensus may, in part be attributable to poor screen resolution used during examination (640x480 pixels), it may also be related to the relative difficulty of the case; Slide 5 achieved the second lowest group consensus (47.1%). It cannot be suggested, based on these individual examples, that group concordant re-diagnosis subsequent to ReplaySuite use will be the rule rather

than the exception. However, these examples are worth noting, as it highlights the fact that using the system can result in diagnostic re-evaluation.

A number of caveats should be considered when reviewing study data, the first being small sample size. There is a significant difference between the number of pathologists invited to participate and those who completed the study (9/70). While not unique in studies involving pathologists use of telepathology systems, the small sample size may be considered a potential source of error. Bamford *et al* (2003) have reported similar difficulties with low participation rates, however, other studies have also attempted to evaluate telepathology tools using small sample sizes (Lee *et al.*, 2003). Low participation rates may be due to a number of factors. Bamford *et al* (2003) cited technical difficulties and pathologists workloads as principal factors for low participation. System speed was highlighted as an issue by a number of participants during the initial VPS evaluation study. As both VPS and ReplaySuite systems utilise similar technology, speed may be a potential contributing factor to low participation rates for this follow-on study. User 36 illustrated this when asked to comment on any technical difficulties encountered during the ReplaySuite evaluation study, stating, *"It took a great deal of time downloading image via dial up network connection"*. User 60 also commented to this effect regarding technical difficulties, commenting *"very slow"*.

Secondly, it is not unreasonable to suggest that the positive evaluation of the ReplaySuite may in part be attributable to bias from participants (4/9) who also previously participated in the VPS validation study. Pathologists who participated in both VPS and ReplaySuite studies expressed greater satisfaction and confidence in the VPS during the VPS validation study than pathologists who participated only in the VPS study. However during the ReplaySuite study, these participants who participated in both did not provide more favourable evaluation of the ReplaySuite than participants who only participated in the ReplaySuite study. In Figure 5.1, the participant who considered the ReplaySuite the most user-friendly did not participate in the VPS evaluation study. In Figure 5.2, all participants who considered the ReplaySuite of greatest benefit in training were participants who did not participate in the VPS evaluation study,

and both participants and non-participants in the VPS evaluation study considered the ReplaySuite of great benefit in quality assurance studies (Figure 5.3).

It also evident that those who participate in studies of this nature are often early adapters and often possess a positive bias towards new technology. This is an issue when evaluating any new software, and in that context, bias is unavoidable. 20 of the 70 participants invited to participate had foreknowledge of the VPS, however none had previously seen the ReplaySuite and, therefore, had no preconceived notions about the software.

The benefits of various online tools and their impact on diagnostic performance have already been highlighted, however, the ReplaySuite possesses a number of practical advantages. Many online tools present the diagnostic opinion of one pathologist. In contrast, the ReplaySuite can replay examinations of the same slide by multiple experts, illustrating a number of different diagnostic pathways that corroborate the same diagnostic hypothesis. Alternatively, reviewing pathologists may observe examinations that disagreed with group consensus, in order to identify the possible sources of disagreement. This is of particular interest for disorders that suffer from a high degree of inter-observer variability. Additionally, interactive tutorials and annotated digital slides require considerable time to create, however, individual authoring time with the VPS/ReplaySuite is around 6 minutes, per user, per slide.

Figure 5.2 & Figure 5.3 illustrate that the more that participants used the system, the greater potential benefit they perceived it having in pathology training and quality assurance. All 3 participants (Users 10,55,74) who replayed more than ten examinations considered the ReplaySuite of some or of great potential benefit in pathology training and quality assurance. While small sample size precludes the significance of a relationship between heavy use and favourable perception, it is not unreasonable to suggest that participants who fully appreciated the capabilities of the ReplaySuite will possess a more considered opinion.

The results of this preliminary evaluation demonstrated that experienced pathologists considered the ReplaySuite a useful pathology tool, and of significant potential in both training and EQA. The ability to replay archived virtual slide examinations lead to 2 participants reassessing their original (discordant) diagnoses, and while this cannot be considered to validate its capability for improving performance, it does suggest that the ReplaySuite may be used in improving scoring uniformity and reducing interobserver variability.

Kronz *et al* (2000) illustrated the potential of web-based applications for improving diagnostic performance in histological assessment. In order to evaluate whether use of the ReplaySuite could produce similar results, a controlled study was required that allowed comparison of pre and post ReplaySuite use diagnostic performance. The following Chapter describes a three-phased EQA study, involving two scoring sessions partitioned by a phase in which participants reviewed electronic resources, including the ReplaySuite.

Chapter 6: EQUALIS External Quality Assurance Study In Chronic Hepatitis

6.1 Introduction

Ensuring high standards of practice in medicine is of crucial importance to both the medical profession and general public. This is no less important in pathology, where the pathology report is frequently the basis for patient treatment, and misdiagnosis can critically affect patient care. Histopathology requires well-developed searching, perception and classification skills (Crowley *et al.*, 2003), however these subjective skills are susceptible to human error. Cases in which the pathological diagnoses are not clear-cut are common, and many studies have reported observer variability in histological assessment (Rakovitch *et al.*, 2004, Rousselet *et al.*, 2005, Winkfield *et al.*, 2003, Krieger *et al.*, 1994, Cocker *et al.*, 1968, Hanby *et al.*, 1992, Teasdale, 1996, Hastrup *et al.*, 1994, Bajema *et al.*, 1996).

Inter-observer variability is unavoidable in histopathology, given its reliance on subjective human interpretation, however histopathologists seek to reduce errors to their lowest attainable frequency through audit and continuing education. It has been stated that while other pathology disciplines have national and international standards against which laboratory performance can be measured (Snell and Brown, 2001, Legg and Hurrell, 1984), the 'correct' diagnosis in histopathology is not always well defined, making precise technical standards harder to establish (Ramsay, 1999). The utilisation of ordinal numeric labels, such as in hepatic staging and grading, is also likely to increase inter-observer variability, requiring pathologists to categorise along an underlying biological continuum (Goldin *et al.*, 1996, Metavir group, 1994, Westin *et al.*, 1999). Additionally, the more categories a classification system contains, the higher the levels of inter-observer variability observed (Goldin *et al.*, 1996, Gronbaek *et al.*, 2002, Westin *et al.*, 1999).

Reducing errors to the lowest achievable frequency and severity requires continued effort on the part of histopathologists. Participation in external quality assurance (EQA) schemes provides a means of assessing standards. In the UK, the United Kingdom Accreditation Service (UKAS) is the sole national

accreditation body recognised by government to assess, against standards set by the European (CEN) or international (ISO) standardisation bodies, organisations that provide certification, testing, inspection and calibration services (UKAS, 2005). Clinical Pathology Accreditation (UK) Ltd (CPA), in partnership with UKAS, provide a means to accredit Clinical Pathology Services and EQA.

Conventional EQA schemes, such as the Scotland and N Ireland General Histopathology EQA (CPA EQA Ref No:047/0143) and National Liver Histopathology EQA (CPA EQA Ref No:072/0171) schemes, involve the distribution of glass slides to participants, with subsequent collation and analysis of submitted diagnoses. However this process is time consuming, incurs risk of loss or damage to glass slides, and precludes the use of biopsies containing small quantities of tissue. The use of telecommunication systems as a means of delivering biopsy images provides the potential to overcome these issues in EQA. This has been previously been exploited for EQA work in other organs such as breast through the use of libraries of pre-selected field images (Ellis *et al.*, 2005, European Working Group for Breast Screening Pathology, 2005), and digital slides (Medical Solutions PLC, 2005), however, while digital images have been used for the quantification of fibrosis in liver biopsies (Dahab *et al.*, 2004, Caballero *et al.*, 2001, O'Brien *et al.*, 2000), their use in liver quality assurance programs has been limited (University of Leeds, 2005).

The use of virtual slide technology (Demichelis *et al.*, 2002, Fujita and Crowley, 2003, Dee *et al.*, 2003) for EQA schemes possesses a number of intrinsic advantages over traditional use of glass slides alone:

- Utilisation of biopsies with limited tissue, such as needlecore biopsies.
- Examination of the same slides by multiple examiners in different locations, at the same time.
- Removal of inherent problems associated with physical delivery of glass slides. Images delivered via Internet are not at risk of physical loss or damage

- Electronic submission of scoring, removing inherent problems and overheads associated with physical delivery of paperwork.
- Electronic surveying may be linked to a database, enabling extensive, rapid interrogation of study data, and the reduction of error associated with data input. Also allows real-time analysis of the dataset.
- Examinations may be tracked, allowing the diagnostic process used during an examination to be evaluated, and potential sources of diagnostic error to be identified

Hepatitis, inflammation of the liver, is caused by a number of factors such as viral infection, alcohol or drug abuse. Viral Hepatitis is sub-categorised (A, B, C, D and E) according to the viral pathogen strain (HAV, HBV, HCV, HDV and HEV respectively) (CDC, 2005). Hepatitis C virus (HCV) is the most common cause of liver hepatitis, and infects about 170 million people worldwide, and 4 million in Europe (WHO, 1997). Its prevalence in Sweden is estimated to be 0.1-0.5%, corresponding to approximately 40,000 people (Wejstal *et al.*, 2000). Chronic hepatitis refers to inflammation lasting six months or longer. Most HCV infected individuals' become chronic carriers, of which some 20% develop liver cirrhosis after 20/30 years. Approximately one-third remain asymptomatic with only mild liver injury, while 15/40% spontaneously resolve their acute HCV infection. (Wejstal *et al.*, 2003).

While serum HCV RNA detection by qualitative PCR testing is indicative of ongoing infection, further evaluation is required to assess disease progression and appropriate course of treatment. While non-invasive methods have been proposed (Afdhal, 2003), histological reporting of liver biopsy remains the gold standard for assessing the severity of chronic hepatitis. Classification systems differ in complexity (Batts and Ludwig, 1995, Knodell *et al.*, 1981, Scheuer, 1991, Ishak *et al.*, 1995), but each assess two key parameters; inflammation and fibrosis. Fibrosis is the formation of scar tissue after injury; Lesions of fibrosis and parenchymal remodeling are referred to as 'stage' and indicate long-term disease progression. Necroinflammation (grade), meaning cell death and inflammation, is not only a measure of severity, but also of ongoing disease

activity, and the parameter most potentially responsive to therapy. Grade may fluctuate with disease activity or therapeutic intervention, but stage is considered relatively constant (Scheuer, 2003). Figure 6.1 and Figure 6.2 depict Batts and Ludwig's (1995) scoring system. Investigators acknowledge the distinctions between the lesions of panels Figure 6.2 C and D in liver biopsy material as an area of potential difficulty, illustrating the inherent difficulties associated with histological assessment of liver biopsies.

In 1999, a Swedish national expert panel developed national guidelines for treatment of chronic hepatitis C, recommending Interferon (IFN) in combination with ribavirin as standard treatment for liver biopsies showing fibrosis Stage ≥2 and ongoing inflammation, or fibrosis stage 1 and inflammation Grade ≥2 (Wejstal *et al.*, 2000). The subsequent introduction of pegylated (peg) IFN resulted in the panel recommending combination treatment with peg-IFN and ribavirin as standard treatment (Wejstal *et al.*, 2003). Adverse effects during combination therapy are common and result in discontinuation of treatment in up to 15% of patients, and dose reductions in more than 20% of patients (Wejstal *et al.*, 2000).

In order to exploit the advantages of using digital slides for liver EQA, a scheme was commissioned with the primary objective of ascertaining the level of consistency in grading and staging of chronic hepatitis in Sweden. The scheme was commissioned by EQUALIS (External Quality Assurance in Laboratory Medicine in Sweden), an accredited provider of external quality assurance (EQA) schemes in Sweden (EQUALIS, 2005). In contrast to conventional liver EQA schemes, all biopsies were presented digitally; examination and scoring performed via the Virtual Pathology Slide (VPS) system.

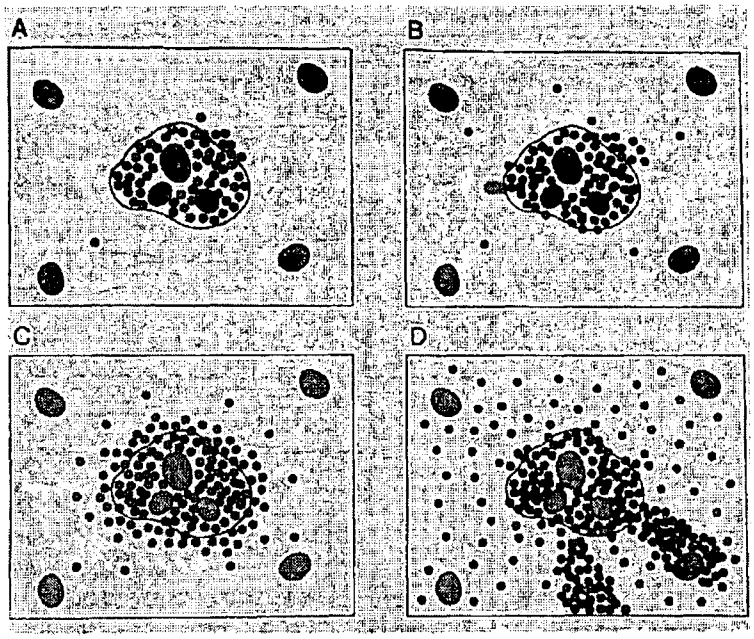


Figure 6.1 In these four panels, the increasing severity of portal inflammation, interface hepatitis, and lobular necroinflammatory lesions in chronic hepatitis are shown. Reprinted from Batts and Ludwig (1995)

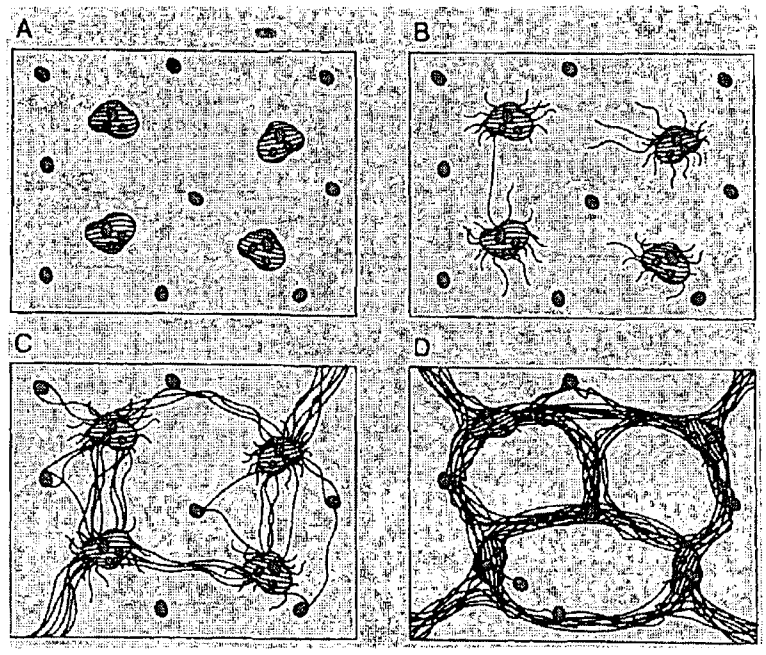


Figure 6.2 These panels graphically portray the progression of fibrosis from portal expansion (A) through septal fibrosis (B and C) to complete cirrhotic remodeling (D). Reprinted from Batts and Ludwig (1995)

In addition to evaluating the uniformity of chronic hepatitis scoring, a secondary objective of the scheme was to evaluate the 'learning effect' of conducting two scoring sessions, with an intermediate session providing electronic reference material. Studies have shown the use of online training tools can enhance the development of diagnostic skills (Heidger *et al.*, 2002, Dee *et al.*, 2003). The complexity of such resources can vary from static images to virtual slide libraries, however the most common consist of interactive tutorials, displaying a limited number of pre-selected images per case, often with accompanying notation (Kronz *et al.*, 2004, College of Medicine University of Illinois at Urbana-Champaign, 2005, Kronz *et al.*, 2000, University of Kansas Medical Center, 1996, Mitros, 1996). Two resources were provided to participants. The first comprised an online library of pre-selected static fields with accompanying annotation. The second was the ReplaySuite.

Finally, subsequent to study completion, an expert pathologist reviewed examination replays via the ReplaySuite in order to potentially identify sources of error. Examinations were evaluated on the basis of examination technique, and where possible, a source of error (missed pathology or misinterpretation of observed pathology) was identified.

The primary objective of this Chapter was to evaluate participant diagnostic performance in relation to gold standard staging and grading before and after utilisation of the available electronic resources, in order to ascertain whether an improvement was observed. The secondary objective was to establish whether an expert liver pathologists assessment of participant examinations via the ReplaySuite could be used to identify sources of error that contributed to discordant diagnosis.

6.2 Study Design

6.2.1 Case Selection

Liver needlecore biopsies were provided by the KVAST group (Quality and Standardisation) for liver disease. The group consists of 4 members, each a Swedish pathologist with a special interest in liver pathology, appointed by the Swedish Society of Pathology.

From approximately 80 cases collected during their routine work, the KVAST group selected 20 to include in the study, which covered the different grades and stages. These were separated into two groups (10 cases per group), both groups being of similar diagnostic difficulty. Each case consisted of 2 glass slides from the same needlecore biopsy; one stained with Haematoxylin and Eosin, the other with Sirius red or van Gieson's stain. A 'gold standard' stage and grade was provided by the KVAST group for each case, based on group consensus after examining digital slides for each case. The Batts and Ludwig classification system was used by participants to score (Brunt, 2000, Batts and Ludwig, 1995).

6.2.2 Study Participants

Participants consisted of 26 Swedish pathology laboratories. Participants were required to examine cases as a group, and submit scoring according to group consensus.

6.2.3 Study Architecture

The study consisted of 3 Phases:

- **Phase I:** Examination and scoring of 10 digitised needlecore liver biopsies via the VPS. Participants were not permitted to re-examine cases once scored.
- **Phase II:** Utilisation of electronic resources.

- The ReplaySuite was used to review Phase I gold standard and group concordant scoring data, replay examinations and SnapShots.
- The Swedish Society of Pathology reference library provided access to reference images of pathological changes associated with different scores.
- **Phase III:** Examination and scoring of 10 digitised needlecore liver biopsies via the VPS, and continued utilisation of the ReplaySuite.

6.2.4 Electronic Resources

Participants were provided with two electronic references during Phase II. Accessible online, these two resources were the ReplaySuite 2.0 and the Swedish Society of Pathology Image reference library.

6.2.4.1 ReplaySuite 2.0

The ReplaySuite version 2.0 was provided to participants to enable them to review Phase I data. Chapter 4 has previously described the ReplaySuite (2.0) in detail. Subsequent references to ReplaySuite in this Chapter refer to ReplaySuite Version 2.0.

6.2.4.2 Swedish Society of Pathology Reference Image Library

The library provides a number of pre-selected images, representative of visual features associated with particular staging and grading scores, at a range of magnifications (Swedish Society of Pathology, 2004). Images were not interactive; navigation within images and increasing/decreasing magnification is not possible. The library contains over 60 images, and provides reference images for all stage and grades. Figure 6.3 illustrates an example of a reference images displayed by the website.

Figure 6.3 Example of a reference image concordant with Stage 1, as displayed on the Swedish Society of Pathology Reference Image Library.

6.2.5 Statistical Methods

Cohen's kappa is a measure of agreement between two sets of data. Kappa values are measured between 0 and 1, 1 indicating 'complete' agreement, and 0 indicating 'no agreement better than chance' agreement (Cohen, 1960). Negative values indicate 'worse than chance' agreement. Kappa statistics may be weighted or un-weighted; Un-weighted kappa only measures the level of exact agreement, whereas weighted kappa also acknowledges close, but not exact agreement. The use of un-weighted kappa alone only provides the level of complete concordance between participants and KVAST scoring. The inclusion of linear weighted kappa provides an indication of how close participants were to achieving concordance with KVAST, of particular benefit considering the low percentage of concordance observed in Phase I. Kappa statistics were calculated using MedCalc™ (Schoonjans, 2005). Un-weighted kappa is calculated using the following equation:

$$Kappa = \frac{ObservedAgreement - ChanceAgreement}{TotalObserved - ChanceAgreement}$$

While linear weights were calculated as follows:

$$w_i = 1 - \frac{i}{k-1}$$

Where k is the number of categories (i.e. 0-4), and i the degree of difference (ie 0-4). With 5 categories, the weights in the linear set are 1, 0.75, 0.50, 0.25 and 0, when there is a difference of 0 (=total agreement) and 1, 2, 3 and 4 staging categories respectively. Table 6.1 below provides a guide to interpreting level of agreement using kappa (Landis and Koch, 1977).

It has been commented upon (Byrt, 1992) that kappa may not be the most appropriate method of quantifying 'levels of agreement', as low kappa ratings may be achieved despite high observed agreement. The magnitude of the kappa coefficient represents the proportion of agreement greater than that expected by chance. The interpretation of the coefficient, however, is not so straightforward, as there are other factors that can influence the magnitude of the coefficient, or the interpretation that can be placed on a given magnitude. Two such factors that can influence the magnitude of kappa are prevalence (distribution) and bias (level of disagreement). Populations with an uneven distribution can achieve lower kappa than those with equal distribution, even if levels of agreement are the same. In contrast, when a large bias exists, kappa is greater than when there is low or no bias. The lack of adjustment of such factors prohibits comparison of kappa across studies, procedures or populations (Sim and Wright, 2005).

It is possible to correct for prevalence and bias using PABAK (Prevalence Adjusted Bias Adjusted Kappa), however Hoehler (Hoehler, 2000) is critical of the use of PABAK because he believes that the effects of bias and prevalence on the magnitude of kappa are themselves informative and should not be adjusted for. Thus, the PABAK could be considered to generate a value for kappa that does not relate to the situation in which the original ratings were made.

Despite the shortcomings however, kappa statistics are firmly rooted in medical literature, and remain the statistic of choice to assess inter-observer variability.

Table 6.1 Strength of agreement between two sets of data indicated by kappa statistic values

Kappa Value	Level of agreement
< 0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1	Very Good

6.3 Monitoring Electronic Resource Use and Impact on Performance

In order to ascertain whether electronic resource use was associated with improved participant performance, it was required to determine whether the two electronic resources were utilised, and evaluate levels of use for each.

6.3.1 Monitoring ReplaySuite Use

ReplaySuite use was monitored as described in Chapter 4, providing a means of identifying light, medium and heavy users during the study. Table 6.2 illustrates participant use patterns, displaying the number of individual functions called and the total number of ReplaySuite functions called.

6.3.2 Society of Swedish Pathology Website

Participants were surveyed subsequent to the study to determine their perceptions of the Swedish society of Pathology image library. The survey asked participants if they used the image library, and if so, how useful did they find it on a five-point Likert scale: not at all, not very, neutral, a little, very. While it is difficult to quantify the degree of activity the image library experienced based on the post-study survey results, it does provide a rough indication of 'heavy' users.

6.3.3 Comparison of Heavy and Light Resource User Performance

A greater improvement in staging was observed in heavy ReplaySuite users than other users, where heavy users are those that used the ReplaySuite ≥ 21 times. Mean un-weighted, and linear weighted staging kappa improved for 'heavy' ReplaySuite users by 0.357 (+0.326 Vs –0.031) and 0.223 (+0.159 Vs –0.064) more than 'non-heavy' users respectively. However, at the same time, negligible difference was observed in grading based on ReplaySuite use, with heavy users improving by 0.064 (0.301 Vs 0.238) and 0.1 (0.328 Vs 0.229) for un-weighted, and linear weighted grading kappa respectively. With such small sample sizes, it is difficult to make assumptions based upon the data available.

Table 6.2 Use of the VPS and ReplaySuite by users in Phases I-III. Phase I & III values refer to the number of cases scored. Phase II sub-category values (Replays, SnapShots, Graphs, Lists) refer to the number of times each function was utilised, and 'Any Use' relates to the sum of all ReplaySuite sub-categories.

Users	Phases						
	I	II					III
	Exams	Replays	SnapShots	Graphs	Lists	Any use	Exams
5	10		10		2	12	10
6	10	1	6	5	13	25	10
7	10						
8	10						
9		1	2		14	17	
10					3	3	
11	10	7	6		8	21	10
13	10	12		3	16	31	10
14	10	3			24	27	10
15	10		16		6	22	10
17	10	1	7	10	5	23	10
19	10	3			6	9	10
20				4	20	24	
22	1		2		3	5	10
23	10	4	4		31	39	9
25	10						
26					1	1	
28	10				7	7	10
29	10						
30	10		1		9	10	5
40					1	1	
48			1	2	8	11	
84	10						
Total no. Participants	**17**	**8**	**10**	**5**	**18**	**18**	**12**
Total no. functions	**161**	**32**	**55**	**24**	**177**	**288**	**114**

Table 6.3 Participant use of electronic resources and improvements in un-weighted and linear weighted kappa statistics for staging and grading between scoring sessions

User ID	Image Library use	User perception of library's helpfulness	Frequency of *ReplaySuite* function use	Kappa improvement between scoring phases			
				Stage		Grade	
				Un-weighted	Weighted	Un-weighted	Weighted
5	N/A*		12	0.174	0.047	0.04	0.086
6	Yes	A little helpful	25	0.769	0.606	-0.054	0.077
11	Yes	A little helpful	21	0.273	0.11	0.489	0.525
13	Yes	Very helpful	31	0.713	0.367	0.453	0.326
14	Yes	A little helpful	27	-0.076	-0.105	0.43	0.185
15	No		22	0.591	0.297	0.384	0.751
17	Yes	N/A*	23	-0.316	-0.32	0.106	0.106
19	N/A*		9	0.123	0	0.31	0.236
28	Yes	Very helpful	7	-0.391	-0.238	0.363	0.364

* Do not remember

6.4 System Review

In order to complete each summary report form submitted at the end of an examination, participants were required to evaluate the quality of VPS images and system speed, using a 4-point scale. While omission of a neutral 'middle point' in the evaluation scale introduced the potential for positive bias, the scale was used to maintain evaluation criteria established during validation of the original VPS (Costello, 2004).

6.4.1 Image Quality

Figure 6.4 illustrates user perceptions of VPS images for all examinations. Less than 10% of cases were considered to have poor images quality, and almost 40% of cases were deemed to have good or excellent quality images. Perceptions of image quality did not differ significantly between DVD and remote server sources.

6.4.2 System Speed

Participants submitted a rating of VPS system speed with each summary report form, using a 4-point scale. As Figure 6.5 illustrates, overall, participants considered the majority of VPS examinations to be slow (67%). Participants using slow Internet connections in particular may have found the system unacceptably slow, due to the time required to retrieve images from the remote server. 25% of Phase III examinations utilised DVD/CD-ROM images, and for those who downloaded images from both DVD and web sources, the majority of examinations using DVD's (57%) were of acceptable speed. With regards to speed of lateral navigation, light microscopy will always outperform digital slides. Accordingly, using light microscopy as a reference when evaluating system speed will always disfavour digital slides.

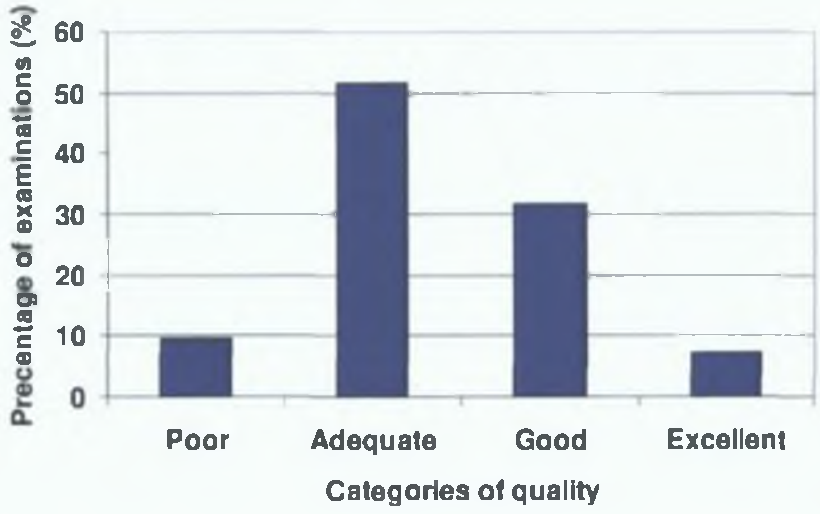


Figure 6.4 Participant perceptions of image quality when using the VPS to examine cases

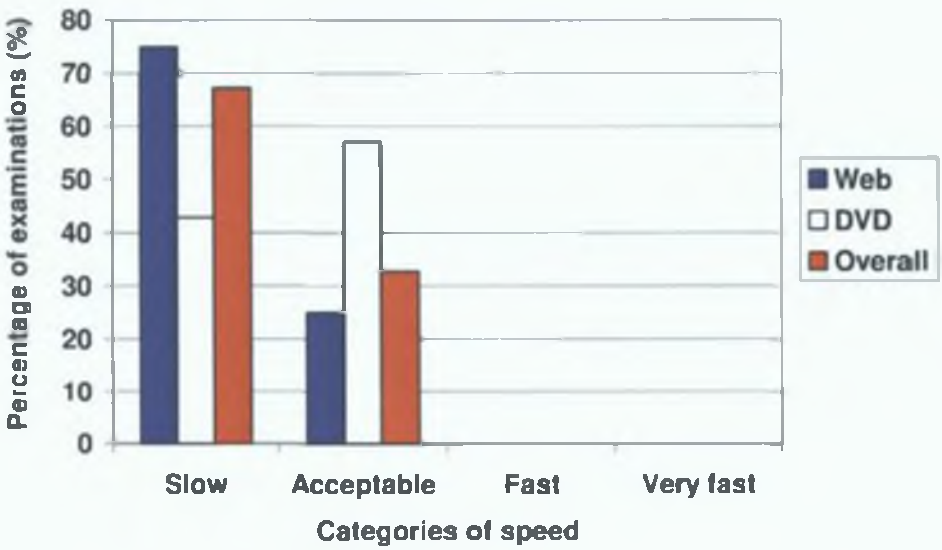


Figure 6.5 Perceptions of VPS speed for all participants, and perceptions of participants who downloaded images from both DVD and Web sources

6.5 Comparison of Phase I and Phase III Staging Performance

Participant staging was assessed for both the entire population, and for participants that completed the study by examining and scoring all cases in both scoring phases.

6.5.1 Comparison of Phase I and Phase III Staging for Entire Population

Improvement in performance was evaluated for both the entire sample population, and the sub-population that examined all cases and reviewed the electronic resource material. Staging and grading were assessed independently. Table 6.4 also displays the percentage of all participant examinations that concurred with KVAST group staging, for each case. Mean consensus with KVAST staging increased from 49% to 69% (+20%), indicating that participant concordance with KVAST grading improved considerably.

It should be highlighted that non-uniform numbers of participants submitted scores; 17 scored cases in Phase I (161 examinations: 16x10 cases and 1x1 case) while 12 scored Phase III cases (114 examinations: 10x10cases, 1x9 cases and 1x5 cases), making direct comparison of mean consensus agreement with KVAST difficult.

6.5.2 Distribution of Staging with Respect to Gold Standard

As Figure 6.6 illustrates, between scoring sessions:

- Overall concordance with KVAST staging increased by 20% (49% - 69%).
- The percentage of discordant examinations staging within one category of KVAST stage decreased from 48% to 25%.

Increase in discordant staging within 2 or more was negligible, increasing from 2.5% to 5%, suggesting that while performance improved, improvement was in stage 'fine-tuning.'

In particular this modification may be perceived as correcting a predilection to over-stage in Phase I, as illustrated in Figure 6.6; almost 38% of Phase I examinations staged 1 higher than KVAST. In Phase III, this Figure dropped to 7%. In responding to this trend of over-staging, a number of participants may have over compensated. While the number of examinations that over-staged decreased from 40% to 8%, under-staging increased from 11% to 23%.

Table 6.4 KVAST and participant group consensus staging for Phases I & III.

Phase I					Phase III				
		Group Consensus		Percentage			Group Consensus		Percentage
Case	KVAST Stage	Stage	Percentage Consensus	Agreement with KVAST	Case	KVAST Stage	Stage	Percentage Consensus	Agreement with KVAST
10	2	2	82	82	12	2	2	92	92
2	3	3	75	75	15	2	2	92	92
7	4	4	69	69	11	4	4	91	91
3	2	3	63	38	16	3	3	83	83
6	2	3	63	38	13	0	0	82	81
8	1	2	63	19	14	1	1	75	75
9	3	4	63	31	18	2	1	55	36
4	1	1	56	56	20	2	2	55	55
1	0	0/1	44	44	17	3	2/3	45	45
5	1	1	38	38	19	3	2/3	36	36
	Mean		**62**	**49**		**Mean**		**71**	**69**

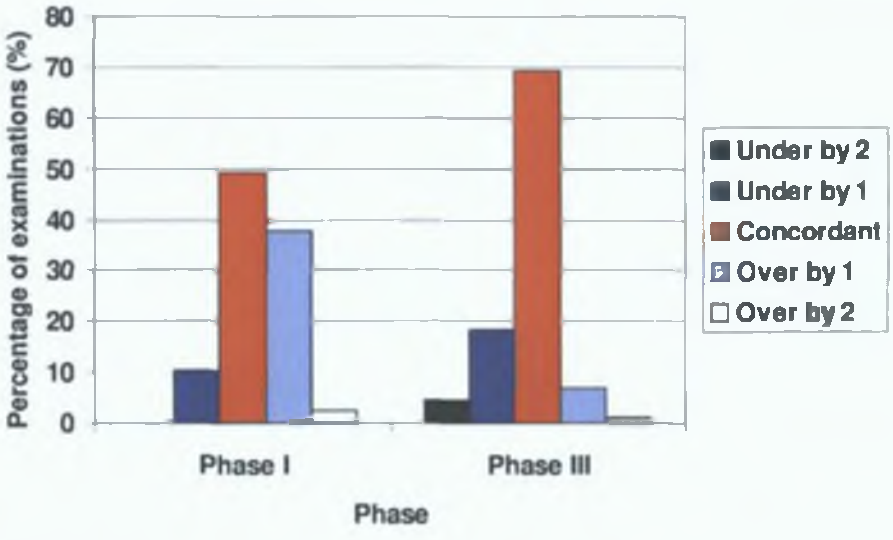


Figure 6.6 Comparison of Phase I & III distribution of staging with respect to KVAST staging for all examinations

A non-uniform distribution of KVAST scoring occurred within each phase, for example in Phase I, KVAST scored 3 cases as Stage 2, whilst only scoring 1 case Stage 4. When the distribution of staging for all examinations is normalised, the percentage concordance/discordance relative to specific stages can be identified. In Table 6.5 & Table 6.6, each row is associated with a specific KVAST score. The percentage distribution of participant scoring for cases KVAST scored accordingly is indicated in each cell within the row. The diagonal indicates the percentage of examinations in which participant scoring and KVAST scoring agreed.

As Table 6.5 illustrates, the highest Phase I staging concordance was in cases KVAST staged 4 (69%), while the lowest was for cases KVAST staged 1 (38%). Additionally, the greatest spread of participant staging occurred for cases staged 1 by KVAST. As might be expected, the greatest level of over-staging was for cases KVAST staged 0, whilst the greatest level of under-staging occurred for those staged 4.

As Table 6.6 shows, the highest staging concordance in Phase III was again for cases staged 4 by KVAST (91%). In contrast to Phase I, however, the lowest concordance was in cases KVAST staged 3 (56%). As with Phase I, the greatest incidence of over-staging occurred with cases staged 0 by KVAST, however the highest level of under-staging was for Stage 3 cases, which also had the greatest spread of participant staging.

Table 6.5 Normalised percentage distribution of participant staging in Phase I with respect to KVAST staging, for all examinations.

	Participant Staging				
KVAST staging	**0**	**1**	**2**	**3**	**4**
0	**44**	44	13	0	0
1	23	**38**	35	4	0
2	0	0	**53**	47	0
3	0	0	3	**53**	44
4	0	0	0	31	**69**

Table 6.6 Normalised percentage distribution of participant staging in Phase III with respect to KVAST staging, for all examinations.

	Participant Staging				
KVAST Staging	**0**	**1**	**2**	**3**	**4**
0	**82**	9	9	0	0
1	8	**75**	17	0	0
2	2	22	**70**	6	0
3	0	12	26	**56**	6
4	0	0	0	9	**91**

6.5.3 Comparison of Phase I and Phase III Staging for Participants to Complete the Study

All participants that completed the study utilised the ReplaySuite. Reviewing only participants to examine all cases in both phases permits direct comparison of mean group consensus with KVAST between scoring phases. As Table 6.7 illustrates, mean group consensus with KVAST increased from 49% to 67% (+18%), while mean group consensus (irrespective of KVAST agreement) improvement was marginal; +5% (62% - 67%). While group consensus staging agreed with KVAST in only 50% of cases in Phase I, it agreed for all cases in Phase III.

As Figure 6.7 illustrates, between scoring sessions, concordance with KVAST staging increased by 18% (49% - 67%). This increased concordance is attributed primarily to a 20% (49% - 29%) decrease in discordant staging within 1 of KVAST staging. In particular, participants tended to be over cautious in their Phase I scoring, as indicated by 41% of examinations over-staging by 1 with respect to KVAST, which decreased to 9% in Phase III. Fluctuation in discordant staging by 2 or more was negligible, increasing 2% (2% - 4%).

Table 6.7 KVAST and group consensus staging for Phases I & III, for participants who completed the study.

Phase I					Phase III				
Case	KVAST Stage	Group Consensus Stage	Percentage Consensus	Percentage Agreement with KVAST	Case	KVAST Stage	Group Consensus Stage	Percentage Consensus	Percentage Agreement with KVAST
7	4	4	78	78	11	4	4	89	89
10	2	2	78	78	12	2	2	89	89
2	3	3	67	67	15	2	2	89	89
4	1	1	56	56	13	0	0	78	78
6	2	3	56	44	16	3	3	78	78
5	1	1	44	44	14	1	1	67	67
1	0	1	56	33	17	3	2/3	44	44
8	1	2	56	33	18	2	1/2	44	44
9	3	4	56	33	19	3	2/3	44	44
3	2	3	78	22	20	2	2	44	44
	Mean		**63**	**49**		**Mean**		**67**	**67**

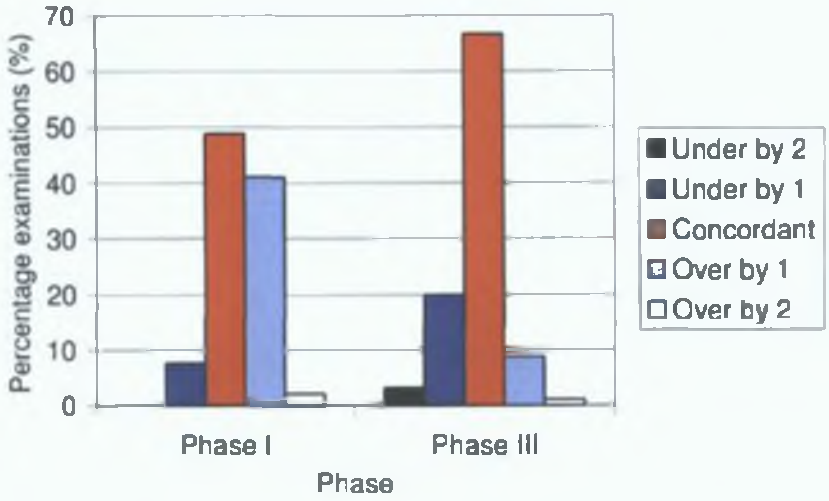


Figure 6.7 Comparison of Phase I & III distribution of staging with respect to KVAST staging by participants who completed the study

As previously indicated, non-uniform distribution of KVAST scoring occurred within each phase. When normalised, as with all participants, the highest Phase I staging concordance was in cases KVAST staged 4 (78%), and lowest KVAST staged 0 (33%), as Table 6.8 illustrates. As with the entire population, the greatest spread of participant staging occurred for cases staged 1 by KVAST, greatest level of over-staging for cases KVAST staged 0, and under-staging for cases staged 4.

As Table 6.9 shows, the highest staging concordance in Phase III was again for cases staged 4 by KVAST (89%). In contrast to Phase I, however, the lowest concordance was in cases KVAST staged 3 (56%). As with Phase I, the greatest incidence of over-staging occurred with cases staged 0 by KVAST, however the highest level of under-staging was for stage 3 cases, which also had the greatest spread of participant staging.

Table 6.8 Normalised percentage distribution of participant staging in Phase I with respect to KVAST staging, for examinations performed by participants who completed the study.

	Participant Staging				
KVAST Staging	**0**	**1**	**2**	**3**	**4**
0	33	56	11	0	0
1	15	44	37	4	0
2	0	0	48	52	0
3	0	0	6	50	44
4	0	0	0	22	78

Table 6.9 Normalised percentage distribution of participant staging in Phase III with respect to KVAST staging, for examinations performed by participants who completed the study.

	Participant Staging				
KVAST Staging	**0**	**1**	**2**	**3**	**4**
0	78	11	11	0	0
1	11	67	22	0	0
2	3	22	67	8	0
3	0	7	30	56	7
4	0	0	0	11	89

6.5.4 Comparison of Phase I and Phase III Staging Kappa Statistics

A substantial improvement in mean group un-weighted kappa was observed, increasing from 0.347 to 0.554 (+0.207), or from 'fair' to 'moderate' agreement. This improvement was not reflected in linear weighted staging kappa, increasing from 0.603-0.688 (+0.085). This suggests that improvement in staging was due to 'fine-tuning' of staging with respect to KVAST staging. Figure 6.8 & Figure 6.9 illustrate the performance of users in Phases I & III via un-weighted, and linear weighted kappa statistics respectively.

Kappa is a measure of agreement between two sets of data, and the strength of agreement can be interpreted by associating values with categories, indicating the level of agreement, as indicated in Table 6.2. Figure 6.10 & Figure 6.11 graphically compare the levels of participant agreement with KVAST between scoring sessions, using un-weighted kappa and linear weighted kappa respectively. For each figure, the inner donut denotes Phase I, while the outer represents Phase III.

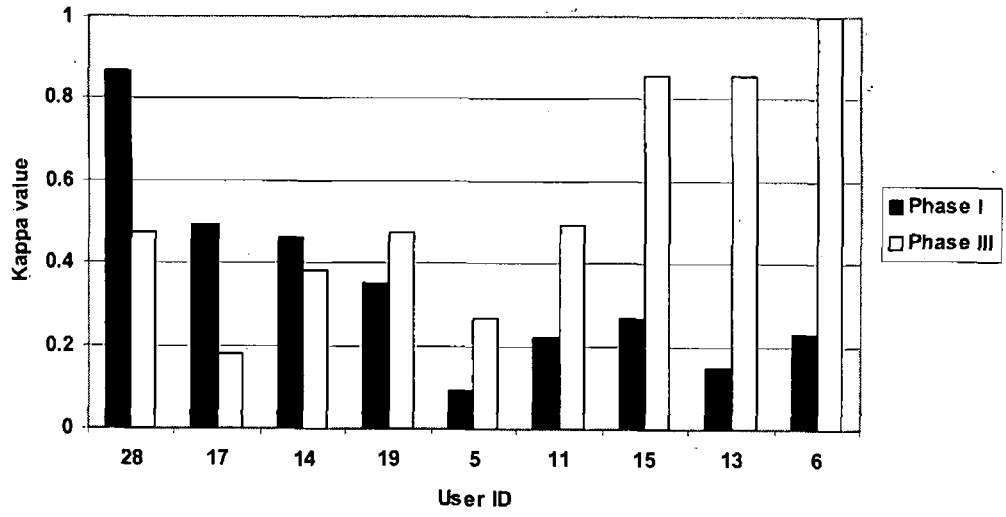


Figure 6.8 Strength of participant agreement with KVAST staging of digitised liver biopsies in scoring Phases I & III using un-weighted kappa statistics, ordered by improvement

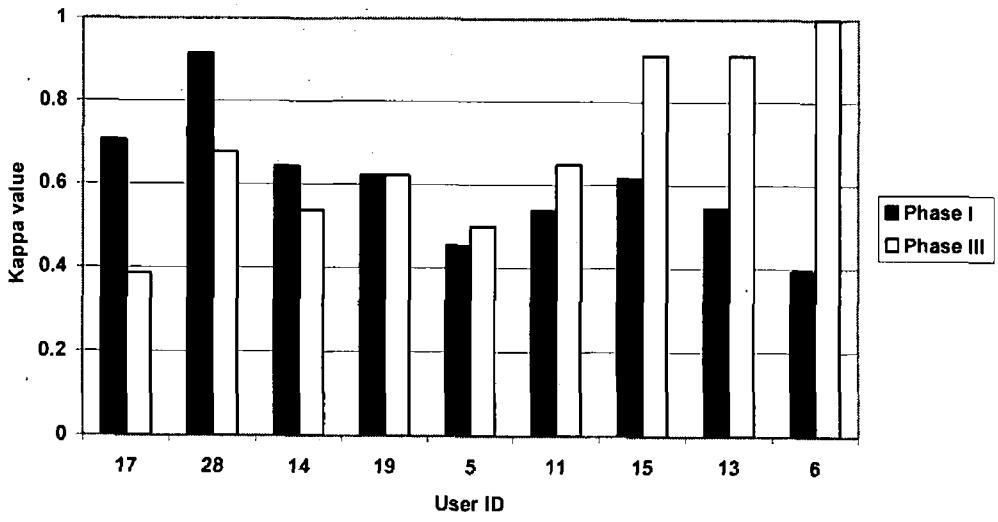


Figure 6.9 Strength of participant agreement with KVAST 'gold standard' staging of digitised liver biopsies in scoring Phases I & III using linear weighted kappa statistics, ordered by improvement

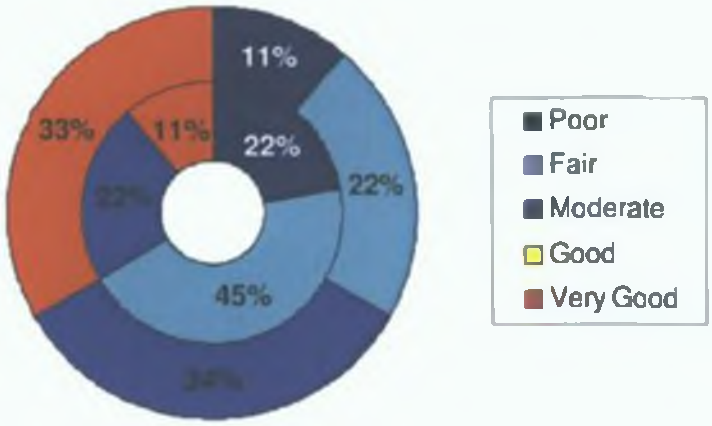


Figure 6.10 Strength of participant agreement with KVAST staging using un-weighted kappa, for participants who completed the study. Inner ring corresponds to Phase I & outer ring corresponds to Phase III.

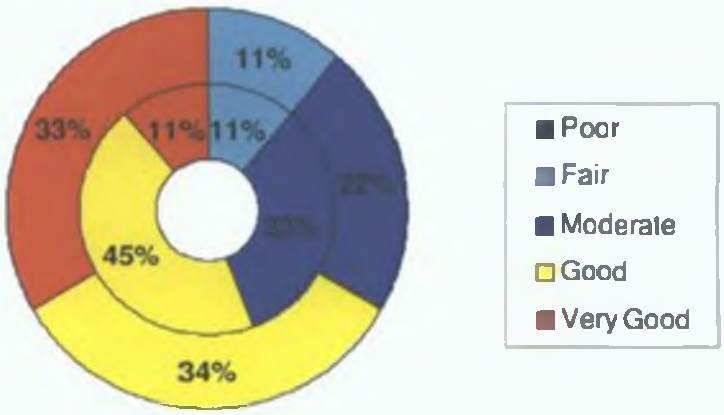


Figure 6.11 Strength of participant agreement with KVAST staging using linear weighted kappa for participants who completed the study. Inner ring corresponds to Phase I & outer ring corresponds to Phase III

6.6 Comparison of Phase I and Phase III Grading Performance

Participant grading was assessed for both the entire population, and for participants that completed the study by examining and scoring all cases in both scoring phases.

6.6.1 Comparison of Phase I and Phase III Grading for Entire Population

For all participants, mean group concordance with KVAST staging increased by 19%, while mean group consensus increased only 2%. As with staging, it is difficult to directly compare mean data for all participants, as the number of participants to examine each case varied. Table 6.10 provides a means of comparing mean group consensus, and mean group consensus with KVAST for participants who examined all cases.

Table 6.10 KVAST and participant group consensus Grading for Phases I & III

Phase I					Phase III				
		Group diagnosis		Percentage			Group diagnosis		Percentage
Case	KVAST Grading	Grading	Percentage Consensus	Agreement with KVAST	Case	KVAST Grading	Grading	Percentage Consensus	Agreement with KVAST
10	2	2	65	65	13	0	0	82	82
4	1	1	56	56	19	4	4	73	73
8	2	2	50	50	17	3	3	64	64
5	1	1/2	38	38	20	1	1	64	64
7	3	3/4	38	38	12	2	2	58	58
2	3	4	63	31	15	2	2	50	50
1	1	2	50	25	14	1	0/1	42	42
3	2	3	63	19	11	2	2/4	36	36
9	3	4	75	19	18	2	1	64	36
6	2	3	63	6	16	3	2	50	33
		Mean	**55**	**35**			**Mean**	**58**	**54**

As Figure 6.12 illustrates, between scoring sessions, concordance with KVAST grading increased by 19%. The over-staging trend observed in Phase I was repeated with grading. 42% of examinations over-graded by 1; this exceeded the percentage of examinations that concurred with KVAST. While the percentage of examinations that discordantly graded within 1 of KVAST grading decreased from 55% to 42% (-13%), this reduction is not as substantial as that observed in staging (23%)

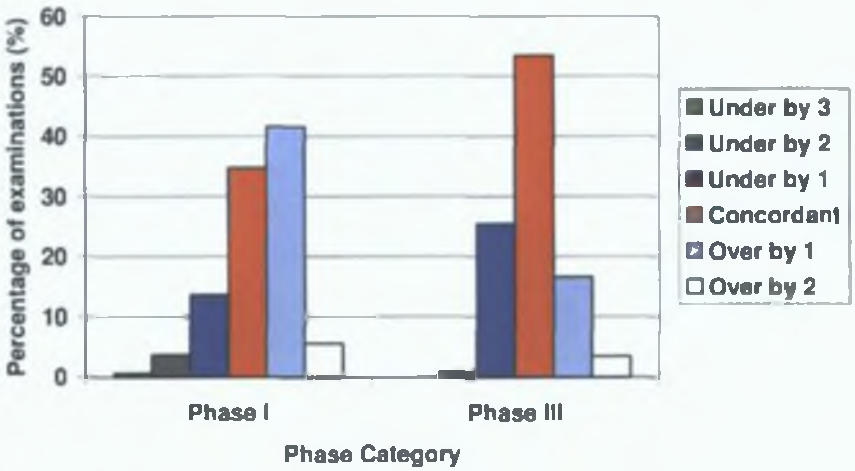


Figure 6.12 Comparison of Phase I & III distribution of grading with respect to KVAST grading for all examinations

Additionally, the reduction in the percentage of examinations discordantly grading within 2 or more degrees of KVAST decreased by 6%, as opposed to a 2.5% increase for staging. This suggests that, in comparison to staging, in which concordance increase was linked primarily with a decrease in scoring within 1, improvement in grading concordance is associated with decrease in all degrees of discordance. This suggests a slight disparity in the nature of improvement that occurred in staging and grading. While staging improvement may have been due to a 'fine-tuning' in scoring strategy, grading improvement appears to be due to a more general alignment with KVAST grading.

As Table 6.11 illustrates, the highest grading concordance occurred in cases KVAST graded 1 (40%), while the lowest was for cases KVAST graded 3 (29%). Grade 3 cases also recorded the greatest spread of participant grading. Phase I cases were graded 1-3 by KVAST, and the greatest level of over-grading was for cases KVAST graded 3, whilst the greatest level of under-staging occurred for those graded 2.

As Table 6.12 shows, the highest grading concordance in Phase III was again for cases graded 0 (82%) by KVAST, and the lowest concordance for cases graded 2 (46%). The greatest incidence of over-grading occurred with cases graded 2 by KVAST, and the highest level of under-grading was for grade 3 cases.

Table 6.11 Normalised percentage distribution of participant grading in Phase I with respect to KVAST grading, for examinations.

	Participant Grading				
KVAST Grading	**0**	**1**	**2**	**3**	**4**
0	**0**	**0**	**0**	**0**	**0**
1	17	**40**	35	8	0
2	8	15	**35**	34	8
3	2	2	8	**29**	58
4	0	0	0	0	**0**

Table 6.12 Normalised percentage distribution of participant grading in Phase III with respect to KVAST grading, for examinations.

	Participant Grading				
KVAST Grading	**0**	**1**	**2**	**3**	**4**
0	**82**	18	0	0	0
1	30	**52**	17	0	0
2	0	24	**46**	22	9
3	0	4	35	**48**	13
4	0	0	0	27	**73**

6.6.2 Comparison of Phase I and III Grading for Participants to Complete the Study

An improvement in mean group consensus with KVAST grading of 20% (34% - 54%) was observed, between Phases I & III. At the same time, mean group consensus increased negligibly, from 56% to 59 (+3%). As Table 6.13 indicates, group consensus agreement with KVAST grading increased from 4/10 to 8/10 cases.

Table 6.13 KVAST and group consensus grading for Phases I & III, for participants that completed the study. % KVAST agreement indicates the percentage of participants to agree with KVAST, irrespective of concordance

Phase I					Phase III				
Case	KVAST Grading	Group Grading	Concordance Percentage Consensus	Percentage Agreement with KVAST	Case	KVAST Grading	Group Grading	Concordance Percentage Consensus	Percentage Agreement with KVAST
4	1	1	56	56	13	0	0	89	89
7	3	3	56	56	19	4	4	78	78
8	2	2	56	56	17	3	3	67	67
10	2	2	56	56	20	1	1	67	67
9	3	4	56	33	11	2	2	44	44
1	1	2	56	22	12	2	2	44	44
2	3	4	67	22	14	1	1	44	44
3	2	3	44	22	15	2	2/3	44	44
5	1	2	44	22	18	2	1	56	44
6	2	3	67	0	16	3	2	56	22
	Mean		55.8	34.5		Mean		58.9	54.3

As Figure 6.13 illustrates, between scoring sessions, concordance with KVAST grading increased by 20%. The over-staging trend observed in Phase I was repeated with grading. 41% of examinations over-graded by 1; exceeding the

of examinations that discordantly graded within 1 of KVAST grading decreased from 53% to 41% (12%), this reduction is not as substantial as that observed in staging (20%).

Additionally, the reduction in the percentage of examinations discordantly grading by 2 or more degrees of KVAST decreased by 8% (12% to 4%), as opposed to a 2% increase for staging. This suggests that, in comparison to staging, in which concordance increase was linked primarily with a decrease in scoring within 1 of KVAST, improvement in grading concordance is associated with a decrease in all degrees of discordance. This suggests a slight disparity in the nature of improvement that occurred in staging and grading.

As Table 6.14 shows, the highest grading concordance in Phase I occurred for cases KVAST graded 3 (37%), while the lowest was for cases KVAST graded 1/2 (33%). A significant degree of scoring spread was observed in participant grading; each of the five possible grades were submitted at least once by participants, while KVAST only grades cases between 1 & 3.

As Table 6.15 shows, the highest grading concordance in Phase III was again for cases graded 0 (89%) by KVAST, and the lowest concordance for cases graded 2 (46%). The greatest incidence of over-grading occurred with cases graded 2 by KVAST, and the highest level of under-grading was for Grade 3 cases.

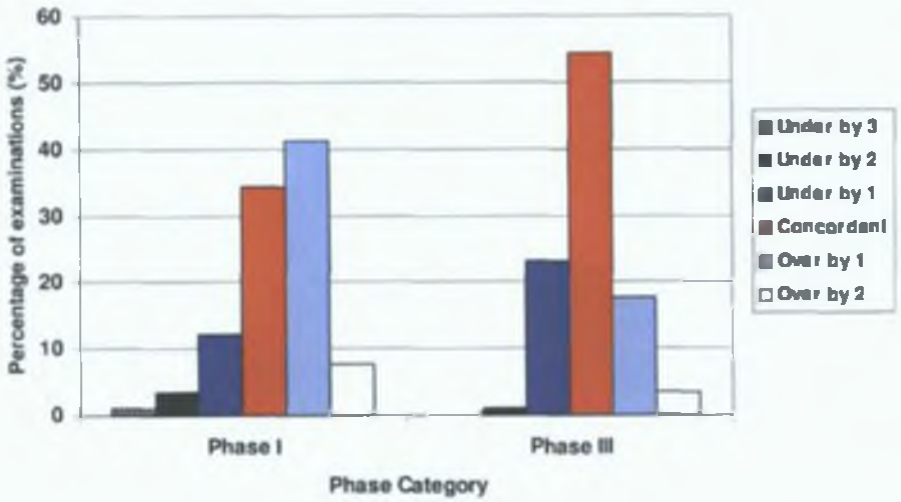


Figure 6.13 Comparison of Phase I & III distribution of grading with respect to KVAST grading for all examinations by participants to complete the study

Table 6.14 Normalised percentage distribution of participant grading in Phase I with respect to KVAST grading, for examinations performed by participants who completed the study

	Participant Grading				
KVAST Grading	**0**	**1**	**2**	**3**	**4**
0	**0**	0	0	0	0
1	11	**33**	41	15	0
2	8	17	**33**	33	8
3	4	0	7	**37**	52
4	0	0	0	0	**0**

Table 6.15 Normalised percentage distribution of participant grading in Phase III with respect to KVAST grading, for examinations performed by participants who completed the study

	Participant Grading				
KVAST Grading	**0**	**1**	**2**	**3**	**4**
0	**89**	11	0	0	0
1	22	**56**	22	0	0
2	0	25	**44**	22	83
3	0	56	33	**44**	17
4	0	0	0	22	**78**

6.6.3 Comparison of Phase I and Phase III Grading Kappa Statistics

An even more pronounced improvement in kappa values was observed in grading than staging. Mean group un-weighted kappa improved by 0.280, from 0.132 to 0.412, or from 'poor' to 'moderate' agreement. In contrast to staging, linear weighted grading kappa increased from 0.328-0.624 (+0.295), or from 'fair' to 'good' agreement. As Figure 6.14 & Figure 6.15 illustrate, improved performance was observed amongst the majority of participants, 66% (6/9) obtaining an improvement of greater than 0.3 in un-weighted kappa, and 55% (5/9) recording an increase greater than 0.23 in linear weighted kappa. In contrast to staging, no participant achieved complete concordance with KVAST grading.

Figure 6.16 categorises participants according to strength of agreement with KVAST grading (un-weighted kappa), and compares levels of agreement between scoring sessions. The inner ring denotes Phase I, while the outer ring represents Phase III. Figure 6.17 compares strength of agreement in Phases I & III using linear weighted kappa.

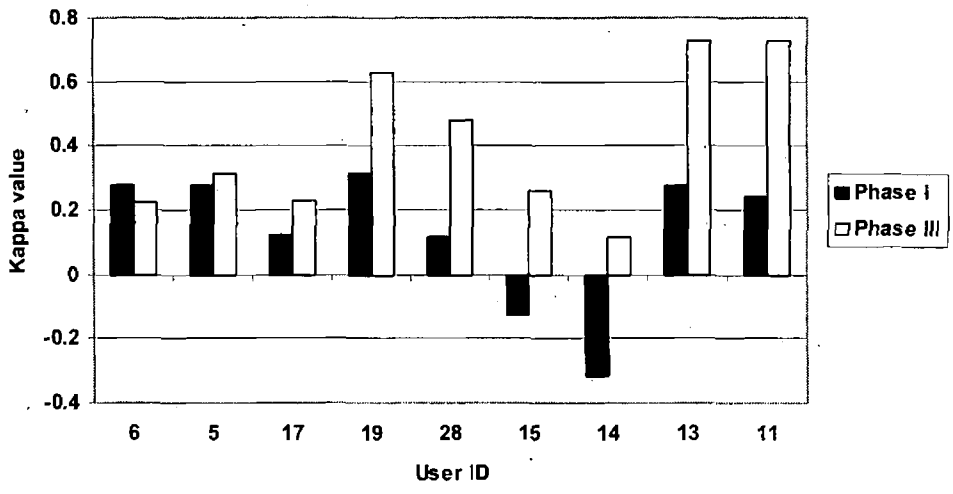


Figure 6.14 Strength of agreement with KVAST 'gold standard' grading of digitised liver biopsies in Phases I & III, using un-weighted kappa statistics, ordered by improvement.

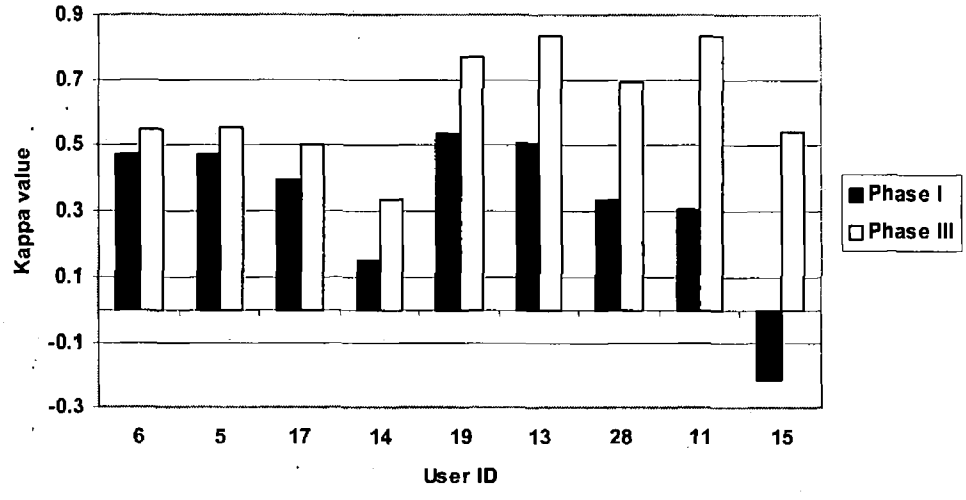


Figure 6.15 Strength of agreement with KVAST 'gold standard' grading of digitised liver biopsies in Phases I & III, using linear weighted kappa statistics, ordered by improvement.

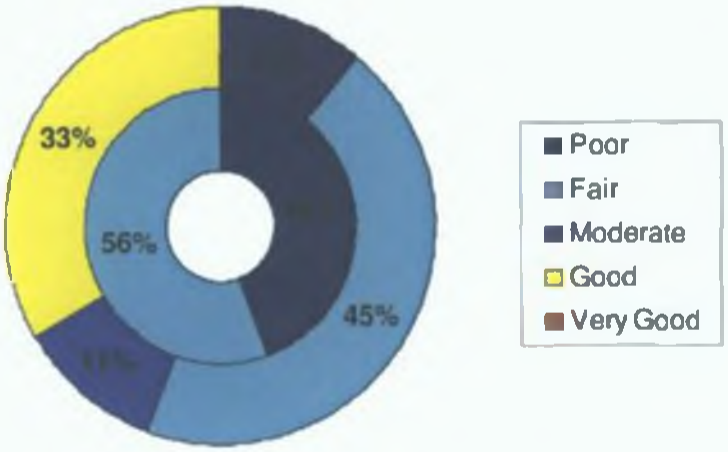


Figure 6.16 Strength of participant concordance with KVAST using un-weighted kappa. Inner ring corresponds to Phase I & outer ring corresponds to Phase III.

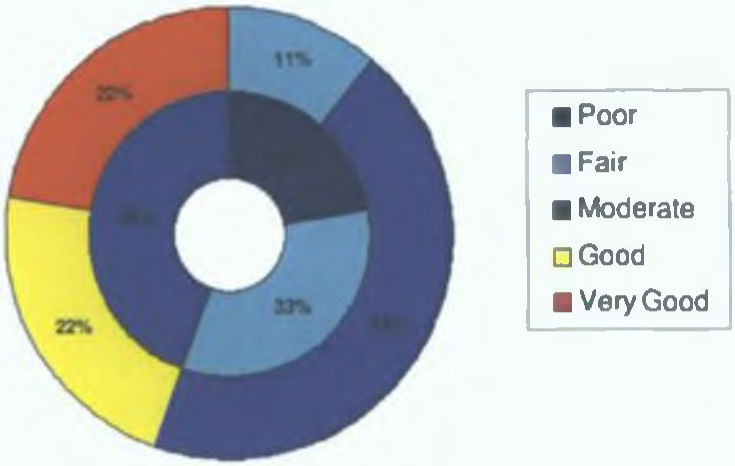


Figure 6.17 Strength of participant concordance with KVAST using un-weighted kappa. Inner ring corresponds to Phase I & outer ring corresponds to Phase III.

6.7 Expert Review of Examination Replays

Using the ReplaySuite, a representative sample of examinations from the chronic hepatitis staging and grading study in Sweden study were replayed and evaluated by an expert liver pathologist.

6.7.1 Evaluation Procedure

Examinations from both Phase I (pre-electronic resource utilisation) and Phase III (post-electronic resource utilisation) were evaluated (Table 6.16) and for a large number of participants (Table 6.17). Examinations for each case were selected randomly, with the exception of Users 11, 19 and 13. All available examinations were evaluated for these participants, as these participants recorded significant improvements in performance between scoring phases. This was conducted in order to assess whether a correlation existed between improvements in kappa and improvement in diagnostic technique.

Table 6.16 Number of VPS 2.0 examinations evaluated by the expert pathologist.

	Phase		
	I	III	Total
Total number of examinations	161	114	275
Number of examinations evaluated	51	50	101
Percentage of phase examinations evaluated (%)	32	44	37

Table 6.17 Number of participants to have VPS 2.0 examinations evaluated by the expert pathologist.

	Phase		
	I	III	Total
Total number of participants	17	12	29
Number of participants evaluated	13	10	23
Percentage of participant evaluated (%)	76	83	79

Table 6.18 Number of examinations evaluated for each participant, and observed improvement in kappa for staging and grading. Kappa statistics only included for participants that examined and scored all 20 cases.

User	Number of examinations evaluated	Un-weighted Kappa Statistics*			
		Staging		Grading	
		Phase I	Phase III	Phase I	Phase III
11	20	0.221	0.494	0.241	0.73
19	20	0.351	0.474	0.315	0.625
13	19	0.146	0.859	0.277	0.73
23	9	-	-	-	-
5	3	0.091	0.265	0.275	0.315
6	5	0.231	1	0.275	0.221
14	5	0.459	0.383	-0.316	0.114
17	5	0.494	0.178	0.125	0.231
29	4	-	-	-	-
15	3	0.268	0.859	-0.125	0.259
25	1	-	-	-	-
84	1	-	-	-	-

* Kappa values calculated on all examined cases

6.7.2 Reviewing Pathologist

Examinations were reviewed by Professor Lennart Franzen, a Swedish pathologist with a special interest in liver pathology, and member of the KVAST group (Quality and Standardisation) for liver disease, appointed by the Swedish society of pathology.

6.7.3 Evaluation Criteria

Staging and Grading were evaluated separately, with staging technique based upon examination of Sirius red/Van Guisen stained slides, and grading examination technique on H&E stained slide examinations. Prior to evaluating each examination, the reviewer was informed of the submitted Stage & Grade, and gold standard (KVAST) case scoring. Examinations were replayed with the ReplaySuite and evaluated according to following criteria:

- Could a source of error be identified for examinations?
- Was sufficient tissue area examined during the examination?
- Was appropriate magnification used during the examination?

Potential sources of error were categorised into (1) misinterpretation of observed pathology (2) oversight – missed pathology (3) don't know/reason not identified. *Adequate tissue coverage* and *appropriate magnification use* criteria were evaluated either 'ok', 'no' or 'N/A' for incidences in which the appropriately stained slide was not examined. Examinations were considered N/A for staging if the examiner failed to review the Sirius red/Van Guisen stained slide, and N/A for grading if they failed to examine the H&E slide. This data allows a diagnostic evaluation tree to be developed (Figure 6.18), which attempted to describe the identification of sources of diagnostic discordance.

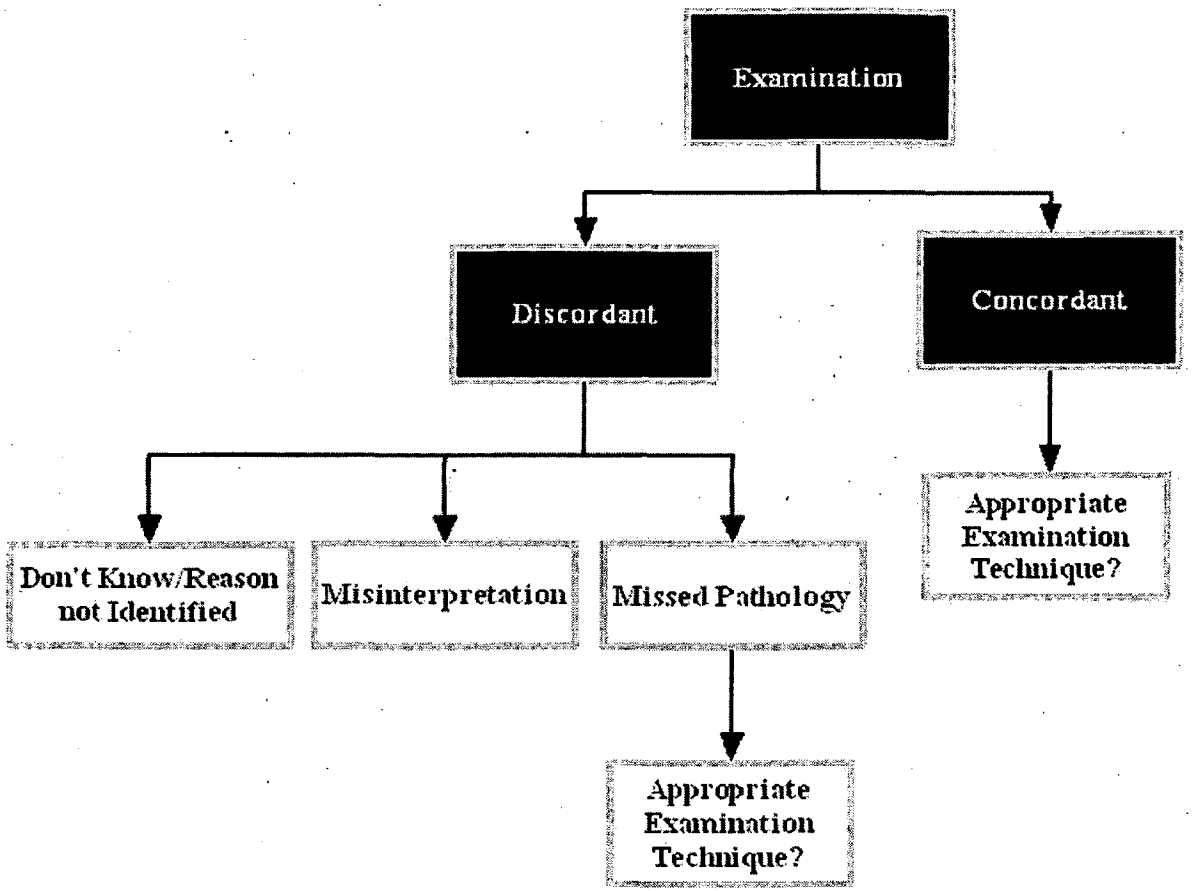


Figure 6.18 Assessment of sources of diagnostic error and deficiencies in diagnostic technique when examining digital slides using the VPS 2.0.

6.7.4 Comparison of Examination Technique for Concordant and Discordant Examinations

Table 6.19 illustrates that for all evaluated examinations (Phases I&III), overall, participants that scored concordant with KVAST used appropriate examination technique 28% (grading) and 24% (staging) more often than participants that scored discordantly. When grading, participants examined adequate tissue AND examined at appropriate magnification in almost half of examinations that agreed with KVAST, compared to less than 20% of examinations that disagreed with KVAST. Almost 60% of concordantly staged examinations used appropriate examination technique, in contrast to 35% of discordant examinations. Overall, participants only used appropriate examination technique in 33% of grading and 46% of staging examinations.

6.7.5 Comparison of Examination Technique in Different Scoring Phases

Between scoring phases (I & III), participant concordance with KVAST staging and grading improved by 20% & 19% respectively. However, as Table 6.19 illustrates, improvement in overall examination technique between phases was negligible. For both concordant and discordant grading, the number of examinations to cover sufficient tissue and examine at appropriate magnification remained the same (-1%), while it improved by 7% for staging.

Table 6.19 Percentage of evaluated examinations to use adequate examination technique, where examination technique is defined as examining adequate areas of tissue at appropriate magnification

	Grading examinations to use adequate examination technique (%)	Staging examinations to use adequate examination technique (%)
Concordant examinations	46	59
Discordant examinations	18	35
All examinations	33	46
Phase I examinations	33	43
Phase III examinations	32	50

6.7.6 Examination Technique and User Performance

To identify a correlation between improvements in kappa statistics and improved tissue area coverage and adequate magnification use, a large sample, for which the following data must have available:

- Kappa statistics based on examinations of all 20 cases
- Extensive evaluation of tissue coverage and magnification use

Of the 17 participants to examine cases, only 9 examined all 20 cases. Of these 9 participants, evaluation of tissue coverage and magnification use for all 20 case examinations occurred for only 2 participants, Users 11 & 19, while 19 examinations were evaluated for Users 13. While the lack of a large number of eligible participants prohibits analysis to elucidate a potential relationship between tissue coverage/magnification use and improvement in kappa statistics, it still enables a cursory evaluation of these users diagnostic technique in relation to performance.

User 11

As Table 6.20 illustrates, between Phases I & III, User 11 un-weighted kappa improved for both staging and grading. Grading un-weighted kappa statistics increased from 0.241 to 0.73 (+0.489), or from 'fair' to 'good' agreement with KVAST grading. This was the largest grading improvement observed. At the same time, the number of examinations to exhibit appropriate examination technique improved from 30-50% (+20%).

Between phases, un-weighted kappa improved from 0.221 to 0.494 (+0.273), or from 'fair' to 'moderate' agreement with KVAST staging, during which time the

percentage of examinations to use appropriate examination technique increased from 30 – 40% (+10%).

User 13

Between Phases I & III, un-weighted kappa improved significantly for both staging and grading, as shown in Table 6.20. Grading un-weighted kappa statistics increased from 0.277 to 0.73 (+0.453), or from 'fair' to 'good' agreement with KVAST grading. This was the second highest improvement in un-weighted grading kappa, after User 11 (+0.489). At the same time, an improvement in the number of examinations to use adequate examination technique was observed, however this improvement was minimal (+4%).

In contrast, a significant improvement in examination technique was observed in staging. Between scoring phases, the percentage of examinations that examined appropriate areas of tissue and at appropriate magnification increased from 40% to 100%. At the same time, staging un-weighted kappa statistics increased from 0.146 to 0.859 (+0.713), or from 'poor' to 'very good' agreement with KVAST grading, the second highest improvement in un-weighted, after User 6 (+0.769).

User 19

As Table 6.20 illustrates, staging and grading un-weighted kappa statistics for User 19 improved by 0.123 and 0.31 respectively, between scoring phases. Grading kappa increased from 0.315 to 0.625, or from 'fair' to 'good' concordance with KVAST grading. However, no change in examination technique was observed regarding grading, with the percentage of examinations to examine enough tissue at the appropriate magnification remaining static at 20%.In contrast, while staging kappa increased from 0.351 to 0.474, or from 'fair' to 'moderate concordance with KVAST staging, the percentage of examinations to exhibit appropriate examination technique decreased, from 50% to 30%.

Table 6.20 Evaluation of User 's 11, 13 & 19 concordance with KVAST staging and grading, using un-weighted kappa statistics, and adequate examination technique (adequate area coverage AND appropriate magnification use)

User		Staging		Grading	
		Un-weighted kappa	Examinations to use adequate examination technique (%)	Un-weighted kappa	Examinations to use adequate examination technique (%)
11	Phase I	0.221	30	0.241	30
	Phase III	0.494	40	0.73	50
13	Phase I	0.146	40	0.277	40
	Phase III	0.859	100	0.73	44
19	Phase I	0.351	50	0.315	20
	Phase III	0.474	30	0.625	20

6.7.7 Evaluation of Potential Sources of Error

In addition to evaluating examination technique on the basis of appropriate/inappropriate magnification use and adequate/inadequate tissue coverage, the reviewing pathologist attempted to identify possible diagnostic errors that may have resulted in discordant diagnosis. Discordant examinations categorised as 'missed pathology' were those considered to not have reviewed sufficient areas of tissue at an appropriate magnification, resulting in failure to locate pertinent visual features. Those categorized as 'possible misinterpretation of observed pathology' were considered to have potentially misinterpreted visual features, which may have contributed to the submission of discordant diagnoses. Examinations for which a potential source of error were not identified were categorized as 'don't know/reason not identified.'

Figure 6.19 illustrates the distribution of identified sources of error for all evaluated examinations, irrespective of phase. The source of error was identified for 50% of evaluated grading examinations and 41% of staging examinations. Missed pathology and misinterpretation of observed pathology were observed to be the source of error in 31% and 19% of discordant grading examinations respectively. The source of staging error was roughly distributed equally between missed pathology (21%) and misinterpretation (20%), where a source was established, with a source not identified in 59% of evaluated examinations.

Figure 6.20 below illustrates an example of an examined field, considered a possible source of misinterpretation by the expert reviewer. The reviewer stated that this field contained

> *"Two areas may possibly have been misinterpreted in this field, possibly as regenerative nodules. However these are really inflamed portal tracts"*

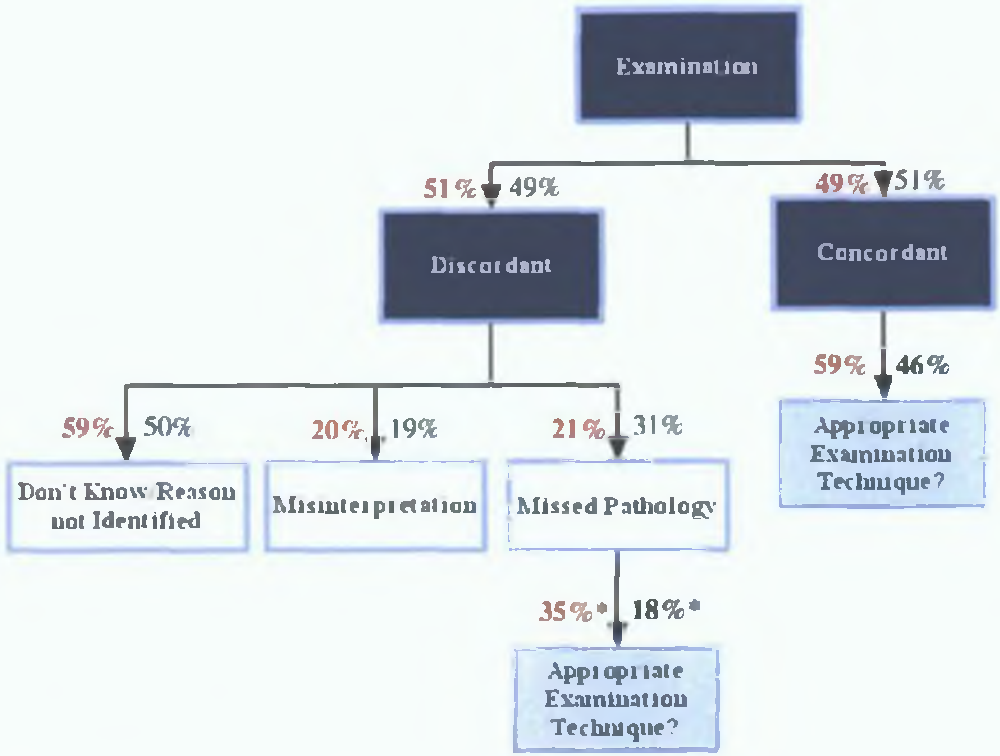


Figure 6.19 Assessment of sources of diagnostic error and deficiencies in diagnostic technique. Red values correspond to staging. Black values to grading.

* Values correspond to all discordant diagnoses

Figure 6.20 Possible misinterpretation of inflamed portal tracts as regenerative nodules by User 11 while examining Case 16.

6.8 Conclusions

Threat of litigation is the proverbial sword of Damocles that hangs over the head of medical practitioners. Histopathology is no exception, and the need to reduce errors by improving the quality, accuracy and consistency of diagnoses has meant that today's histopathologist must continually assess their diagnostic skills. Participation in conventional EQA schemes facilitates this, however the intrinsic benefit provided to participants is limited; discordance may be identified, but clear elucidation of the reasons for disagreement is difficult.

Two fundamental components of evolving any skill set are identification of errors, and understanding of how these errors occurred, in order to prevent repetition. Within the context of this study, these facilities were provided via the electronic resources available. Using the ReplaySuite and Image library, participants could:

- Identify incidences of discordance, by comparing diagnoses with Phase I gold standard (KVAST) scores
- Potentially elucidate diagnostic error sources, by replaying examinations concluding discordant diagnoses
- Gain insight into gold standard scoring, via Swedish society of pathology reference images, replays and SnapShots of concordant examinations

Previous work has shown that replaying examinations via the ReplaySuite can result in pathologists reconsidering their original diagnosis, and re-scoring concordant with group consensus and clinical diagnosis (Chapter 5). The purpose of this study, however, was to determine if the utilisation of this reference material could improve participant concordance with gold standard (KVAST) scoring of chronic hepatitis. Between scoring sessions, concordance increased for both staging (+18%) and grading (+20%), suggesting that use of electronic resources may improve concordance. However, the apparent type of improvement differed. While mean un-weighted staging kappa increased by 0.207, the increase in mean linear weighted staging kappa was negligible

(+0.085), suggesting that improvement in staging was due to 'fine-tuning'. In contrast, both mean un-weighted and linear weighted grading kappa improved by 0.280 and 0.295 respectively, suggesting that improvement in grading concordance is associated with decrease in all degrees of discordance.

The most significant outcome of the study has been the improved performance of a number of participants. User 6 recorded a marked improvement in staging concordance: un-weighted and linear weighted kappa improving by 0.769 and 0.606 respectively. In Phase I, User 6's staging concordance with KVAST was 40%, with over-staging occurring in the remaining 6 cases, by 1 (40%) and by 2 (20%). In Phase III complete staging concordance was achieved. User 13 experienced a similar improvement, recording an increase in un-weighted and linear weighted kappa of 0.713, and 0.367 respectively. In Phase I, concordance with KVAST staging was 30%, compared with 90% in Phase III. As with User 6, this user over-staged in 60% of Phase I cases, however, in all of these, over-staged only by 1; in contrast to User 6, under-staging the remaining case by 1. The only Phase III case staged discordant with KVAST was under-staged by 1.

An even more pronounced improvement in kappa values was observed in grading. This improvement was observed amongst the majority of participants, 66% (6/9) obtaining an improvement of greater than 0.3 in un-weighted kappa, and 55% (5/9) recording an increase greater than 0.23 in linear weighted kappa. In particular, User 15 achieved an improvement in un-weighted and linear weighted grading kappa of 0.751 & 0.384 respectively.

While improved performance observed suggests the use of supplementary electronic resources may be of benefit in improving concordance with gold standard scoring in chronic hepatitis, a number of caveats must be acknowledged. Firstly, due to low completion levels the sample size is low. Of the 26 institutions, while 23 participated in some form, only 9 completed the study. While this precludes extensive analysis of the data to correlate degree of scoring improvement with level of resource use, it does not invalidate observations regarding overall use. As previously highlighted in Chapter 4, other studies have reported similar difficulties with low participation rates (Bamford *et*

al., 2003). Low study completion here may be attributable to a number of factors. The often mentioned workload requirements placed upon pathologists is not conducive to participation in studies such as this, and a number of laboratories stated that work commitments prohibited their participation. Additionally, system speed was rated as slow for 67% of all examinations, which may have been a disincentive to continued participation.

Secondly, the lack of a control group prohibits assumptions that improved KVAST concordance was directly attributable to utilisation of the available resources. While Table 6.3 shows un-weighted staging/grading improvement for those who used the ReplaySuite extensively and considered the image library to be helpful, the absence of direct comparison with participants that did not utilise the software limits the degree of confidence in such assertions. It may be suggested that those who completed the study may represent a biased subset, however comparing the Phase I performance of study completion and non-completion participant subsets suggests that their performances do not differ significantly.

Thirdly, while participants were strongly recommended to examine cases with all members of the laboratory present, it is impossible to confirm that the cohort of pathologists within a given laboratory were the same for each phase. Given the workloads of pathologists, and the acknowledgement that some laboratories would be unable to participate in all phases due to other commitments, it is possible to speculate that internal participation within laboratories might fluctuate. The lack of direct interaction with participants prevents confirmation or disconfirmation of this however, and it is not unreasonable to suggest that the significance of this in individual laboratories would be negated between all the participating laboratories, over the course of the study.

Finally, while KVAST considered the two sets of cases of equal difficulty, this is a subjective assessment. Ascertaining whether this is an accurate statement is a difficult proposition, without external evaluation. One approach that may have helped to address this issue would have been to divide participants into two groups, with one group examining Phase I cases first, while the other group

examined Phase III cases. After utilising the electronic resources during Phase II, each group would then examine their remaining cases. However, given the lower than expected level of participation, it would have been difficult to distribute participating laboratories equally been scoring groups.

These issues illustrate the difficulty in conducting software validation studies involving pathologists. In an ideal scenario, pathologists would examine cases in a controlled environment, such as a computer room available exclusively for the purpose. This would allow the management of pathologists or groups of pathologists to ensure that the same cohort examine cases at all times, the explicit identification and management of a control group, and standardisation of PC and Internet facilities used, to ensure that all participants were exposed to the same experience. Unfortunately however, given the general constraints involved in working with pathologists (limited available time) and the specific issues pertaining to this study (location of participants), this was impossible to implement in this study, and would be almost impossible to achieve in similar studies.

A histopathologists' ability to accurately diagnose is dependant on his/her skill, dedication and experience. Ultimately, the process is subjective, and liable to human error. EQA and supplementary development programmes are in place to attempt to identify such errors, provide feedback and support. to ensure diagnostic standards are maintained, and errors reduced. While the use of virtual microscopy in EQA faces opposition from some quarters, its use opens up a plethora of avenues of analysis unavailable with conventional microscopy. In addition to removing the inherent delays and risk of loss/damage associated with physical delivery of slides and enabling the utilisation of small tissue samples, it provides a potential means of elucidating sources of diagnostic error by reviewing the diagnostic process. This may be the most substantial legacy of using virtual slide technology, identifying and correcting erroneous diagnostic techniques, as opposed to mere evaluation of laboratories' diagnostic proficiency.

It is reasonable to assume that accurate diagnosis based on histological evaluation requires methodical and thorough microscopic examination, however it is difficult to quantify this assumption in an external quality assurance context utilising glass slides. In contrast, the capability to review examinations by replaying them with the ReplaySuite creates a powerful method for evaluating examination technique, and identifying potential sources of diagnostic error.

As Table 6.19 illustrates, examinations scoring concordant with KVAST were observed to exhibit acceptable examination technique more frequently than discordantly scoring examinations. When grading, 28% (46% - 18%) more concordant than discordant examinations were considered to have viewed sufficient tissue, and at the appropriate magnification. This was similar to the 24% (59% - 35%) discrepancy between concordantly and discordantly staging examinations, suggesting that examination technique is important both when determining the degree of necroinflammation within a biopsy, and when ascertaining the extent of fibrosis.

It should be noted that while examinations by 8 of the 9 participants to examine all cases were evaluated, roughly 60% of those reviewed were performed by three of the twelve users: Users 11, 13 & 19. This introduces a potential source of bias, as all three exhibited improved performance, however the sample size is too small to justify their omission.

When viewed in isolation, portal tracts within the same liver biopsy may be graded differently. Therefore, it is generally considered essential to examine several portal tracts and grade according to the overall evaluation (Scheuer, 2003). Failure to observe a sufficient number of portal tracts can lead to grading that, while representative of the portal tract(s) examined, is not representative of the entire biopsy. 63% of evaluated discordant grading examinations, and 52% of discordant staging examinations identified as 'oversight errors'. While it is difficult to hypothesis with such a small sample, the data suggests that oversight errors may be the more common source of error in histological examination.

For example, User 13 graded Case 3 Grade 4, discordant with the KVAST Grade 2, and was considered to have examined insufficient tissue, and at too low a magnification. During evaluation of the examination, the reviewer observed that the examiner only looked at one portal tract, which contained heavy inflammation. Taken in isolation, this portal tract was considered by the reviewer to be at least Grade 3. However, two portal tracts within the biopsy that were not examined were considered to be Grade 2 and 2-3. The reviewer stated that aggregation of the 3 portal tract grades should lead to an overall biopsy Grade 2.

For Users 11, 13 & 19, improved mean diagnostic performance coincided with a mean increase in the percentage of examinations that adequately examined slides. Mean un-weighted kappa for staging increased between the three participants by 0.37 (0.24–0.61), or an improvement from a 'fair' to 'good' level of agreement with KVAST staging. At the same time, the mean percentage of examinations to display appropriate examination technique improved by 17%, from 40 to 57% of examinations. An increase in mean un-weighted kappa from grading of 0.417 (0.278 – 0.695) ('fair' to 'good' level agreement with KVAST) coincided with a 8.15% overall improvement in the 3 users examination technique. This suggests that not only is appropriate examination technique important in accurate grading and staging, but that improvement in technique may improve scoring concordance with gold standard grading.

Overall, misinterpretation was cited as the possible reason for discordant grading in 38% of evaluated examinations. For example, the reviewer's evaluation of User 5's grading examination of Case 13 concluded that connective tissue might have been misinterpreted as inflammation. However, the reviewing pathologist's interpretation of where misinterpretation may have occurred is subjective, and without extensive examiner commentary indicating whether connective tissue was truly misinterpreted as inflammation, such conclusions are speculative.

Oversight was cited as the source of error for 63% of discordant grading examinations evaluated, with the examiner considered to have missed diagnostically important histological features. As with grading, failure to

examine sufficient portal tracts can lead to staging not representative of the biopsy. For example, User 11 was observed to have examined broad collagen during the examination of Case 5. The reviewer stated that, in isolation, this would suggest Stage 2, the score submitted for the examination. However examination of other fields in addition to examination of this connective tissue would have been indicative of Stage 1, the score submitted by KVAST.

Misinterpretation of observed pathology was cited as the possible source of error in 48% of evaluated examinations that staged discordantly, where a source was identified. Figure 5.1 illustrates a field viewed by User 11 while examining Case 16. The reviewer stated that inflamed portal tracts within this field might have been misinterpreted as regenerative nodules. Regenerative nodules may form in the fibrous septa but they are not necessary for the histological diagnosis of cirrhosis: as nodules alone without fibrosis do not constitute cirrhosis (Loyola University Chicago, 2005).

Whilst acknowledging that oversight error was cited as a significant source of observed error, it should be mentioned that virtual slide examinations are restricted by the technological limitations of the system being utilised, and as such differ from their glass counterpart. Figure 6.5 illustrates, 2/3 of examinations were considered slow while using the VPS. In such circumstances it is understandable that examinations may not be as extensive as those, conducted using glass slides, resulting in a reduction the area examined at high magnification.

The reviewing process utilised to evaluate the examination techniques used is a subjective one, introducing a potential source of error in the evaluation process itself. On several occasions the reviewer changed his mind as to whether examinations had been conducted satisfactorily, and often it was difficult to categorise a potential source of error. Examinations that misinterpreted observed pathology may also have missed important pathological changes, however error categories were mutually exclusive. In such cases, the source was categorised as "Don't know/Reason not identified."

In addition, the subtle differences between some categories made the conclusion of some discordant diagnoses understandable, and ascertaining whether a significant diagnostic error occurred (i.e. oversight or misinterpretation) difficult. For example the reviewer stated that it is often difficult to differentiate between Stages 0 & 1, and that the scoring categories are semi-quantitative; staging or grading often falling between two categories and requiring consensus amongst the examining group to decide a score. This is not an uncommon phenomenon, as illustrated by lack of discreet division between B3 & B4 classifications in breast core biopsy assessment (Chapter 5).

Most significantly, it should be acknowledged that while there are only 2 possible sources of error in histological assessment (misinterpretation of observed pathology, missed pathology), the expert reviewer did not successfully identify a source of error in 59% of Staging, and 50% of Grading examinations reviewed. It is not unreasonable to suggest that, by default, if an examiner has examined a sufficient area of tissue at appropriate magnification then the error must be due to misinterpretation of observed pathology. Alternatively, if it cannot be established that the examiner missed relevant pathology, but that they did not examine a sufficient area of tissue at appropriate magnification, then the source of error is most likely missed pathology. Given this fact, it should be feasible to identify a source of error in all examinations, and it must be acknowledged that in attempting to ensure that discordance was not assigned to specific sources unless the expert was confident, the experimental design was not ideal.

The results of this study illustrate the potential added value that ReplaySuite functionality provides to External Quality Assurance. While other researchers have made reference to examination tracking capabilities (Brauchli *et al.*, 2002), this is the first work to utilise such capabilities to elucidate sources of error in histological assessment. This is of enormous potential benefit, both to EQA organisers for identifying diagnostic trends, and individual pathologists for identifying reasons for diagnostic discordance. Participants in future computer-based EQA schemes will be able to examine digital slides online, immediately receive kappa statistics evaluating their performance, then review 'gold standard' examination replays highlighting important diagnostic cues. It is even feasible to

identify important diagnostic cues that were missed during examination, by cross-referencing participant and expert SnapShot data. Such extensive feedback and analysis provides participants with a much more comprehensive 'service' than conventional, glass-based schemes. In addition, committee review of examinations may be able to identify frequently occurring errors that may contribute to inter-observer variability. This may be the greatest legacy of digital slides in pathology EQA, particularly in areas where inter-observer variability is high. While conventional EQA has always been concerned with identifying problems, virtual EQA may be able to identify the reasons for such problems.

Chapter 7: Irish Trainee Pathologists Perceptions Of Computer Assisted Learning

7.1 Introduction

Computer-assisted learning in pathology education is finding greater acceptance in a supplementary role, with online pathology resources such as Pathnet (Rous, 2005) and Tapir (Rous *et al.*, 2005) evolving to meet the specific requirements of pathology trainees. Shared by histopathology training schools across the UK, Pathnet allows both consultants and trainees to contribute presentations, tutorials, cases and discussion topics, while Tapir acts as a repository of teaching images and tutorials. Legal and ethical considerations regarding the use of patient images for educational purposes has been raised, however if patient anonymity is maintained, then use of such images for educational purposes is considered permissible (Tranberg *et al.*, 2003).

Online resources can reduce the burden on formal teaching, encouraging independent learning and provide trainees with access to material that would not normally be so readily available. Research in the US (Klatt, 1997, Jones *et al.*, 2002) and UK (Reid *et al.*, 2000) has previously illustrated that the incorporation of CAL into formal training has been greeted with enthusiasm, while Blake *et al* (2003) observed that students at the University of South Carolina preferred the use of digital slides over glass. Other researchers had successfully integrated digital slides into a seminar setting (Dee *et al.*, 2003), with attendees at the annual pathobiology of cancer workshop at the University of Iowa enthusiastic about the technology. One attendee commented that:

> *"Digital slides are the digital equivalent of a multi-headed microscope for the entire class."*

This illustrates that trainee pathologists in other countries were amenable to the use of digital slides in formal teaching environments, and also to supplementary computer-assisted learning. However to date no research had attempted to evaluate Irish pathology trainee's perceptions towards similar CAL systems.

The impact of ReplaySuite use on diagnostic performance in an EQA setting has already been described (Chapter 6). Prior to this, it had been determined that

evaluation of the technology in a training setting would be beneficial. However, subsequent to the work conducted in chronic hepatitis EQA, it was decided that such an evaluation would be too similar to previous work to provide pertinent new information. However, it was still considered relevant to evaluate pathology trainee's perceptions of virtual slide based computer-assisted learning (CAL).

The opportunity to participate in the Dublin-region postgraduate pathology seminar series provided such an environment to assess Irish trainee's perceptions. It was proposed that digital slides would replace glass slides during weekly/bi-monthly seminar presentations, which would subsequently be made available online to trainees, with additional reference material.

7.2 Limitations of Existing Technology

During the course of the seminar series, a large volume of slides (>80) were required to be digitised. This created a number of logistical problems. While redevelopment of scanning algorithms improved on existing digitisation methodologies considerably (Chapter 3), the system still necessitated human intervention during the digitisation process, was susceptible to crashing (due to heavy workload) and most significantly, continued to require several hours to successfully scan and publish a slide. The procurement of access to an ultra-fast slide scanner (Aperio ScanScope T3) provided a means of rapid digitisation, however, slides would be incompatible with the VPS and ReplaySuite applications, prohibiting their use.

It was therefore decided to utilise 'off the shelf' commercial software to deliver functionality, and follow common practice of other institutes (Lundin *et al.*, 2004, Feit *et al.*, 2005) by providing a virtual laboratory as the available online resource, subsequent to presentations. The online system used, the Digital SlideBox™, provided digital slides and tutorials, incorporating annotations with extensive descriptions. This provided a number of benefits over the VPS and ReplaySuite:

- *Smooth scrolling.* The virtual slide viewer component facilitated smooth lateral motion

- *Improved Stabililty.* The VPS and ReplaySuite customised browser technology had encountered technical difficulties with hospital firewalls. The use of a commercial application that had been fully tested and validated significantly reduced the likelihood of similar difficulties occurring again.

- *Expert Annotations.* The Digital SlideBox™ provided access to expert slide annotations and detailed explanations of observed pathology

The following Chapter describes the replacement of glass slides with digital slides in the Dublin based postgraduate seminar series, and subsequent provision of online access to seminar material, via commercial applications. The objective of the chapter was to assess Irish trainee pathologist's perceptions of a virtual laboratory CAL system (Digital SlideBox™), based on cases encountered during seminars.

7.3 Hardware and Software

A number of commercial hardware and software applications were used to provide digital slides and online resources to seminar attendees.

7.3.1 ScanScope Ultra-Fast Slide Digitisation

The ScanScope™ T3 is an ultra-fast slide scanner purpose built for high-throughput slide digitisation. Developed to facilitate 'one touch scanning,' it is capable of digitising up to 5 slides without human intervention beyond initiation of the scan. Capturing images via 20x/40x (40x: 0.75/0.95 Plan Apochromat) objective, it achieves optimal focus by automatically making rapid real-time focus adjustments to compensate for subtle changes in specimen topology (Aperio Technologies, 2005).

The ScanScope uses line scanning to digitise slides, rather than previously utilised image tiling. Line Scanning invokes the use of a linear array detector and specialised motion control components (nanomotors), as opposed to a fixed imaging device. The resulting scanning procedure keeps the slide in constant motion during data acquisition, rather than pausing between frames as most conventional scanners do. Line scanning holds a fundamental advantage over stitching and tiling; the capture and alignment of a small number of images is far more efficient and less prone to misalignment and overlapping than the capture of thousands of smaller images. In addition, captured images are free from optical aberrations along the scanning axis, resulting in digital slide images that have minimal/no tiling artefacts. Digital slides are stored in true colour (32-bit) TIFF image file format, with JPEG or JPEG2000 (wavelet) compression (Aperio Technologies, 2005). Figure 7.1 illustrates the graphical user interface of the ScanScope system.

7.3.2 ImageScope

The commercial virtual slide viewer utilised was ImageScope (Aperio Technologies, 2005). It interfaces with Aperio's ImageServer™ software using network protocols to access remote slides, or locally from CD/DVD/local hard-drive. Figure 7.2 illustrates the ImageScope GUI used to facilitate virtual slide examination. Using ImageScope, users could:

- Pan (smooth scrolling) and zoom to any region and magnification of a multi-gigabyte virtual slide.
- View multiple digital slides concurrently, in synchronized windows.
- Apply image enhancements, in real-time, for contrast, brightness and gamma.
- Annotate regions of interest.

It provides a number of navigational methodologies, such as 'drag and drop' for lateral motion, a slider on the control panel to control magnification, and ZoomBox capabilities. ImageScope was used during case presentation during seminars, as it allowed digital slides to be examined offline (without an Internet connection), in contrast to the Digital SlideBox™. It also had annotation capabilities, which were used to generate expert annotation tutorials.

7.3.2.1 ImageScope Annotation

ImageScope provided annotation capabilities that allowed the examiner to mark regions and attach comments. Areas could be marked with a number of shapes (rectangles, ellipses, lines and arrows). Figure 7.3 illustrates using ImageScope to mark areas with (a) a rectangle and (b) an arrow. The authoring tool creates XML files to store annotation data, with tags within the XML (eXensible Markup Language) file recording the geometrical coordinates of the annotations and comments. A more detailed explanation of XML is provided in Appendix A.

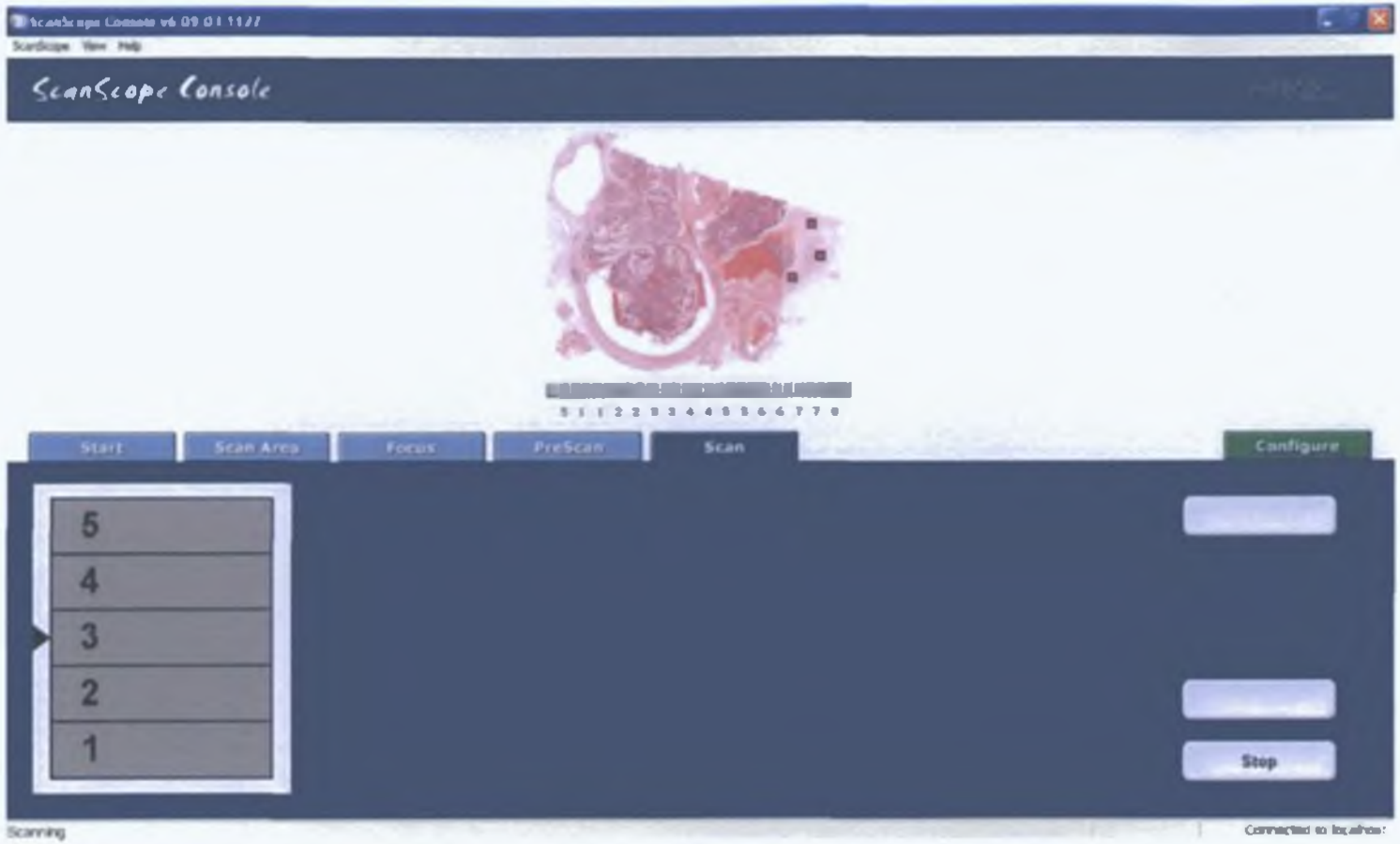


Figure 7.1 The Aperio ScanScope T3 scanner Graphical User Interface while digitising a slide

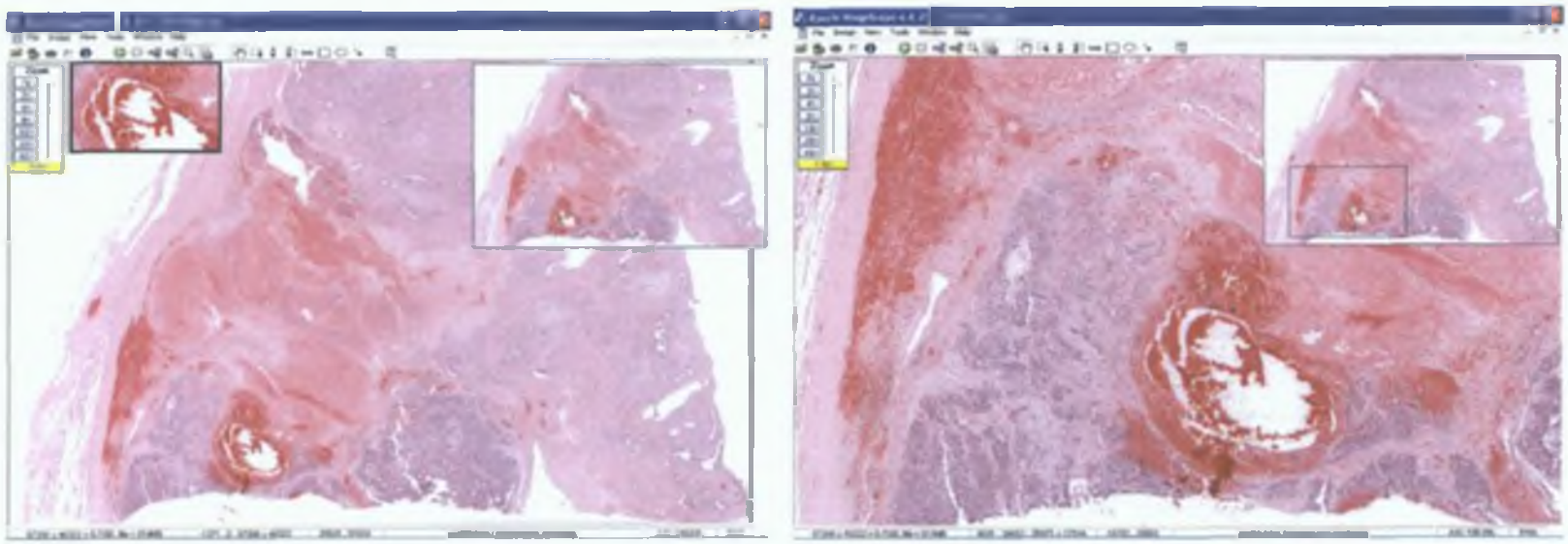


Figure 7.2 Examining a biopsy using ImageScope. (a) Examining the biopsy at low magnification (b) Examining a specific region at higher magnification

(a) (b)

Figure 7.3 Annotating regions of interest using ImageScope. (a) Marking a rectangular region (b) Marking a visual feature with an arrow

7.3.3 Digital SlideBox

The Digital SlideBox™ is a virtual slide management system designed for application in pathology External Quality Assurance. It facilitates rapid EQA study set-up, by enabling sets of cases to be quickly associated with a group of users, and a scoring regimen. Users using the system are only able to access cases from studies they are registered for, and are able to submit diagnoses and upload annotations. This system can be exploited for educational purposes, allowing related cases to be grouped into a 'study,' with cases containing multiple digital slides, and expert annotations: tutorials displaying regions selected via annotation using ImageScope.

Figure 7.4 & Figure 7.5 illustrate the graphical user interface (GUI) of the Digital SlideBox™. As with the ReplaySuite, users were required to log in (Figure 7.5 (a)). Once successfully logged in, available studies are displayed (Figure 7.5 (b)). This provided two options for each study: the user may review clinical data for all cases in a study (Figure 7.5 (c)) or review gold standard diagnoses and a list of tutorials (if available) for cases within the study (Figure 7.5 (d)). From either option, the user could examine case slides using the Zoomify slide viewer (Zoomify Inc., 2004) (Figure 7.5(e, f)).

The Digital SlideBox™ permitted users to submit diagnoses after examination, using summary report forms dynamically defined during the study set-up process (Figure 7.5(g)). ImageScope (Aperio Technologies, 2005) annotations could also be uploaded with the report form. If the uploading user was assigned 'gold standard' status during study setup, these annotations are automatically converted into a tutorial, which displays each annotated field, along with diagnostic comments attached to the field (Figure 7.5 (h)).

The Digital SlideBox™ used a similar architectural design to the ReplaySuite, using PHP to facilitate communication with the database and dynamically generate HTML for use as the GUI. The most significant difference relates to the use of a mySQL database (Appendix A), rather than an Oracle database as used by the VPS and ReplaySuite.

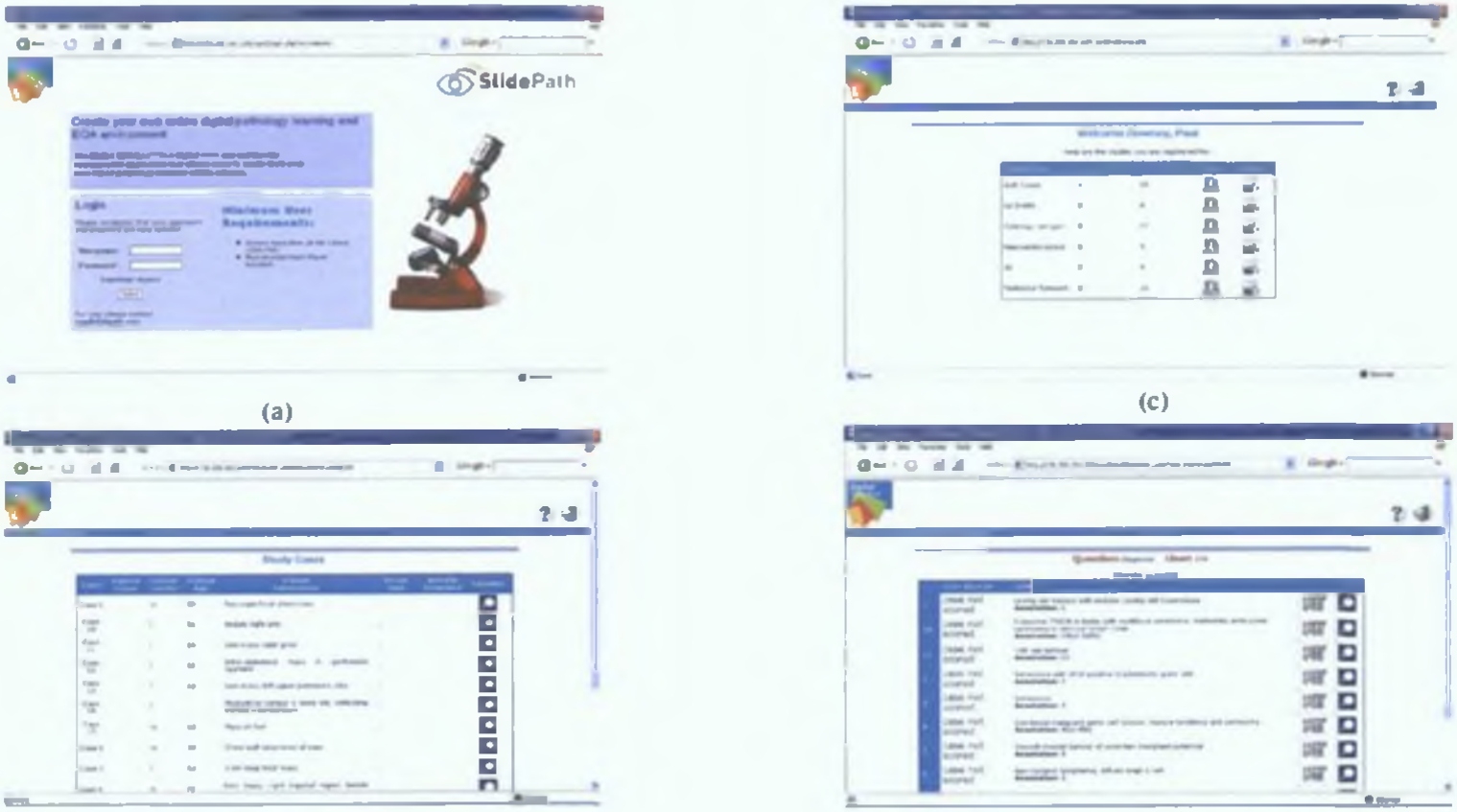


Figure 7.4 Digital SlideBox GUI. (a) Logging onto the system (b) Available Seminars (c) Clinical data for cases within a seminar (d) Reviewing case diagnoses.

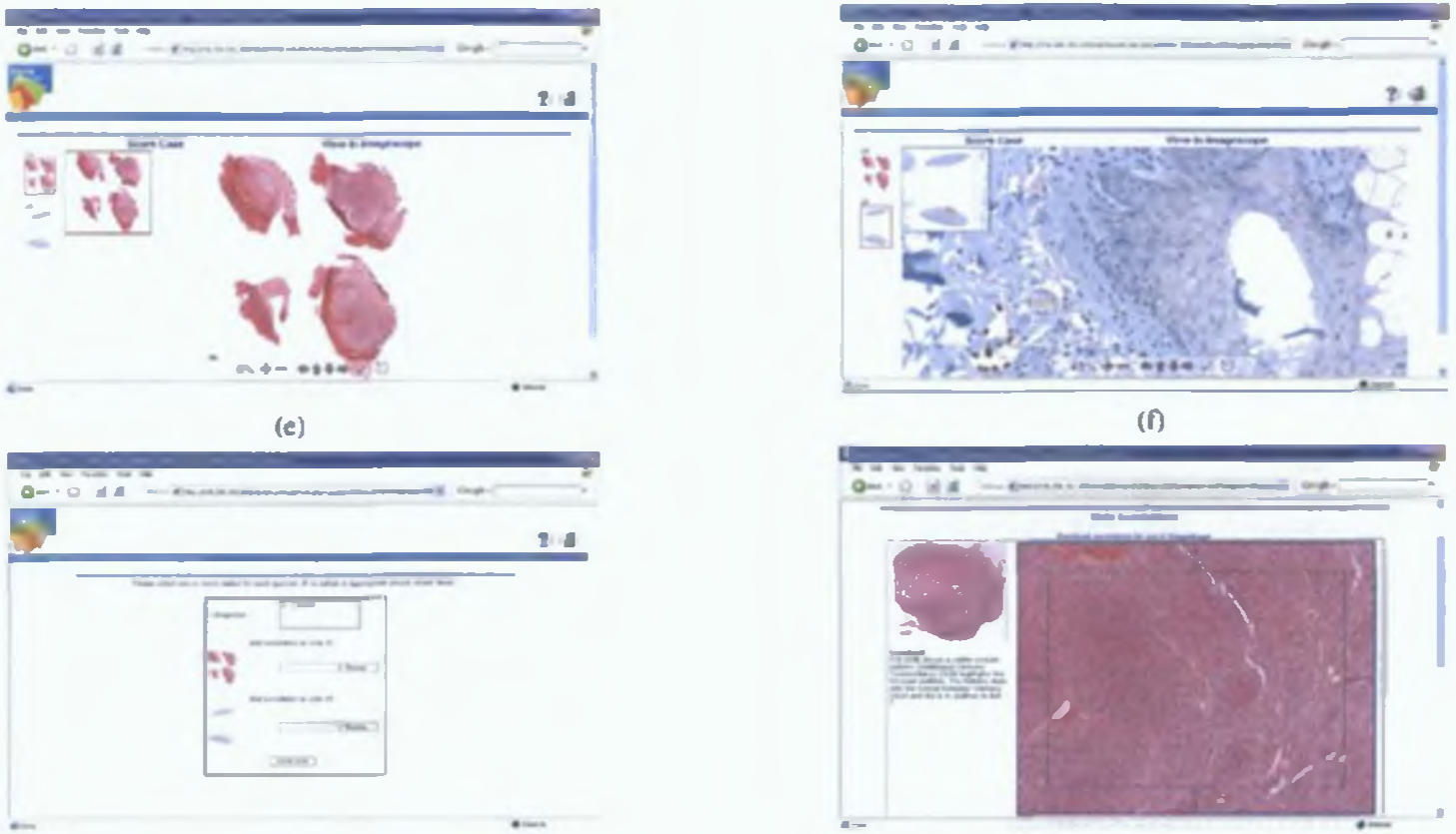


Figure 7.5 Digital SlideBox GUI. (e) Examining a virtual slide (f) Examining an immuno-stained slide for the same case (g) Submitting a diagnosis (h) Reviewing expert annotated fields.

7.4 Postgraduate Seminar Series

The postgraduate seminar series is an annual set of roughly 30 seminars presented throughout the year, and attended by trainee pathologists based in Dublin hospitals. Weekly/Bimonthly, a consultant pathologist from one of the participating institutions presents each seminar, on a range of topics such as testicular tumours, urology, soft tissue, breast and haematolymphoids. The location of the seminars rotates between participating hospitals, and attendance fluctuates, with trainee numbers varying from 10 to 20. Trainees of varying levels of experience attend, and the duration of each seminar is usually 1-2 hours.

Each seminar is structured and executed in an approximately equivalent manner. Prior to commencement, trainees are permitted to microscopically examine glass slides for cases to be discussed during the seminar. The number of cases varies, but is typically in the range of 8-15 cases. Multi-headed microscopes are often used, requiring one trainee to 'drive' the examination. Seminars usually begin with a presentation on the topic in question by the consultant, however this does not always occur. Trainees are subsequently invited to discuss and diagnose the previously examined cases. A CCD camera is connected to a light microscope, allowing the slides to be projected onto a screen and examined. Clinical details and immunochemistry images may also be provided, depending on the consultant and topic. Figure 7.1 provides a description of seminars that utilised digital slides.

7.4.1 Cases

Cases presented were encountered by the consultant in clinical practice, and considered to be of interest to trainees. Immunohistochemistry stained slides were provided when present with case material.

Table 7.1 Postgraduate seminars for which digital slides were provided

Topic	Consultant	Cases digitised	Slides digitised
Soft Tissue	Dr Mairin McMenamin	15	21
Testicular Tumours	Dr Barbara Loftus	11	13
Non-Gynaecological Cytology	Dr Michael Jeffers	10	10
Haematolymphoid	Dr Sean O'Briain	9	9
Gastrointestinal	Dr Paul Crotty	9	14
Gastrointestinal Pathology	Dr Kieran Sheahan	8	11
Total		62	78

7.4.2 Seminar Attendees

Seminar attendees were trainee pathologists, ranging in experience up to 5 years. Attending trainees were based in the following hospitals in the Dublin area:

- St. Vincent's University Hospital, Dublin
- Beaumont Hospital, Dublin
- AMNCH (Adelaide and Meath Hospital, Dublin, incorporating the National Children's Hospital)
- Mater Misericordiae University Hospital, Dublin
- St. James's Hospital, Dublin

7.5 Utilisation of Computer Based Applications

Consultants presented digital slides during seminars and where then required to use the digital slides to create tutorials on each case.

7.5.1 Presentation of Digital slides During Seminars

For participating seminars, digital slides replaced glass slides during case presentation. The virtual slide under discussion was displayed on a large screen,

from a PC via a connected projector. Examination of case slides was facilitated using the ImageScope slide viewer (Aperio Technologies, 2005).

7.5.2 Expert Tutorial Generation

Subsequent to each seminar, the presenting consultant was requested to examine and annotate each case slide, marking regions of diagnostic importance and describing the pathological changes observed. Consultants were requested to mark regions of interest at low magnification, before describing observed pathology at higher magnification. In this manner it attempted to closely replicate the process of tracking an examination, however only recording fields of interest. Annotation files were uploaded to the Digital SlideBox™, where they were associated with the 'gold standard' diagnosis, for the appropriate case.

7.5.3 Online Seminar Material

Subsequent to conclusion of seminars, the Digital SlideBox™ system provided trainees with online access to digital slides, expert annotations, clinical details and diagnoses from cases presented during seminars. Participants were provided with unique usernames and passwords, which were required to access seminar material. The Digital SlideBox™ grouped seminar cases into 'studies' allowing all information from one seminar to be stored in one location. Trainees were able to examine case slides, observe clinical and diagnostic data and review expert annotations in the form of a static image tutorial, with extensive descriptions.

7.6 Trainee Electronic Survey

Subsequent to the conclusion of the seminar series, trainees were requested to complete an electronic survey, in order to evaluate their perceptions of using of digital slides during seminars, and the online resource (Digital SlideBox™). This comprised a HTML form that submitted data back for analysis via PHP's email capabilities.

The survey contained fourteen compulsory multiple-choice questions (MCQ), and two optional open-ended questions. Two of the MCQ required Yes/No responses, while the remaining twelve allowed the user to answer questions using a 5-point Likert scale (Likert, 1932). The complete survey is available in Appendix C.

7.7 Evaluation of Available Resources

In order to assess trainee's perceptions of digital slide use during seminars, and online material available through the Digital Slidebox™, trainees completed an online electronic survey. Survey questions were answered using a 5-point Likert scale to and open text. The survey is presented in Appendix C.

7.7.1 Digital slides in Seminar Presentations

While a large proportion (56%) of surveyed students were neutral regarding their preference of glass slides or digital slides, 75% of participants that expressed a preference preferred the use of digital slides to present cases during seminars (Figure 7.6). It should be noted however that two of the five categories for this question were erroneously titled. As Figure 7.6 'Not very' and 'Not at all' were presented instead of 'Preferred glass' and 'Preferred digital.'

All surveyed trainees considered it beneficial to have digital slides available for examination prior to seminars, with 67% considering it of great benefit (Figure 7.7), and almost 90% of surveyed trainees preferred to have digital slides available for between 1 and 2 weeks, prior to seminars (Figure 7.8).

7.7.2 Digital SlideBox

Trainees were asked how frequently they accessed the Digital SlideBox™ after seminars. 33% stated they used it occasionally, 44% an average amount and 22%

used it frequently. As Figure 7.9 illustrates, all users found the system easy to use, with 44% considering it very easy to use. 89% of surveyed trainees favourably evaluated image quality, with the majority (56%) considering it very good (Figure 7.10). However, perceptions of speed were more greatly distributed, with less than 50% of participants considering speed good, and 22% considering it not very good Figure 7.11.

7.7.3 Evaluation of Current and Potential Functionality

As Figure 7.12 illustrates, all participants considered digital slides and expert annotations to be of benefit as an online resource. 67% of surveyed trainees considered digital slides of great benefit, while 89% considered expert annotations of great benefit as an online resource.

All trainees considered it beneficial to be able to search for digital slides according to organ, diagnosis or pathological feature to be of benefit, with 67% considering it of great benefit. In contrast, only 44% considered expert examination videos of great benefit, while 33% expressed no preference (neutral) Figure 7.13. When asked to state Yes or No as to whether trainees considered the incorporation of Multiple Choice Questions (MCQ) and links to relevant publications, 89% said Yes.

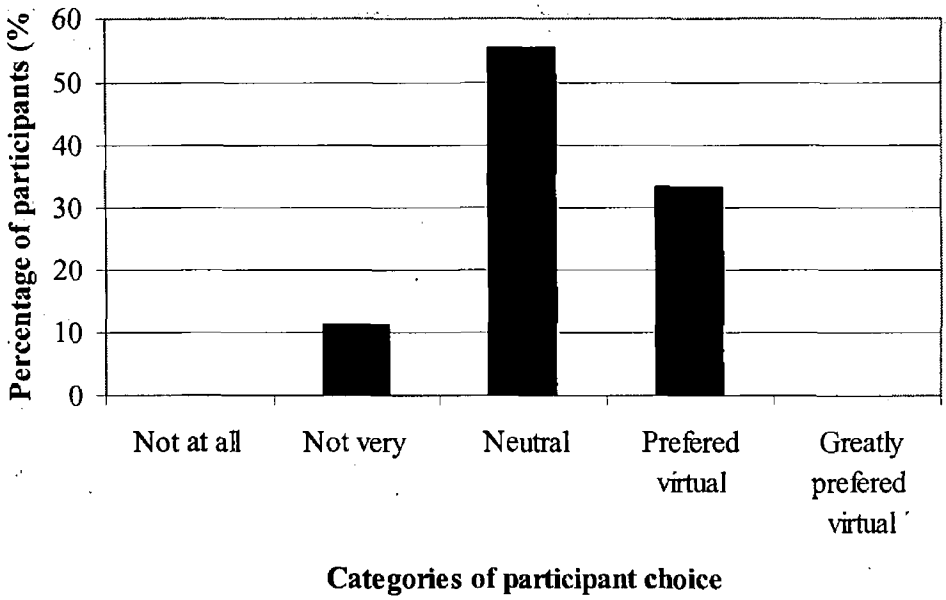


Figure 7.6 Trainee pathologist's preference of digital slides over glass slides during postgraduate seminars.

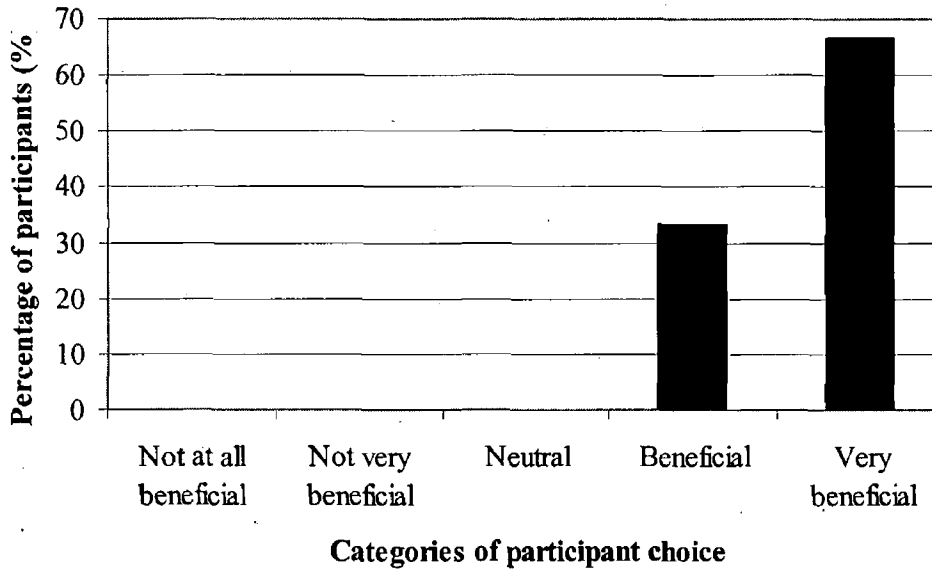


Figure 7.7 Trainee pathologist's perceptions of the benefit of having digital slides available before postgraduate seminars for review.

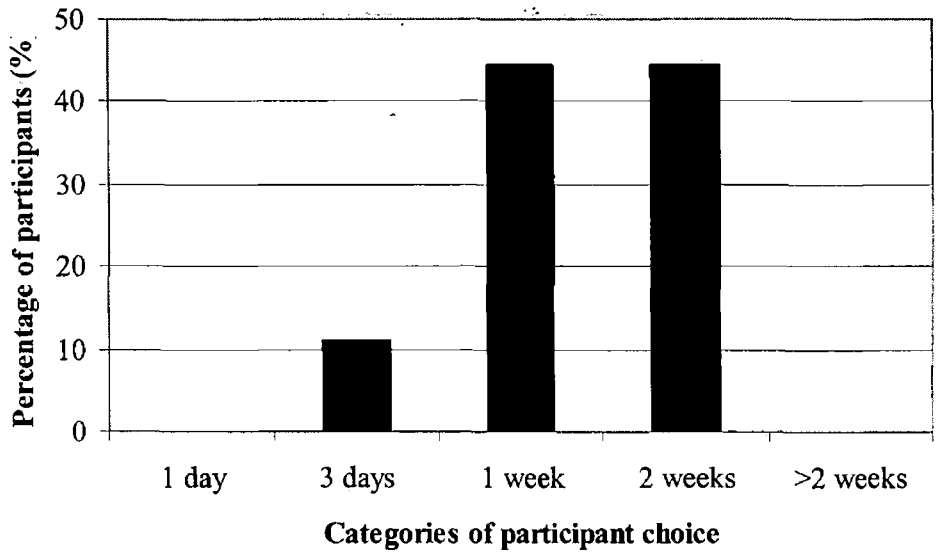


Figure 7.8 Trainee pathologist's perceptions of how much time prior to postgraduate seminars they would like to have access to digital slides.

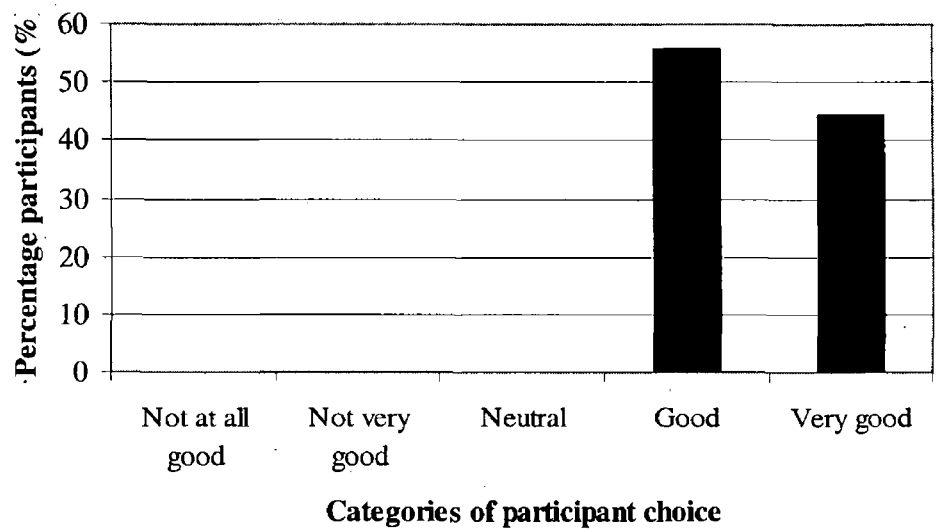


Figure 7.9 Trainee pathologist's rating of Digital SlideBox ease of use.

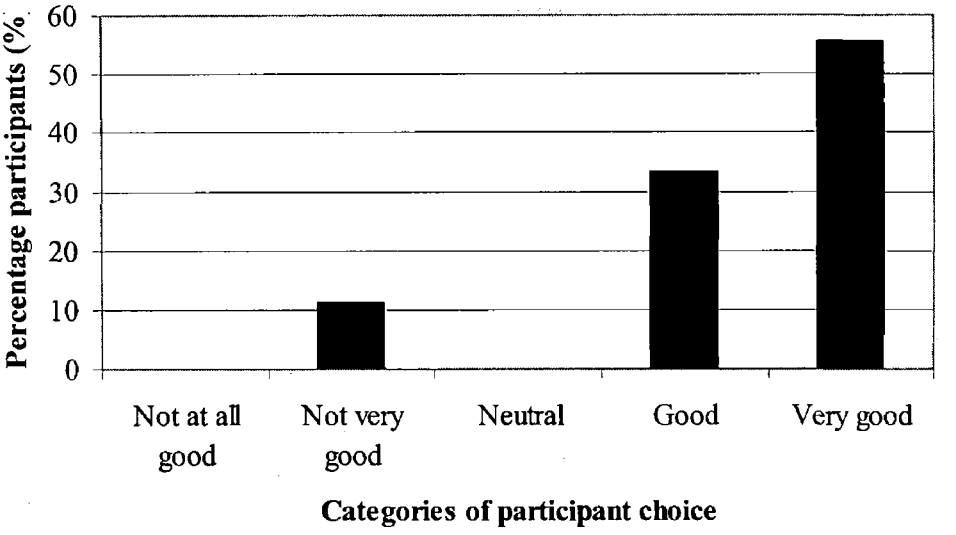


Figure 7.10 Trainee pathologists rating of Digital SlideBox image quality.

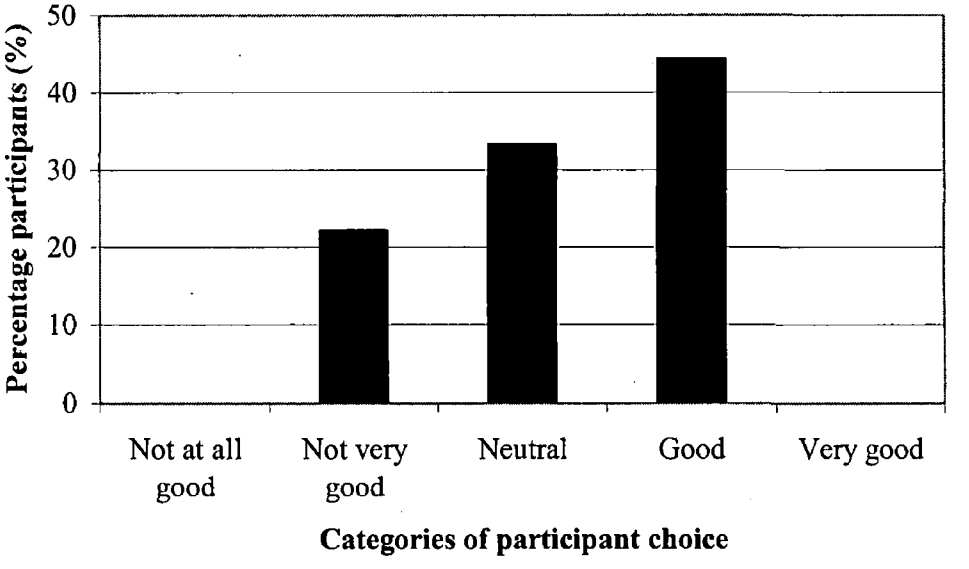


Figure 7.11 Trainee pathologists rating of Digital SlideBox speed.

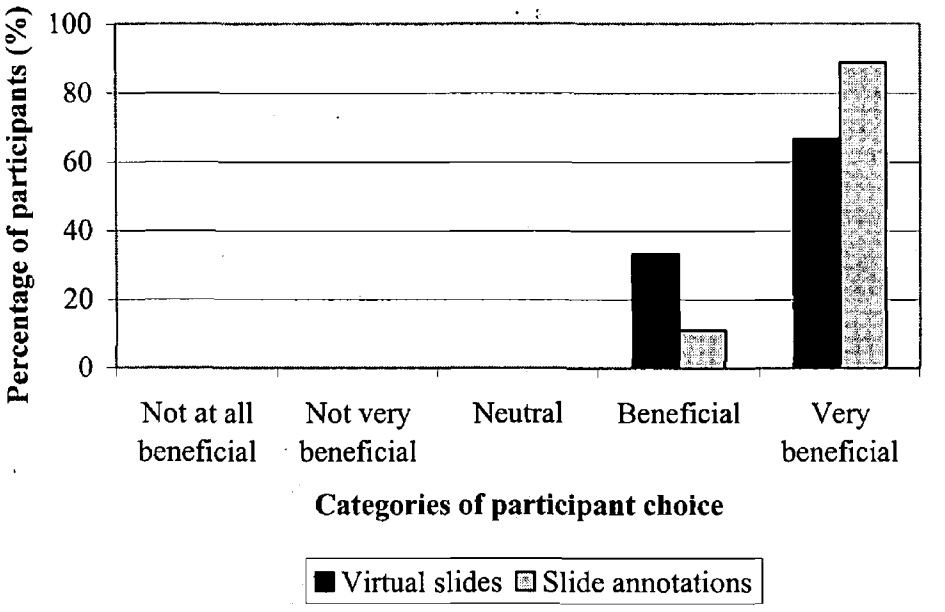


Figure 7.12 Trainee pathologists perceptions of the benefit of existing Digital SlideBox functionality.

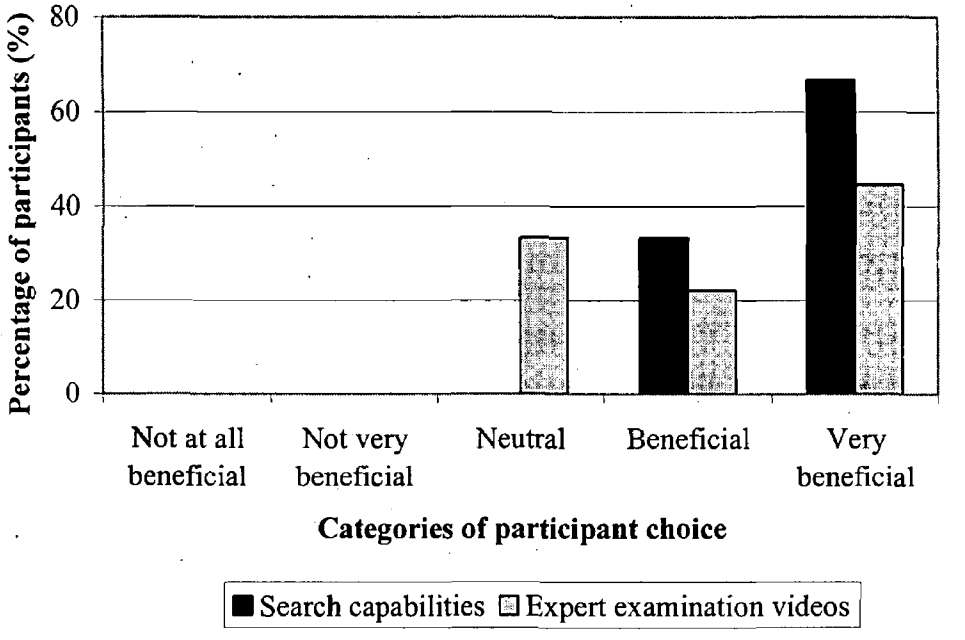


Figure 7.13 Trainee pathologist's perceptions of the benefit of potential Digital SlideBox functionality.

7.8 Conclusion

Studies have shown that both students and faculty are enthusiastic about the use of computer-assisted learning systems in histology training (Reid *et al.*, 2000, Klatt, 1997, Jones *et al.*, 2002, Blake *et al.*, 2003). Dee *et al* (2003) observed that the use of digital slides enabled faculty to better point out cells and lesions, and enhanced students ability to learn from glass slides at the microscope, while Klatt (1997) noted that CAL reduced formal teaching times by up to 30%, providing students with more time for self-directed learning.

For faculty, online virtual laboratories are significant teaching tools, as they effectively function as information repositories, accessible for teaching purposes when required, and updated at will. This issue was highlighted by one consultant in the Dublin seminar series, who commented that having this year's material online would make it much easier to construct next years teaching material (O'Briain, 2005).

Trainees were enthusiastic about having access to material before formal presentation, with all surveyed trainees considering it of benefit to access to digital slides prior to their presentation (Figure 7.7), and most stating a preference for having slides available 1 (45% of trainees) or 2 (45% of trainees) weeks before seminars (Figure 7.8). It was observed that while it is obvious that trainees would like to have seminar slides available to them as early as possible, workload constraints result in many consultants only preparing cases for presentation several days before a seminar (O'Briain, 2005).

As an educational resource subsequent to seminar presentation, virtual laboratories provide quick access to course material quicker and easier than conventional approaches, with material presented during seminars available for review and re-examination at anytime, anywhere on an Internet-enabled PC. One consultant stated that trainees often request access to slides again, prior to examinations (MRCPath), but that it is often difficult to locate the slides quickly,

and access is restricted to only one person at a time (Jeffers, 2005) (Jeffers, 2005). Having material available online removes these issues.

All surveyed trainees considered the Digital SlideBox™ easy to use (Figure 7.9) and found virtual slide and expert annotations of benefit in training (Figure 7.12). One trainee commented:

> *"On the whole I think it's excellent. A more realistic way to learn than just reading."*

However the system is not without drawbacks. Two trainees commented that the Digital SlideBox™ was an:

> *"Excellent facility with lots of potential- only drawback currently is speed (v. glass slides)"*

and

> *"the only problem is the length of time to load images."*

When comparing speed of virtual to glass slides, virtual will always compare unfavourably, especially if the user has little experience with virtual slides, as a period of time is usually required to become familiar with new technology. However, as computers and bandwidth improves, speeds will improve. During the inception of telepathology it took several minutes to transmit static images via email, whereas broadband now enables virtual slide images to be refreshed in a fraction of a second (SlidePath Ltd, 2004). It is not unreasonable to suggest that, given the continual improvements in computer processor speeds, computer memory, graphics cards and bandwidth, that virtual slide image refresh rates will compare favourably with light microscopy.

Trainees were favourable regarding expert annotations, with almost 90% considering them of great benefit. It is interesting to note that trainees expressed a greater preference for expert annotations rather than expert examination videos.

As illustrated in Chapter 5, almost 80% of surveyed expert pathologists considered examination replays of benefit in pathology training, while 66% of trainees considered then of benefit. However it should be highlighted that trainees did not have any experience of examination replays, and 33% of trainees did not express a preference (Figure 7.13). 89% trainees favoured the incorporation of both Multiple Choice Questions (MCQ) and links to relevant publications. In open-ended questions, a number of trainees requested access to seminar presentations.

Ultimately, this study was a pilot study to assess Irish trainee pathologist's perceptions of computer-assisted learning. Feedback was positive, illustrating that Irish trainees express a similar enthusiasm for supplementary electronic resources as their international counterparts (Klatt, 1997, Kronz *et al.*, 2000). The General Medical Council's education committee (General Medical Council, 1993) recommended that modern technologies should be utilised in pathology education where a discernable benefit is observed, and this study suggests that continued utilisation of computer-assisted learning may be of long-term benefit to Irish pathology training.

Ultimately, CAL will never replace one-to-one human tutoring, however it will be increasingly used to supplement, rather than replace existing teaching methodologies. The expansion of online facilities nationally, or even internationally, increases the potential of online facilities exponentially. Such networked systems would provide trainees with access to a broader education, and improve uniformity of training standards nationally and internationally.

Chapter 8: Conclusion

Conclusion

The light microscope remains, as it was 100 years ago, the principal diagnostic tool in the pathologists' armoury. The searching, perception and problem-solving skills demonstrated by pathologists during microscopic examination have also remained predominately unchanged, however the methodologies used to develop and maintain these skills are constantly evolving.

Many institutions have modified their approach to medical education, migrating from a teaching centred approach, in which the emphasis is on the teacher and what they teach, to a student-centred approach, were the focuses lies on students and what they learn. This is intended to promote self-directed learning, encourage deep learning of underlying principles of topics rather than learning 'by rote,' and ensure that students take an active role in their education.

Continuing medical education (CME) also encourages experienced pathologists to develop lifelong learning skills, in order to adapt to the changing needs of both patients and the healthcare profession.

Computer-assisted learning (CAL) is ideally suited for both the training and continuing education of pathologists. CAL systems can inform, assess and support users, and require pathologists/trainee pathologists to actively participate in the learning process.

In 1986, Weinstein proposed a 15-year timeline for the refinement, validation and implementation of telepathology. While telepathology has not yet been widely integrated into pathology practice, technological improvements are rapidly making it possible that telepathology will follow in the pioneering footsteps of radiology. Images that previously took minutes to transmit over the Internet now take less than a second. Advancements in computing technology mean that large quantities of visual data can be rapidly processed and displayed on large, high definition screens in true colour, and in high resolution. The evolution of ultra-fast slide scanners has facilitated high-throughput, automated

slide digitisation that can feasibly be incorporated into the pathology laboratory, raising the possibility of an entirely virtual pathology laboratory.

However, pathologists are still reluctant to use telepathology in everyday practice (Dennis *et al.*, 2005). Studies have shown diagnostic accuracy with telepathology systems to be comparable to glass slide diagnosis (Cross *et al.*, 2002), yet few pathologists regularly use a such systems. While pathologists cite concerns over ethico-legal issues and image quality, much is due to lack of experience and knowledge of telepathology.

Studies have shown that age and experience of telepathology can affect acceptance and performance with telepathology systems, with those more computer literate more eager to exploit new technologies. Some commentators have even cited prior experience with video games as a possible indicator of favourable perceptions towards telepathology systems. If early exposure to computers is associated with improved acceptance of telepathology, the incorporation of computer-based resources into pathology training may result in a creation of a generation of young pathologists enthusiastic about telepathology.

Two key objectives of this body of work were (1) to identify whether supplementary electronic resources could improve performance, or concordance with a gold standard (2) whether VPS tracking data could be utilised to identify sources of error in histological assessment. In order to achieve these objectives, a number of software applications were developed during the course of this work, and subsequently evaluated in appropriate environments.

The ReplaySuite was developed as a telepathology tool for use in training and education. Designed to replicate the functionality of a double-headed microscope, it enables one pathologist to review an examination conducted by another pathologists in a manner similar to watching a video. Importantly however, it removes temporal and spatial issues that surround the use of double-headed microscopes, by allowing examinations to be reviewed at different times and in different locations to the original examination.

The ReplaySuite was initially evaluated by a group of expert EU breast pathologists who assessed its easy of use and potential applications in pathology. All participants who replayed at least one examination agreed or strongly agreed that the ReplaySuite was user-friendly, while all participants that replayed more than 10 examination replays considered the ReplaySuite of benefit, or of great benefit in both pathology training and External Quality Assurance. Significantly, two pathologists reconsidered their original diagnoses subsequent to reviewing examination replays, with both re-diagnosing either concordant with group consensus or glass slide diagnosis.

The results obtained from the preliminary evaluation of the ReplaySuite (Chapter 5) provided an interesting insight into its potential to make pathologists reassess discordant diagnoses. Subsequently, the ReplaySuite was incorporated into an External Quality Assurance study of staging and grading of chronic hepatitis in Sweden. While the scheme was conducted to assess the level of uniformity in staging and grading, it was also conducted to ascertain whether the use of supplementary electronic material could improve diagnostic performance. Access to two electronic resources, the ReplaySuite and an online image reference library (atlas) provided by the Swedish Society of Pathology, were provided between two scoring sessions. Diagnoses were recorded, enabling comparison between diagnostic performances in different scoring sessions.

Between scoring sessions, mean staging un-weighted kappa improved from 0.347 to 0.554 (+0.207), or from 'fair' to 'moderate' exact agreement with KVAST staging, while mean grading un-weighted kappa increased from 0.132 to 0.412 (+0.280), or from a 'poor' to 'moderate' level of exact agreement with KVAST. Mean group staging and grading improved for both staging (+18%) and grading (+20%), suggesting that the use of supplementary online resources can improve diagnostic performance in grading and staging of chronic hepatitis.

In order to assess the ReplaySuite in an External Quality Assurance setting, it was required to digitise a large numbers of slides using a customized microscope with motorised stage. Pre-existing slide scanning algorithms were insufficient to deliver the volume of slides required, and the redeveloped of scanning

algorithms was necessary to provide a more rapid, robust and flexible scanning regimen. The incorporation of a PreScan component to the revised system significantly reduced scanning times, and subsequently, it has been observed that this approach is now widely used by a number of ultra-fast slide scanners.

A new Virtual Pathology Slide (VPS 2.0) slide viewer was developed in tandem to facilitate online examination of slides digitised with the new scanning algorithms. Additional modifications were incorporated, the most significant modifications relating to annotation capabilities, and alteration of the system to display slides digitised with the new slide scanning algorithms. A web application delivered via a customised Internet browser, it utilised a number of client and server side Internet technologies, such as PHP, Oracle 8*i* relational database management system, SQL, DHTML and JavaScript. These technologies were originally selected for use in the development of VPS 1.0 (Costello, 2004) due to their open source nature and widespread use, and were subsequently retained for development of VPS version 2.0 (Chapter 4). The incorporation of StyleSheets facilitated the separation of content (information displayed onscreen) from presentation (colours, fonts and layouts used). Internet Explorer 5.0 or higher was required to use the VPS 2.0. At the time, Internet Explorer possessed 95% of the web-browser market, however this limitation still prohibited Netscape, Mozilla and Apple users from accessing the system.

A new version of the ReplaySuite (version 2.0) was also developed specifically for the EQA study. It retained the original web-based, customized browser architecture, but incorporated new features pertinent to the EQA study. In addition, novel SnapShot functionality was incorporated that enabled 'heat maps' to be generated from examination data. These maps comprised colour-coded regions overlayed on slide overviews to indicate areas examined, with different colours used to differentiate between regions examined at different magnifications.

Subsequent to the chronic hepatitis study, an expert pathologist reviewed study examinations via the ReplaySuite. This served two purposes:

- Ascertain whether concordant examinations utilised appropriate examination technique more frequently than discordant diagnoses, where examination technique pertained to adequate tissue coverage at appropriate magnification.

- Identify sources of error in histological assessment of chronic hepatitis, where sources of error are identified as oversight (missed pathology) and misinterpretation.

Examination reviews identified that, when grading, 28% (46% - 18%) more concordant than discordant examinations were considered to have viewed sufficient tissue, and at the appropriate magnification. 24% (59% - 35%) more concordant (Vs discordant) examinations were conducted appropriately, suggesting that examination technique is important both when determining the degree of necroinflammation within a biopsy, and when ascertaining the extent of fibrosis.

Additionally, the expert reviewer was able to identify a source of error in 58% and 50% of staging and grading examinations respectively. Missed pathology (oversight) was observed as the most likely source of error in 63% of discordantly grading examinations where a source was identified, while this value was lower for staging (52%). While it is acknowledged that it should be possible to identify a source of error in all examinations and that experimental design may have prevented this somewhat, the findings illustrate that it is possible to use VPS tracking data to identify sources of error in histological assessment.

This study highlighted the significant potential of the ReplaySuite in EQA, and its enormous potential to provide added value to EQA studies. While conventional schemes are restricted to identifying the occurrence of errors, use of VPS and ReplaySuite applications facilitated not only the identification of errors, but also sources of error in diagnostic technique.

Subsequent to the chronic hepatitis study it was determined that it would be beneficial to evaluate the technology in a training environment. Participation in a postgraduate pathology seminar series facilitated this, however such an undertaking raised several significant issues:

- A large number of digital slides were required in a relatively short period of time.

- The existing slide scanning algorithms was not able to manage the required throughput.

- Use of an ultra-fast slide scanner would be capable of providing the required throughput, but would prohibit the utilisation of ReplaySuite technology.

It was decided that use of an ultra-fast slide scanner was the only feasible approach, as the customised scanning system developed in-house was at this time outmoded and somewhat obsolete, with respect to ultra-fast scanner technology. While this precluded ReplaySuite evaluation in a training environment, it facilitated the assessment of Irish trainee pathologist's perceptions of computer-assisted learning, with particular emphasis on digital slides and expert annotations. 78 slides were digitised and presented during formal teaching seminars, and subsequently made available online to trainees, in addition to extensive expert annotations. Trainees were favourable regarding both digital slides and expert annotations, with all trainees considering digital slides of some or of great benefit, and almost 90% considering slide annotations of great benefit.

This work illustrates the benefit of computer-assisted learning in both training and ongoing education. Both expert (Chapter 5) and trainee pathologists (Chapter 7) considered such systems of benefit, concurring with other studies that have encountered similar enthusiasm towards online resource (Klatt, 1997, Jones et al., 2002), illustrating that such tools are not being forced onto an

unwilling student population. Quite the contrary, some researchers have observed that trainees actually prefer the use of digital slides to glass slides (Dee et al., 2003).

Secondly, and more significantly, using supplementary online material can improve the quality of education provided and improve diagnostic performance. Two expert pathologists reconsidered their original diagnoses after reviewing examinations via the ReplaySuite (Chapter 5), while improvement in both staging and grading was observed between chronic hepatitis scoring sessions (Chapter 6). Similar observations have been published regarding CAL; Kronz *et al* (2000) illustrated that the use of a Web-based tutorial is capable of improving diagnostic performance in Gleason grading of prostate carcinoma, while Klatt (1997) commented on improvement in pathology examination scores subsequent to the incorporation of computer-assisted learning in the University of Utah. However not all electronic resources may be considered success stories; Smith *et al* (2000) concluded that use of the Bethesda system atlas by itself does not appear to improve the reproducibility or accuracy of cytological diagnoses.

With regard to the type of content provided by online resources, it is interesting to note that while trainee's evaluating the Digital Slidebox^{TM} expressed a greater preference for expert annotations than expert examination videos (Chapter 7). As illustrated in Chapter 5, almost 80% of surveyed expert pathologists considered examination replays of benefit in pathology training, while 66% of trainees considered then of benefit (Chapter 7). However it should be highlighted that during their evaluation of the technology, trainees did not have any experience of examination replays, and 33% of trainees did not express a preference (Figure 7.13). Also, preliminary evaluation of the ReplaySuite version 1.0 (Chapter 5) indicated that experts who replayed 11 or more examination replays were more appreciative of their benefits than those that replayed 5 or less, suggesting that the more replays are accessed, the more beneficial they are seen to be.

As previously stated, image quality is often cited as a concern regarding telepathology systems. During the course of this work, three separate groups of pathologists, expert pathologists preliminary evaluating the ReplaySuite 1.0

(Chapter 5), laboratories participating in the chronic hepatitis EQA scheme (Chapter 6) and Irish trainee pathologists in the postgraduate seminar series (Chapter 7) each utilised a different software application (or version of the same application) to review digital slides. In addition, each set of slides examined was digitised using a different scanning methodology. However for each group, image quality was favourably evaluated. This illustrates that it is possible to provide images of sufficient quality that pathologists are confident in their diagnosis, validating the observations of other researchers (Cross *et al.*, 2002, Costello, 2004).

Almost 80% of participants evaluating ReplaySuite 1.0 (Chapter 5) considered the system to be user-friendly, while all trainees considered the Digital Slidebox™ easy to use (Chapter 7), and almost 40% considering it very easy to use. In contrast, system speed remains an issue. The majority of pathologists evaluating the ReplaySuite 1.0 (Chapter 5) that expressed a non-neutral opinion of system speed considered it slow. The majority of laboratories using the VPS 2.0 and ReplaySuite 2.0 (Chapter 6) also considered the system slow, however those that utilised DVD's provided a more favourable evaluation, with the majority considering the speed acceptable. Trainees using the Digital Slidebox™ were more complementary (Chapter 7), with slightly more trainees positive about system speed than negative. Direct comparison between digital slide and glass slide speed will always favour glass slide examinations. However, it should be highlighted that evaluation of ReplaySuite 1.0 (Chapter 5) occurred at a time when broadband speeds were not commonly available, and participants in the chronic hepatitis study (Chapter 6) were accessing the server located in Ireland from Sweden. Trainee's utilising the Digital Slidebox™ (Chapter 7) however were not subjected to either of these limiting factors, and their perceptions of speed may be a more accurate reflection of current technology.

The future development of the Digital SlideBox™ may provide an indication of the path that computer-assisted learning systems in pathology may take. It is proposed that the Digital SlideBox™ will be redeveloped to integrate VPS and ReplaySuite functionality, provide access to seminar presentations, relevant publications and MCQ self-assessments to provide a comprehensive resource to

trainee pathologist's (O'Shea, 2005). The method of creating annotations files will be re-evaluated, as the use of ImageScope is less than intuitive, and requires additional uploading of annotation files to Digital SlideBox™. Additionally, the possibility of incorporating decision support software (DSS) has tremendous potential for the system, as it would facilitate self-directed active learning by requiring users to actively participate with the material provided.

It is proposed that this application will become a multi-functional resource, providing a wealth of facilities to trainees and experienced pathologists alike. It may ultimately evolve into a dichotomous application for Continuing Medical Education (CME) and postgraduate training, with the following two examples illustrating different utilisations of effectively the same system:

An experienced pathologist logs onto the Digital SlideBox™ from home and registers for an accredited EQA scheme, paying by credit card. He/she examines the available clinical data and digital slides for each case, focusing through a clump of tissue using z-stack capabilities, before submitting a diagnosis. Feedback on diagnostic performance is rapidly received from the central server, in addition to extensive evaluation of their examination technique; which important areas they successfully/failed to examine, what relevant pathological features they successfully/failed to identify. The pathologist participates in these EQA schemes several times a year (due to rapid execution), with the occasional incorporation of tissue sections that are usually difficult to incorporate into traditional EQA schemes, such as needlecore biopsies (due to limited tissue). Finally, the pathologist reviews links to new and relevant publications that are displayed with the performance data, in addition to several educational presentations provided.

In contrast, a trainee pathologist log onto the Digital SlideBox™ system and searches for a specific organ, diagnosis or pathological feature. The search function retrieves a number of relevant multimedia items: a recent seminar presentation, a tutorial with extensive notation images illustrating important pathological changes and links to relevant publications. Pertinent digital slides are also retrieved that may be examined with annotations switched on or off. The

trainee then accesses the incorporated decision support system (DSS) to assess their ability to accurately evaluate diagnostic clues while examining the digital slides, before using a self-assessment MCQ. Finally, once fully prepared, the trainee logs onto the examination assessment module, which displays a number of digital slides with accompanying clinical data. Pulling on a headset, the trainee examines the slides, whilst describing the diagnostic process being used. The examination data (and audio) is uploaded onto the remote server, from where it will be accessed and evaluated by an experienced pathologist at a later date, in order to assess the student's diagnostic performance. Finally, the trainee accesses a video of an expert's examinations of the cases (only available after the trainee's examination), which provides instantaneous feedback to the trainee as to their performance.

Ultimately, whether this scenario becomes a reality depends on the pathologists who will be required to use it. This work illustrates that computer assisted learning can enhance pathologists training, and improve diagnostic performance. For pathologists to reap the benefits such systems, they must be embraced. If this occurs, and technology continues to evolve at its current speed, the light microscope may finally be supplanted by the humble PC as the most important piece of equipment available to pathologists.

Appendices

Appendix A

Description of technologies and programming languages used to construct and deliver Virtual Pathology Slides and ReplaySuite

Web-Based Languages and Technologies

HTML

Based on SGML (Standard Generalized Mark-up Language), HTML is the standardised programming language used by Internet browsers to display WebPages. HTML documents are text files (denoted by the .htm or .html suffix) that incorporate tags, or elements, that instruct the browser how to display static text and images. Hyperlinks are elements within documents that connect to other elements in the same or other documents, enabling navigation between HTML documents, and the creation of paths between associated information. Tags also allow objects to be embedded within them, such as Java Applets, Cascading Style Sheets (CSS) and JavaScript function libraries. By embedding objects and libraries within documents, a greater degree of control is created, allowing the properties of elements to be modified dynamically, and providing the user with a greater degree of interaction.

JavaScript

JavaScript is a cross platform, object-oriented scripting language that incorporates client side, dynamic user interaction into HTML documents. JavaScript libraries can be embedded directly within the head of HTML documents, or linked as external JavaScript files (denoted by the .js suffix). When a HTML document is loaded, all JavaScript functions and variables embedded with the file are loaded into the cache. These functions may then be called when a client-side event occurs, such as moving the mouse cursor over, or clicking on, a specific element. JavaScript is used to incorporate client side functionality within a webpage, creating a greater level of interactivity than is capable with HTML alone. While JavaScript has the potential to dynamically generate HTML, it is strictly a client side scripting language; it cannot interact with a database, or dynamically generating HTML before the page is loaded into the cache.

StyleSheets

Cascading StyleSheets (CSS) are used to control the appearance of a webpage. They allows greater control and flexibility in how the page is displayed, by separating content (information displayed onscreen) from presentation (colours,

fonts and layouts used). A StyleSheet contains a list of rules governing the way an associated webpage should be displayed. Files accesses StyleSheets (denoted by the .css suffix) by linking to it with the inclusion of the following code:

```
<link rel="stylesheet" type="text/css" href="css/stylesheet.css">
```

Where stylesheet.css is the CSS file. Once linked, the page presentation is governed by the style rules, associated with specific HTML elements or 'tags', such as <body> or <p>. Rules take the following form:

```
selector { property: value }
```

for example:

```
body   {background-color:#2663E0;}
```

Alternatively, user-defined selectors may be created, which are referenced within the html document via the class property. For example the following user-defined selector:

```
.datavals {
        font-family: Arial, Helvetica, sans-serif;
        font-size: 12px;
        font-style: normal;
        color:#FFFFFF;
    }
```

Which is referenced by an element using the class property:

```
<font class="datavals">
```

PHP

PHP (PHP Hypertext PreProcessor) is an open source, multi-platform, server side scripting language. It is capable of executing commands on the server,

enabling it to dynamically generate HTML, and facilitate communication with databases. When a server supporting PHP receives a PHP file (appended with the .php or .php3 extension), the contents are parsed until PHP tags (<?*php* and ?>) are located. Commands within these tags are processed and executed, allowing the insertion of static HTML code as instructed, dynamically generating HTML prior to transmission.

PHP supports a wide range of databases, such as mySQL, Oracle and ODBC. A relational database is a collection of data organised into pre-defined tables, from which data can be accessed or reassembled in many different ways without having to reorganise the database tables. Databases may be located locally on a client PC, or remotely on a server. To interact with relational databases, PHP utilises function libraries that pass questions or *queries* to the database. These queries are written in Structured Query Language (SQL), the standardised language for communicating with relational databases.

XML

XML (extensible markup language) is a W3C mark-up language, similar to HTML. It differs significantly to HTML however, in that instead of being used to *display* data, it is used to *describe* data. Created so that richly structured documents could be used over the web, XML complements HTML rather than replaces it.

SQL

Structured Querying language (SQL) is a an ANSI (American National Standards Institute) standard language for querying and modifying relational databases. Developed by IBM in the 1970's, it supports major keywords such as SELECT, UPDATE, DELETE, INSERT & WHERE, and utilises the following syntax:

SELECT columnname FROM tablename WHERE columnname2 = value

For example:

Select firstname FROM employees WHERE surname = jones

Retrieves all values from the firstname column in the employees table that have the surname Jones. Tables in a database can be related to each other with keys. A primary key is a column with a unique value for each row. For example, two tables, *employees* and *managers,* may contain a common key *department.*

Employees

Firstname	Surname	Salary	Department
John	Cole	30000	IT
Peter	Song	20000	Administration
Mary	Toure	50000	IT
Alex	Jones	25000	IT

Managers

Firstname	Surname	Department
John	Cole	IT
Peter	Song	Administration
Mary	Toure	Sales
Alex	Jones	HR

To identify the names of managers of departments with any employee earning more than 30,000 salary the following query would be used to join the two tables together:

SELECT managers.firstname, managers.surname from managers, employees WHERE employees.salary > 30000

Oracle

Oracle 8i is an object relational, database management system, which can be interacted with in a number of ways, such as via the Oracle Call Interface (OCI), a set of low-level APIs. The OCI code contains all the information required to initiate a SQL dialogue between the ReplaySuite and the Oracle database. It

defines calls to the server to parse SQL statements for syntax validations and open a cursor for the SQL statement. It then binds client application variables into the server shared memory and describes the contents of the fields being returned based on the values in the servers data dictionary. It executes SQL statements within the cursor memory space, and fetches one or more rows of data into the client application, before closing the cursor.

MySQL

MySQL is an open source relational database management system (RDBMS) that uses Structured Query Language (SQL) to read and write data to the database, in a similar manner to Oracle. In contrast to Oracle, however, it does not require a licence to use.

Non Web-Based Languages and Technologies

C++

C++ is a compiled object-oriented programming (OOP) language, derived from C, used to write applications for nearly every available platform. Classes are a fundamental construct in C++. Each class groups its own data, the class members, with functions that act on them, the methods of the class. Members and methods of a class may be accessible to other parts of a program (public) or only directly by other members of the class (private), enabling interaction between parts of a program while protecting class data.

Dynamic Linked Libraries (DLL) are libraries which are linked to application programs when they are loaded or run, rather than as the final phase of compilation. This means that the same block of library code can be shared between several tasks rather than each task containing copies of the routines it uses.

Microsoft Foundation Class (MFC) is a Microsoft library that wraps portions of the Windows API (Application Programming Interfaces) in C++ classes, forming an application framework. MFC applications can use Internet Explorer file libraries to behave as a customised browser. This enables applications to locate the VPS Web server on the Internet and request a page, which is retrieved

through the network and displayed on the client machine. It also informs how to interpret the set of HTML (Hyper Text Mark-up Language) tags within the file, in order to display the page onscreen.

ALI

Analytical Language for Images (ALI) is the macro control language for OPTIMAS. It has a similar syntax to C, but is not an implementation of C. ALI is an interpreted language, with statements executed one-by-one as presented to the macro parser. Despite the interpreter overhead, programs are extremely fast for many problems typical of analysis of image measurements. This is because such problems often can be naturally expressed in vector terms. Vector Language Functions and operators in ALI perform vector operations and return vector results. For example, an area boundary is a vector of coordinate pairs. Thus, many loops found in typical C programs are unnecessary in the equivalent ALI program.

Appendix B

Sample programming code from VPS (v2.0) Slide Viewer and ReplaySuite (v1.0 & 2.0)

Virtual Pathology Slide Version 2.0

The following sample code demonstrate how PHP is used to access the remote Oracle database, write data to the database and generate filepaths for images that are displayed via the VPS 2.0.

Database Access via PHP

Establishing a connection with the database is achieved via Oracle Call Interface (OCI).

```
$connection = ocilogon('useraccount','userpassword', 'webstuff.dcu.ie');
$query = "Select USERID FROM USERMASTER where USERNAME = '$user'
and PASSWORD ='$pass'";

$cursor = OCIParse ($connection, $query);
$result = OCIExecute ($cursor);

while (OCIFetchInto($cursor, &$values)){
        $myvalue=$values(Crowley et al.);
        echo $myvalue;
        }

OCICommit ($connection);
OCILogoff ($connection);
```

Tracking Examinations

The browser intercepts each new field request during an examination. Variable data describing the request is transmitted via CHTTP to a PHP file *(populate_db.php)* which uses an SQL query to write data to the database in two stages:

A new row is created in the *trackingequalis* table via INSERT statement each time a new field is requested. This takes the following syntax:

INSERT INTO tablename (column-1, column-2, ... column-n) VALUES (value-1, value-2, ... value-n)

This is used to insert field coordinates (*stain, mag, x, y*) and examination keys (*sessno, examindex*):

INSERT INTO TRACKINGEQUALIS (SESSNO, EXAMINDEX, STAIN, MAG, X, Y) VALUES ('$Fsessno', '$myvalue', '$Fstain', '$Fmag', '$Fx', '$Fy')"

If this is the first field viewed, the value inserted into *examindex ($myvalue)* is 0. All subsequent *examindex* values are calculated prior to insertion by retrieving the highest *examindex* value for the current session number (*sessno*) and incrementing it by 1.

"Select max(EXAMINDEX) FROM TRACKINGEQUALIS WHERE SESSNO = '$Fsessno'"

$myvalue++;

While information is recorded each time a new field is requested, data pertaining to the last field viewed is also tracked via an UPDATE statement. This is used to update a pre-existing row within the *trackingequalis* table, and takes the following syntax:

UPDATE tablename SET column = expression WHERE conditions

The UPDATE statement is used to insert data into the row that describes the previous field, updating how much time spent was viewing it, what type of navigation was requested to move from it, and whether the field was annotated. *Sessno* and *examindex* are used as keys to locate the correct row within *trackingequalis*, for example:

"UPDATE TRACKINGEQUALIS SET TIME = '$Ftime' WHERE SESSNO = '$Fsessno' AND EXAMINDEX = '$myvalue'"

Image Filepaths generation via PHP

Both VPS frames are passed case and stain variables. In addition, the VPS frame is passed the following:

layer="+layer+"&x="+x+"&y="+y

Where

var layer=0;
var x=101;
var y=101;

The VPS page uses these variables to determine the appropriate images to display, where *layer* is the magnification, and *x* & *y* the image coordinates. Depending on whether images are to be retrieved from the server or from DVD, one of two functions will be called, which will return an array of image filepaths.

If images are to be retrieved from the DVD, *local_image_filepaths()* is called, if not, *image_filepaths()* is called. These functions generate image filepaths for each image to be displayed, based upon the parameters provided.

```
function image_filepaths($rootpath, $localroot, $case, $layer, $x, $y, $xend, $yend) {
    $number=1;
```

```
        for ($yn=$y; $yn<$yend; $yn++)     //sets the y coordinates for initial
images
                {
                for ($xn=$x; $xn<$xend; $xn++) {  //sets the y coordinates for
initial images
        (!file_exists($localroot.$case."/build".$layer."/".$xn."k".$yn.".jpg"))?($i
mage[$number]=$rootpath.$case."/build".$layer."/blank.jpg"):($image[$number
]=$rootpath.$case."/build".$layer."/".$xn."k".$yn.".jpg");          //sets image
src to blank if file doesn't exist or else correct x,y filepath
                                $number++;  //increases $image by one each time
                        }
                }
        return $image;
        }

function local_image_filepaths($rootpath, $localroot, $case, $layer, $x, $y,
$xend, $yend) {
        $number=1;
        for ($yn=$y; $yn<$yend; $yn++)     //sets the y coordinates for initial
images
                {
                for ($xn=$x; $xn<$xend; $xn++) {  //sets the y coordinates for
initial images
      $image[$number]=$rootpath.$case."/build".$layer."/".$xn."k".$yn.".jpg";
                        $number++;      //increases $image by one each time
                        }
                }
        return $image;
        }
```

PHP is used to dynamically create a table, into which image elements are placed. Each image is assigned a name, e.g. 'image1'. Finally, the image elements are linked with the previously generated filepaths, displaying the images onscreen:

```
<?php
```

```
for ($ii=1; $ii<1+pow($lim,2); $ii++)
echo("image$ii.src=imagesArray[$ii].src;");
?>
```

This PHP is placed within JavaScript <script> </script> tags so that the code will be executed client-side. Code execution loads images into the cache, and then displays them onscreen.

Navigation via JavaScript

The following describes how the VPS uses JavaScript to provide client-side navigational functionality within the VPS.

Keyboard Navigation

Pressing a key on the keyboard calls the *KeyPressed()* JavaScript function, via the *onKeyDown* JavaScript event handler. A navigation function is then called via a *Switch()* multi-if statement, using the *keyCode* event property (*window.event.keyCode)* as the option selector. For example, clicking the '6' key request right lateral motion. This calls the *lateral_js(scrll, layr, x, y)* function, where *layr, x & y* correspond to the current fields coordinates, and *scrll* indicates the requested navigation (right lateral motion: *scrll* = 7). Another *switch* statement is used to modify the image coordinates, with *scrll* as the option selector. In this case, only the x coordinate is modified (*x+=i*), and the page refreshed with the new image coordinates.

```
url="vps.php3?scroll="+scrll+"&layer="+layr+"&x="+x+"&y="+y+"&case="+c
aseno+"&time="+time+"&drive="+drive+"&slidenum="+slidenum;

document.location.href = url;
```

Mouse Navigation

The *onmousedown* JavaScript event handler is used to control all mouse navigation. Clicking within the main field of view calls the JavaScript *onmouseclick()* function, which executes code that does the following:

- Determines *x,y* screen coordinates of the point clicked via the *clientX, clientY* event properties.
- Determines the *x,y* coordinates of the image clicked via the image *src* property.
- Ascertains which mouse button was clicked via the *button* event property.
- Uses the button clicked to select *switch* statement options.
- Executes a navigation function based on which zoom method is currently active

Point Select Navigation

If the 'Point Select' zoom function icon is active, clicking on a point will centre the field on that point and double the magnification. The *short_zoom()* function is called, which increments the magnification folder by 1 (*layer)* and doubles the *x,y* coordinates to provide the correct coordinates for the selected area at higher magnification:

```
layer++;
xy_array(Crowley et al.)=2*(xy_array(Crowley et al.)-101)+101;
xy_array[1]=2*(xy_array[1]-101)+100;
```

The new *x,y* & *layer* coordinates are incorporated into a new URL, which is requested via the *href* event property:

```
url="vps.php3?scroll="+scrll+"&layer="+layer+"&x="+xy_array(Crowley et al.)+"&y="+xy_array[1]+"&case="+caseno+"&time="+time+"&drive="+drive+"&slidenum="+slidenum;

document.location.href = url;
```

Maglist Navigation

Accessing the mag-list is achieved via the VPS JavaScript function library. Right mouse-clicking on the main field of view calls the *show_menu()* function, which

displays a semi-transparent div to the right of the location of the mouse cursor. The *BasicImage* Microsoft static visual filter is used to set the opacity level of the div, and CSS *posLeft* and *posTop* object attributes are used to position the div based on the location of the mouse cursor when the right button clicked.

```
style.filter='progid:DXImageTransform.Microsoft.BasicImage(opacity=.7)';
style.posLeft=menu.style.posLeft-50
style.posTop="+ym+"+"+y+"-"+adjusty+";"
style.visibility="visible";
```

Hovering the cursor over 'Maglist', located at the top of the list will call the *show_menu()* function again, but pass the submenu div element in as an argument, displaying the submenu div menu beside the main menu. The current magnification is highlighted in light blue; hovering the cursor over a magnification option will change the background via the *onmouseover* event handler:

```
onMouseOver="this.style.backgroundColor='#B9D1E2'"
```

Selecting a magnification from the list displays the right-mouse clicked point at the magnification selected by calling *zoom_function()* (via *zoom_menu()*). This uses the *pixel_to_coordinate()* function to convert *x,y* coordinates of the selected point to new image x,y coordinates. This achieved via a series of phases. The following description of this code includes an example illustrating how new image coordinates are calculated when navigating from 4x to a selected point at 8x via the mag-list. Figure A.1 graphically illustrates the change in field of view.

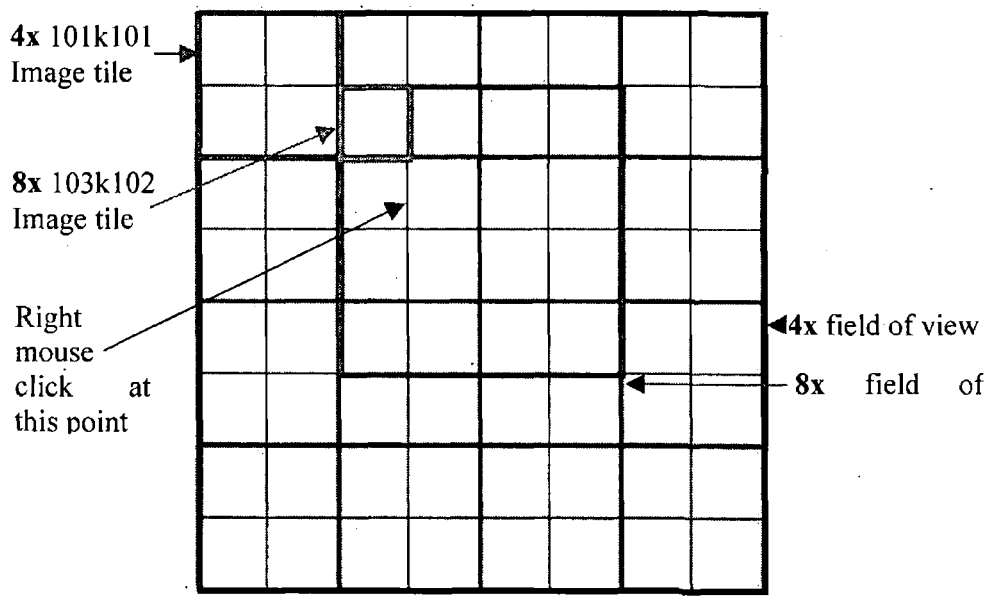


Figure A.1 Using the mag-list to zoom from 4x to 8x

The dimensions of the current and new magnification layers in pixels are determined:

*layerwidth_p = 768*Math.pow(2, layer);*
*newlayerwidth_p = 768*Math.pow(2, newlayer);*

Current layer pixel width: 768 x 2^0 = 768 pixels
New layer pixel width: 768 x 2^1 = 1536 pixels

The width of the current and new magnification layer in image tiles is then calculated:

*layerwidth_c = 4*Math.pow(2, layer);*
*newlayerwidth_c = 4*Math.pow(2, newlayer);*

Current layer image width: 4 x 2^0 = 4

New layer image width: $4 \times 2^1 = 8$

The position of the pixel in relation to the entire layer is calculated using the current image *x* coordinate (*basex*), the onscreen mouse coordinate (*mousex)* and the layer pixel width (*layerwidth_p)*. Where *scan_jum_no[layer]* array is used to compensate for variation in image nomenclature between 20x and 40x scans, and *lim* refers to the onscreen width in image tiles.

xy_array(Crowley et al.) = ((basex-101-scan_jump_no[layer])(768/lim) +mousex)/layerwidth_p;*

$X = [(101 - 101 - 0) \times (768/4) + 250] \div 768 = 0.32$

Pixel positioning is then converted to image file coordinates. The JavaScript mathematical function *Math.ceil* is used to round up the number to the nearest integer (e.g. 1.6123 → 2), and the *inc* variable is used to centre the new images.

xy_array(Crowley et al.)=Math.ceil(xy_array(Crowley et al.)(newlayerwidth_c))+101+(scan_jump_no[newlayer])-inc;*

$X = \text{math.ceiling}(0.32 \times 8) + 101 + 1 - 2 = 103$

A new image *y* coordinate is calculated in the same manner, but height in pixels displayed onscreen (574) replaces width (768). The new *x,y* coordinates are then returned to *zoom_function()* as an array (*xy_array[]*) and incorporated into a new URL, with *layer* set to *newlayer.*

```
url="vps.php3?scroll="+scrll+"&layer="+newlayer+"&x="+xy_array(Crowley
et
al.)+"&y="+xy_array[1]+"&case="+caseno+"&time="+time+"&drive="+drive+
"&slidenum="+slidenum;
document.location.href = url;
```

NewLayer = 1 (8x)

New X = 103

New Y = 102

When the requested magnification is lower than the current magnification, the *macrocoordinates()* function is called, in the place of *pixel_to_coordinate()* to generate new image coordinates. For each stepwise decrease in magnification, this function effectively doubles the coordinates of the required images, making modifications to improve centring, and the JavaScript mathematical function *math.round()* is used to round the calculations to the nearest integer:

:

```
for(i=layer; i>newlayer; i--)  {
        if(i==folder.length-2) {        //when at 1000x
        xy_array(Crowley et al.)--;             //improves centring
                xy_array[1]--;
                }
        if(i<folder.length-2)  {
                xy_array(Crowley et al.)=Math.round((xy_array(Crowley et al.)-
101)/2+99);             xy_array[1]=Math.round((xy_array[1]-101)/2+100);
                }
        }
```

ZoomBox Navigation

1) Creating a ZoomBox

The initial click calls the *div_edit()* function, which sets the top & left position of the ZoomBox div element to the mouse *x,y* coordinates, and width/height to 1:

```
div.style.posLeft=event.clientX;
div.style.posTop=event.clientY;
div.style.posWidth = 1;
div.style.posHeight = 1;
```

The *div_opaque()* function is used to set the div appearance, creating a small semi-transparent blue div beside the cursor:

```
div.style.visibility='visible';
        div.style.borderWidth='0.0cm';
        div.style.borderStyle='groove';
        div.style.borderColor=colour;
        div.style.backgroundColor= colour;
        div.style.filter='progid:DXImageTransform.Microsoft.BasicImage(opacity=.1)';
```

The *show_div_dimensions()* function is then called, which saves the coordinates to a html form *dataform* within the *controlpanel* frame.

```
parent.controlpanel.dataform.zoomx.value = div.style.posLeft;
parent.controlpanel.dataform.zoomy.value = div.style.posTop;
parent.controlpanel.dataform.zoomwidth.value = div.style.posWidth;
parent.controlpanel.dataform.zoomheight.value = div.style.posHeight;
```

2) Resizing the ZoomBox

The *onmousemove* event is used to track the location of the mouse cursor while over the main field of view. The event calls the *move()* function, which changes the size of the box (*style.posWidth, style.posHeight*) as the use moves the mouse cursor, via the *clientX, clientY* event properties.

```
div.style.posWidth = event.clientX - div.style.posLeft-10;
div.style.posHeight = event.clientY - div.style.posTop-10;
```

3) Selecting a region with the ZoomBox

The second left mouse button click calls *div_edit()* again, recording the final box dimensions via *show_div_dimensions()*. It then calculates the new magnification layer by calling *zoombox_mag()*, which calculates the number of magnification layers to increase by relating the width of the ZoomBox in pixels, to the current field of view width (768 pixels). This value is added to the existing

magnification layer by *zoombox_zoom()*, which then calls the generic *zoom_function()*.

New image coordinates are calculated using a JavaScript function library, in a similar manner to mag-list navigation. The *pixel_to_coordinate()* is again used to convert onscreen coordinates to image *x,y* coordinates, however with two differences. Firstly, the new magnification layer to be used is calculated based on the width of the ZoomBox as previously described. Secondly, the *x,y* pixel coordinates of the ZoomBox centre replace the mouse cursor coordinates used by mag-list navigation. The *return_box_centre()* function returns the dimensions of the ZoomBox as an array, based on the dimensions of the ZoomBox div element recorded in the *dataform* html form:

BoxC=parseInt(parent.controlpanel.dataform.zoomx.value)+parseInt(parent.co ntrolpanel.dataform.zoomwidth.value)/2;
boxC[1]=parseInt(parent.controlpanel.dataform.zoomy.value)+parseInt(parent. controlpanel.dataform.zoomheight.value)/2;

This data is then incorporated into a URL and requested from the server in the same manner as mag-list navigation. The variable data is intercepted by the browser and transmitted via CHTTP to *populate_db.php* where it is written to the *trackingequalis* table. Additional data is also written to another *zoomboxequalis* that describes the ZoomBox dimensions:

"INSERT INTO ZOOMBOXEQUALIS (SESSNO, EXAMINDEX, TOP, LEFT, WIDTH, HEIGHT) VALUES ('$sessno', '$myvalue', '$zbt', '$zbl', '$zbw', '$zbh')"

Each time a new field of view is requested, the *add_annotate_zoombox_variables()* function is called. If ZoomBox navigation is utilised, ZoomBox variables are appended to the URL, where the *dimensions[]* array are the coordinates of the ZoomBox.

url+="&zbl="+dimensions+"&zbt="+dimensions[1]+"&zbw="+dimensions+" &zb="+dimensions[3];

On submission of the data to *populate_db.php*, data describing ZoomBox dimensions are written to the *zoomboxequalis* table, where *examindex* and *sessno* are used as keys to associate the data to *trackingequalis*, data.

Annotating

Each time a new field of view is requested, the *add_annotate_zoombox_variables()* function is called. If the field has been annotated, the annotation variables are appended to the URL, where the *dimensions[]* array are the coordinates of the annotation box, and *parent.vps.commentsarea.value* is the diagnostic comment, retrieved from the *commentsarea* text area.

url+="&annotated="+annotated+"&anl="+dimensions+"&ant="+dimensions [1]+"&anw="+dimensions+"&anh="+dimensions[3]+"&ancomments="+pare nt.vps.commentsarea.value;

On submission of the data to *populate_db.php*, data describing any marked region and diagnostic comments are written to the *annotateequalis* table, where *examindex* and *sessno* are used as keys to associate the data to *trackingequalis*, *sessionlistequalis* and *reportequalis* data.

ReplaySuite 1.0

Load image filepaths

When a replay is requested, tracking data required to replicate the diagnostic trace is retrieved from the server and returned as a set of PHP variables. These following variables are returned:

- X,Y,Z coordinates
- Time in seconds per field
- Navigation type (lateral/zoom in/zoom out)
- Diagnostic comments

These PHP variables are fed into a series of JavaScript functions:

```
x_array= new Array("<?php
for($line=1; $line<$moves; $line++){
                    echo $session_array[$line][1];
                    if($line<$moves-1)
                    echo("\",\"");
                    }
?>
")
```

The result is a set of JavaScript arrays of filepaths for images viewed during an examination, with associated comments and information. In order to ensure image downloading does not disrupt the replay process, all images are preloaded into the cache. To do this, a 2 dimensional array of image objects is created, and the source of each object is set to an image filepath.

```
an_image[i][c] = new Image();
an_image[i][c].src =  image_path[i][c];
```

ReplaySuite 2.0

ZoomBox display

```
function box(div, y, x, w, h, colour)    {
        parent.controls.div_opaque(div, colour);
        div.style.filter='progid:DXImageTransform.Microsoft.BasicImage(opacity=.1)';

        div.style.visibility="visible";
        div.style.posLeft = x;
        div.style.posTop = y;
        }
```

SnapShots

SnapShots are created using the following JavaScript function:

```
function mkDiv(x, y, w, h){
        this.htm += '<div style="position:absolute;'+
                'left:' + x + 'px;'+
                'top:' + y + 'px;'+
                'width:' + w + 'px;'+
                'height:' + h + 'px;'+
                'clip:rect(0,'+w+'px,'+h+'px,0);'+
                'background-color:' + this.color +
                (!jg_moz? ';overflow:hidden' : '')+
                ';" class=transparent><\/div>';
}
```

Appendix C

Evaluation Surveys

Preliminary Evaluation of the ReplaySuite Survey

Please state how much you agree with the following statements:	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
Q1 The ReplaySuite is user-friendly	○	○	○	○	○
Q2 The ReplaySuite is slow	○	○	○	○	○
Q3 The ReplaySuite displays information clearly	○	○	○	○	○
Q4 The ReplaySuite has an innovative help file	○	○	○	○	○
Where applicable, please state whether you agree or disagree with the following statements about examination replays:					
Q5 The diagnostic pathway followed by the examining pathologist was apparent	○	○	○	○	○
Q6 Being able to pause, fast-forward and rewind was useful	○	○	○	○	○
Q7 The information panel was not useful	○	○	○	○	○
Q8 Images were of good quality	○	○	○	○	○
Rate the following features of the ReplaySuite on importance:	Not at all important	of little importance	Neutral	A little important	Of great importance
Q9 Examinations list	○	○	○	○	○
Q10 Diagnostic concordance graphs	○	○	○	○	○
Q11 Examination replays	○	○	○	○	○
Q12 Summary report forms	○	○	○	○	○
Q13 Individual examination statistical breakdown	○	○	○	○	○

	Yes	No
Q14 Did you encounter technical difficulties whilst using the ReplaySuite?	○	○

Q15 If yes, please state the difficulty that occurred

Where applicable answer yes or no to the following:	Yes	No	N/A
Q16 Did replaying an examination by another pathologist provide you with an insight into why a diagnosis was concluded?	○	○	○
Q17 Did replaying an examination by another pathologist with a diagnosis different to your own provide you with an insight into why that diagnosis was concluded?	○	○	○
Q18 Did replaying examinations performed by other pathologists make you reconsider your diagnosis for any slides?	○	○	○
Q19 If Yes, did your new diagnosis concur with group consensus?	○	○	○

Rate the benefit of using the ReplaySuite in for the following applications where:	Of no benefit	Of little benefit	Neutral	Of some benefit	Of great benefit
Q20 Training	○	○	○	○	○
Q21 Quality Assurance	○	○	○	○	○
Rate the benefit of the following to telepathology:	Not at all important	of little importance	Neutral	A little important	Of great importance
Q22 Digitised virtual slides	○	○	○	○	○
Q23 Downloading virtual slide images from the Internet	○	○	○	○	○
Q24 Replaying your own examinations	○	○	○	○	○
Q25Replaying other pathologists examinations	○	○	○	○	○

Add any further comments, or have any additonal features you would like to see incorporated into the ReplaySuite, please add these below

Post-graduate seminar series exit survey

Q4 How much did you use the online resource?	○	○	○	○	○
Please answer the following in relation to the online availability of virtual slides for cases to be presented in seminars	Not at all beneficial	Not very beneficial	Neutral	Beneficial	Very beneficial
Q5How beneficial would it be to have virtual slides available beforehand?	○	○	○	○	○
	1 day	3 days	1 week	2 weeks	>2 weeks
Q6How long would you like slides be available for, prior to each seminar?	○	○	○	○	○
Q7What other features/resources would you like to see available online?					
During Seminars, in presentation of virtual slides via the projector:	Not at all important	of little importance	Neutral	A little important	Of great importance
Q8 How would you rate image quality?	○	○	○	○	○
	Not at all	Not very	Neutral	Prefered virtual	Greatly prefered virtual
Q9 How much did you prefer the use of virtual slides over glass?.	○	○	○	○	○
How beneficial would you rate the following as an online resource for trainee pathologists:	Not at all beneficial	Not very beneficial	Neutral	Beneficial	Very beneficial
Q10 Virtual slides, including clinical details and diagnosis	○	○	○	○	○
Q11 Slide annotations (images and descriptions of pathological changes)	○	○	○	○	○
Q12 Capability to search for slides by Organ, Diagnosis or Pathological Feature	○	○	○	○	○
Q13 Capability to watch videos of expert examinations, with audio commentary on the examination process and observed pathological changes	○	○	○	○	○
	Yes	No			
Q14 Multiple choice assessments	○	○			
Q15 Direct links from cases to relevant publications	○	○			
Add any further comments, please use the area below					
Submit					

Bibliography

3DHistech Ltd. *Digital Histology Laboratory.*(2002) Available at: http://www.3dhistech.com/ [Accessed 8-8-5].

Abrioux, D. (1989) Computer-Assisted Language Learning and Distance Education. *Journal of Distance Education/ Revue de l'enseignement à distance,* 4, 20-35.

Afdhal, N. H. (2003) Diagnosing fibrosis in hepatitis C: is the pendulum swinging from biopsy to blood tests? *Hepatology,* 37, 972-4.

Anderson, J., Boyle, C., Corbett, A. & Lewis, M. (1990) Cognitive modeling and intelligent tutoring. *Artificial Intelligence,* 42, 7-49.

Anderson, J., Corbett, A., Koedinger, K. & Pelletier, R. (1995) Cognitve tutors: Lessons learned. *The Journal of the Learning Sciences,* 4, 167-207.

Anderson, J. R. (2000) *Cognitive psychology and its implications,* New York, Worth Publishers.

Aperio Technologies. *virtual microscopy solutions for analysis & discovery.*(2005) Available at: www.aperio.com [Accessed 5-6-5].

Apple Computer Inc. *Apple.*(2005) Available at: www.apple.com [Accessed 21-9-5].

Arene, I., Ahmed, W., Fox, M., Barr, C. E. & Fisher, K. (1998) Evaluation of quick medical reference (QMR) as a teaching tool. *M.D. computing: computers in medical practice.*, 15, 323-6.

Ash, J. *Computerized Physician/Provider Order Entry.*(2003) Available at: http://www.ohsu.edu/dmice/research/cpoe/ [Accessed 7-7-5].

Azevedo, R. & Lajoie, S. (1998) The cognitive basis for the design of a mammography interpretation tutor. *International Journal of Artificial Intelligence in Education,* 9, 32-44.

Bajema, I. M., Hagen, E. C., Hansen, B. E., Hermans, J., Noel, L. H., Waldherr, R., Ferrario, F., van der Woude, F. J. & Bruijn, J. A. (1996) The renal histopathology in systemic vasculitis: an international survey study of inter- and intra-observer agreement. *Nephrology Dialysis Transplantation,* 11, 1989-95.

Bamford, W. M., Rogers, N., Kassam, M., Rashbass, J. & Furness, P. N. (2003) The development and evaluation of the UK national telepathology network. *Histopathology,* 42, 110-9.

Barrows, R. C., Jr. & Clayton, P. D. (1996) Privacy, confidentiality, and electronic medical records. *Journal of American Medical Informatics Association,* 3, 139-48.

Bartlett, F. (1932) *Remembering: A study in experimental and social psychology,* Oxford, Cambridge University Press.

Bartlett, F. (1958) *Thinking,* New York, Basic Books.

Batts, K. P. & Ludwig, J. (1995) Chronic hepatitis. An update on terminology and reporting. *American Journal of Surgical Pathology,* 19, 1409-17.

Bender, H. S., Lockee, B. B., Danielson, J. A., Mills, E. M., Boon, G. D., Burton, J. K., Vermeer, P. J., Zimmerman, K. L. & Hilmer, K. M. (2000) Mechanism-based diagnostic reasoning: thoughts on teaching introductory clinical pathology. *Veterinary Clinical Pathology,* 29, 77-83.

Berner, E. S. (Ed.)) (1999) *Clinical Decision Support Systems*, Springer-Verlag New York, Inc.

Bibbo, M., Bartels, P. H., Pfeifer, T., Thompson, D., Minimo, C. & Davidson, H. G. (1993) Belief network for grading prostate lesions. *Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology.,* 15, 124-35.

Bird, K. (1975) *Telemedicine concept and practice In: Telemedicine: Exporations in the Use of Telecommunications in Health Care,* Springfield, IL, Charles C Thomas.

Blake, C. A., Lavoie, H. A. & Millette, C. F. (2003) Teaching medical histology at the University of South Carolina School of Medicine: Transition to virtual slides and virtual microscopes. *Anatomical record. Part B, New anatomist.,* 275, 196-206.

Bloom, B. (1984) The 2 sigma problem: The search for methods of group instruction as effective as one to one tutoring. *Educational Researcher,* 13, 3-16.

Brauchli, K., Christen, H., Haroske, G., Meyer, W., Kunze, K. D. & Oberholzer, M. (2002) Telemicroscopy by the Internet revisited. *Journal of Pathology,* 196, 238-43.

Brauchli, K., O'Mahony, D., Banach, L. & Oberholzer, M. (2005) iPath - a Telemedicine Platform to Support Health Providers in Low Resource Settings. *Studies in health technology and informatics.,* 114, 11-7.

Brauchli, K., Oberli, H., Hurwitz, N., Kunze, K. D., Haroske, G., Jundt, G., Stauch, G., Banach, L., Wirdnam, M., Mihatsch, M. & Oberholzer, M. (2004) Diagnostic telepathology: long-term experience of a single institution. *Virchows Archiv: an international journal of pathology.,* 444, 403-9.

Brunt, E. M. (2000) Grading and staging the histopathological lesions of chronic hepatitis: the Knodell histology activity index and beyond. *Hepatology,* 31, 241-6.

Burthem, J., Brereton, M., Ardern, J., Hickman, L., Seal, L., Serrant, A., Hutchinson, C. V., Wells, E., McTaggart, P., De la Salle, B., Parker-Williams, J. & Hyde, K. (2005) The use of digital 'virtual slides' in the quality assessment of haematological morphology: results of a pilot exercise involving UK NEQAS(H) participants. *British Journal of Haematology,* 130, 293-6.

Byrt, T. (1992) Problems with kappa. *Journal of Clinical Epidemiology,* 45, 1452.

Caballero, T., Perez-Milena, A., Masseroli, M., O'Valle, F., Salmeron, F. J., Del Moral, R. M. & Sanchez-Salgado, G. (2001) Liver fibrosis assessment with semiquantitative indexes and image analysis quantification in sustained-

responder and non-responder interferon-treated patients with chronic hepatitis C. *Journal of Hepatology,* 34, 740-7.

Catalyurek, U., Beynon, M. D., Chang, C., Kurc, T., Sussman, A. & Saltz, J. (2003) The virtual microscope. *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society.,* 7, 230-48.

CDC. *National Center for Infectious Diseases. Viral Hepatitis.*(2005) Available at: http://www.cdc.gov/ncidod/diseases/hepatitis/ [Accessed 8-8-5].

Chang, P. L., Li, Y. C., Wang, T. M., Huang, S. T., Hsieh, M. L. & Tsui, K. H. (1999) Evaluation of a decision-support system for preoperative staging of prostate cancer. *Medical decision making: an international journal of the Society for Medical Decision Making.,* 19, 419-27.

Chi, M., Glaser, R. & Farr, M. (1988) *The Nature of Expertise,* Hillsdale, NJ, Erlbaum.

Clinical Pathology Accreditation (UK) Ltd. *Clinical Pathology Accreditation (UK) Ltd.*(2005) Available at: http://www.cpa-uk.co.uk/ [Accessed 19-8-2005].

Cocker, J., Fox, H. & Langley, F. A. (1968) Consistency in the histological diagnosis of epithelial abnormalities of the cervix uteri. *Journal of Clinical Pathology,* 21, 67-70.

Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20, 37-46.

Coles, C. (1998) How students learn: the process of learning. IN Jolly, B. & Rees, L. (Eds.) *Medical education in the millennium.* Oxford, Oxford University Press.

College of Medicine University of Illinois at Urbana-Champaign. *Atlas of Histology.*(2005) Available at: http://www.med.uiuc.edu/histo/small/atlas/index.htm [Accessed 8-8-5].

Conn, R. B. (1992) Can continuing medical education prepare the current practitioner for the 21st century? *Archives of pathology & laboratory medicine.*, 116, 602-4.

Corbett, A. T. & Anderson, J. R. (1989) Feedback timing and student control in the LISP Intelligent Tutoring System. *Artificial Intelligence and Education*, 64-72.

Costello, S. P. (2004) Development and evaluation of the virtual pathology slide: a new tool for understanding inter-observer variability in diagnostic microscopy. *Biotechnology.* Dublin, Dublin City University.

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. *Journal of Medical Internet Research,* 5, e11.

Cowper, S. E. *Pathmax, A free site designed by a pathologist, for pathology education.*(2005) Available at: http://www.pathmax.com/main.html [Accessed 7-7-5].

Cross, S. S., Dennis, T. & Start, R. D. (2002) Telepathology: current status and future prospects in diagnostic histopathology. *Histopathology,* 41, 91-109.

Cross, S. S., Stephenson, T. J., Mohammed, T. & Harrisont, R. F. (2000) Validation of a decision support system for the cytodiagnosis of fine needle aspirates of the breast using a prospectively collected dataset from multiple observers in a working clinical environment. *Cytopathology,* 11, 503-12.

Crowley, R., Medvedeva, O., Jukic, D. &, A., 2003 (2003) SlideTutor – A model-tracing Intelligent Tutoring System for teaching microscopic diagnosis. *Proceedings of the 11th International Conference on Artificial Intelligence in Education.* Sydney, Australia, IOS Press.

Crowley, R. S. & Medvedeva, O. (2003) A general architecture for intelligent tutoring of diagnostic classification problem solving. *AMIA. Annual Symposium proceedings [electronic resource] / AMIA Symposium. AMIA Symposium.*, 185-9.

Crowley, R. S., Naus, G. J., Stewart, J., 3rd & Friedman, C. P. (2003) Development of visual diagnostic expertise in pathology -- an information-processing study. *Journal of the American Medical Informatics Association: JAMIA.*, 10, 39-51.

Cuthbert, L., duBoulay, B., Teather, D., Teather, B., Sharples, M. & duBoulay, G. *Expert/Novice differences in Diagnostic Medical Cognition - A review of the Literature.*(1999) Available at: http://www.cogs.susx.ac.uk/users/bend/papers/csrp508.pdf [Accessed 7-7-5].

Dahab, G. M., Kheriza, M. M., El-Beltagi, H. M., Fouda, A. M. & El-Din, O. A. (2004) Digital quantification of fibrosis in liver biopsy sections: description of a

new method by Photoshop software. *Journal of gastroenterology and hepatology.*, 19, 78-85.

Darlington, K. (1996) Basic Expert Systems. *Information Technology in Nursing,* 8, 9-11.

David, D., Miclea, M. & Opre, A. (2004) The information-processing approach to the human mind: Basics and beyond. *Journal of Clinical Psychology,* 60, 353-68.

Dawson, M. & Medler, D. *University of Alberta's Dictionary of Cognitive Science.*(2004) Available at: http://www.bcp.psych.ualberta.ca/~mike/Pearl_Street/Dictionary/dictionary.html [Accessed 8-8-5].

Dee, F. & Leaven, T. *the Virtual Slidebox, virtual microscopy for education.*(2005) Available at: http://www.path.uiowa.edu/virtualslidebox/ [Accessed 5-6-5].

Dee, F. R., Lehman, J. M., Consoer, D., Leaven, T. & Cohen, M. B. (2003) Implementation of virtual microscope slides in the annual pathobiology of cancer workshop laboratory. *Human Pathology,* 34, 430-6.

Degroot, A. (1965) *Thought and Choice in Chess,* The Hague, The Netherlands, Mouton.

Della Mea, V., Cataldi, P., Boi, S., Finato, N., Della Palma, P. & Beltrami, C. A. (1998) Image selection in static telepathology through the Internet. *Journal of Telemedicine and Telecare,* 4 Suppl 1, 20-2.

Della Mea, V., Cortolezzis, D. & Beltrami, C. A. (2000) The economics of telepathology - a case study. *Journal of Telemedicine and Telecare,* 6, S168-169.

Delpierre, C., Cuzin, L., Fillaux, J., Alvarez, M., Massip, P. & Lang, T. (2004) A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach? *International journal for quality in health care: journal of the International Society for Quality in Health Care / ISQua.,* 16, 407-16.

Demichelis, F., Barbareschi, M., Dalla Palma, P. & Forti, S. (2002) The virtual case: a new method to completely digitize cytological and histological slides. *Virchows Archiv: an international journal of pathology.,* 441, 159-64.

Demichelis, F., Barbareschi, M., Dalla Palma, P. & Forti, S. (2002) The virtual case: a new method to completely digitize cytological and histological slides. *Virchows Arch,* 441, 159-64.

Dennis, T., Start, R. D. & Cross, S. S. (2005) The use of digital imaging, video conferencing, and telepathology in histopathology: a national survey. *Journal of Clinical Pathology,* 58, 254-8.

Dervan, P. (2005)

Descour, M., Kärkkäinen, A., Rogers, J., Liang, C., Weinstein, R., Rantala, J., Kilic, B., Madenci, E., Richards-Kortum, R., Anslyn, E., Dupuis, R., Schul, R., Willison, C. & Tigges, C. (2002) Toward the Development of Miniaturized Imaging Systems for Detection of Pre-Cancer. *IEEE Journal of Quantum Electronics,* 38, 122-130.

Diamond, J., Anderson, N. H., Thompson, D., Bartels, P. H. & Hamilton, P. W. (2002) A computer-based training system for breast fine needle aspiration cytology. *Journal of Pathology,* 196, 113-21.

Dmetrix Inc. *Dmetrix. Microscopy multiplied.*(2005) Available at: http://www.dmetrix.net/ [Accessed 19-9-2005].

Doolittle, M. H., Doolittle, K. W., Winkelman, Z. & Weinberg, D. S. (1997) Color images in telepathology: how many colors do we need? *Human Pathology,* 28, 36-41.

Du Boulay, C. (1997) Continuing medical education for pathologists: an evaluation of the Royal College of Pathologists' Wessex pilot scheme. *Journal of Clinical Pathology,* 50, 1022-6.

Ellis, I., Wells, C. & Sowter, C. *St Bartholomew's Histopathology department: EQA Histopathology.*(2005) Available at: http://www.telepathology.qmul.ac.uk/histopath/breast/histology/uk/breqahis.htm [Accessed 20-9-5].

EQUALIS. *External Quality Assurance in Laboratory Medicine in Sweden.*(2005) Available at: www.equalis.se [Accessed 12-12-4].

Eshach, H. & Bitterman, H. (2003) From case-based reasoning to problem-based learning. *Acad Med,* 78, 491-6.

European Working Group for Breast Screening Pathology. *European Working Group for Breast Screening Pathology European Quality Assurance Page.*(2005)

Available at: http://www.telepathology.qmul.ac.uk/euroqa/euroqa1/start.htm [Accessed 10-8-5].

Evans, D. & Patel, V. L. (Eds.) (1989) *Cognitive Science in Medicine,* Cambridge, MA, MIT Press.

Feit, J., Jedličková, H., Vlašín, Z., Burg, G., Kempf, W. & Matyska, L. *Atlas of Dermatology.*(2005) Available at: http://atlases.muni.cz [Accessed 7-7-5].

Firestone, A. R., Sema, D., Heaven, T. J. & Weems, R. A. (1998) The effect of a knowledge-based, image analysis and clinical decision support system on observer performance in the diagnosis of approximal caries from radiographic images. *Caries research.,* 32, 127-34.

Fleege, J. C., van Diest, P. J. & Baak, J. P. (1992) Quality control methods for data entry in pathology using a computerized data management system based on an extended data dictionary. *Human Pathology,* 23, 91-7.

Flynn, B. & Davidson, M. *Nikon MicroscopyU. CCD Resolution for Optical Microscopy.*(2005) Available at: http://www.microscopyu.com/tutorials/flash/pixelcalc/ [Accessed 7-8-5].

Foran, D. J., Meer, P. P., Papathomas, T. & Marsic, I. (1997) Compression guidelines for diagnostic telepathology. *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society.,* 1, 55-60.

Freedman, R., Cho, B., Glass, M., Zhou, Y., Kim, J., Mills, B., Yang, F. & Evens, M. (2001) Adaptive Processing in a Medical Intelligent Tutoring System. *Workshop on Adaptation in Dialogue Systems, NAACL.* Pittsburgh.

Fujita, K. & Crowley, R. S. (2003) The Virtual Slide Set - a curriculum development system for digital microscopy. *AMIA Annu Symp Proc*, 846.

General Medical Council (1993) Tomorrows Doctors. Recommendations on undergraduate medical education. London, GMC.

Glatz-Krieger, K., Glatz, D. & Mihatsch, M. J. (2003) Virtual slides: high-quality demand, physical limitations, and affordability. *Human Pathology,* 34, 968-74.

Glatz-Krieger, K., Glatz, D. & Mihatsch, M. J. (2005) [Virtual microscopy: first applications.]. *Pathologe.*

Goldin, R. D., Goldin, J. G., Burt, A. D., Dhillon, P. A., Hubscher, S., Wyatt, J. & Patel, N. (1996) Intra-observer and inter-observer variation in the histopathological assessment of chronic viral hepatitis. *Journal of Hepatology,* 25, 649-54.

Gronbaek, K., Christensen, P. B., Hamilton-Dutoit, S., Federspiel, B. H., Hage, E., Jensen, O. J. & Vyberg, M. (2002) Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. *Journal of Viral Hepatology,* 9, 443-9.

Halliday, B. E., Bhattacharyya, A. K., Graham, A. R., Davis, J. R., Leavitt, S. A., Nagle, R. B., McLaughlin, W. J., Rivas, R. A., Martinez, R., Krupinski, E. A. &

Weinstein, R. S. (1997) Diagnostic accuracy of an international static-imaging telepathology consultation service. *Human Pathology,* 28, 17-21.

Hamilton, P. W., Anderson, N. H., Diamond, J., Bartels, P. H., Gregg, J. B., Thompson, D. & Millar, R. J. (1996) An interactive decision support system for breast fine needle aspiration cytology. *Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology.,* 18, 185-90.

Hamilton, P. W., Montironi, R., Abmayr, W., Bibbo, M., Anderson, N., Thompson, D. & Bartels, P. H. (1995) Clinical applications of Bayesian belief networks in pathology. *Pathologica,* 87, 237-45.

Hamza, S., Anderson, P., Reddy, V. V. & Siegal, G. P. (2001) Use of the Internet in pathology resident training and education. *Advances in anatomic pathology.,* 8, 290-7.

Hanby, A. M., Hall, P. A., Rooney, N., Dennis, P., James, P., Richman, P., Buk, S., Levison, D. A. & Gregory, W. M. (1992) An inter-observer and intra-observer variability study on the diagnosis of lymph node biopsy specimens. *European journal of cancer.,* 28A, 1858-62.

Harmon, B. J., Wah, R. & Inae, T. (2003) The Military Health System Computer-based Patient Record. *AMIA. Annual Symposium proceedings [electronic resource] / AMIA Symposium. AMIA Symposium.*, 1068.

Harris, T., Leaven, T., Heidger, P., Kreiter, C., Duncan, J. & Dick, F. (2001) Comparison of a virtual microscope laboratory to a regular microscope laboratory for teaching histology. *The Anatomical record.,* 265, 10-4.

Hastrup, N., Clemmensen, O. J., Spaun, E. & Sondergaard, K. (1994) Dysplastic naevus: histological criteria and their inter-observer reproducibility. *Histopathology,* 24, 503-9.

Haug, P. J., Gardner, R. M., Tate, K. E., Evans, R. S., East, T. D., Kuperman, G., Pryor, T. A., Huff, S. M. & Warner, H. R. (1994) Decision support in medicine: examples from the HELP system. *Computers and biomedical research, an international journal.,* 27, 396-418.

Heidger, P. M., Jr., Dee, F., Consoer, D., Leaven, T., Duncan, J. & Kreiter, C. (2002) Integrated approach to teaching and testing in histology with real and virtual imaging. *The Anatomical recorder,* 269, 107-12.

Heidger, P. M., Jr., Dee, F., Consoer, D., Leaven, T., Duncan, J. & Kreiter, C. (2002) Integrated approach to teaching and testing in histology with real and virtual imaging. *Anat Rec,* 269, 107-12.

Helin, H., Lundin, M., Lundin, J., Martikainen, P., Tammela, T., Helin, H., van der Kwast, T. & Isola, J. (2005) Web-based virtual microscopy in teaching and standardizing Gleason grading. *Human Pathology,* 36, 381-6.

Hoehler, F. K. (2000) Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol,* 53, 499-503.

Howe, W. *A brief history of the Internet.*(2004) Available at: http://www.walthowe.com/navnet/history.html [Accessed 8-8-5].

i-Path Diagnostics. *i-Path Diagnostics. Objectivity. Standardisation. Control.*(2004) Available at: http://www.i-path.co.uk/ [Accessed 7-7-5].

Ishak, K., Baptista, A., Bianchi, L., Callea, F., De Groote, J., Gudat, F., Denk, H., Desmet, V., Korb, G. & MacSween, R. N. (1995) Histological grading and staging of chronic hepatitis. *Journal of Hepatology,* 22, 696-9.

Jamieson, S. (2004) Likert scales: how to (ab)use them. *Medical Education.,* 38, 1217-8.

Jeffers, M. (2005) Department of Pathology, Adelaide and Meath Hospital, Dublin

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 medical informatics and decision making,* 5, 10.

Jones, K. N., Kreisle, R., Geiss, R. W., Holliman, J. H., Lill, P. H. & Anderson, P. G. (2002) Group for research in pathology education online resources to facilitate pathology instruction. *Arch Pathol Lab Med,* 126, 346-50.

Jones, K. N., Kreisle, R., Geiss, R. W., Holliman, J. H., Lill, P. H. & Anderson, P. G. (2002) Group for research in pathology education online resources to facilitate pathology instruction. *Archives of pathology & laboratory medicine,* 126, 346-50.

Kaplan, K. J., Burgess, J. R., Sandberg, G. D., Myers, C. P., Bigott, T. R. & Greenspan, R. B. (2002) Use of robotic telepathology for frozen-section diagnosis: a retrospective trial of a telepathology system for intraoperative consultation. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc.,* 15, 1197-204.

Kassirer, J. P. (1989) Diagnostic reasoning. *Annals of internal medicine.,* 110, 893-900.

Kawamoto, K., Houlihan, C. A., Balas, E. A. & Lobach, D. F. (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *British Medical Journal,* 330, 765.

Kayser, K., Kayser, G., Radziszowski, D. & Oehmann, A. (1999) From telepathology to virtual pathology institution: the new world of digital pathology. *Revue roumaine de morphologie et embryologie.,* 45, 3-9.

Kayser, K. & Schlegel, W. (1982) Pattern recognition in histo-pathology: basic considerations. *Methods of information in medicine.,* 21, 15-22.

Kent, T. H. (1977) The Group for Research in Pathology Education. *Arch Pathol Lab Med,* 101, 279.

Khuwaja, R. & Patel, V. (1996) A Model of Tutoring: Based on the Behavior of Effective Human Tutors. *Intelligent Tutoring Systems*, 130-136.

Klatt, E. C. (1997) Web-based teaching in pathology. *JAMA: the journal of the American Medical Association.,* 278, 1787.

Klatt, E. C. & Dennis, S. E. (1998) Web-based pathology education. *Arch Pathol Lab Med,* 122, 475-9.

Knodell, R. G., Ishak, K. G., Black, W. C., Chen, T. S., Craig, R., Kaplowitz, N., Kiernan, T. W. & Wollman, J. (1981) Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. *Hepatology,* 1, 431-5.

Koedinger, K. & Anderson, J. R. (1997) Intelligent Tutoring Goes To School in the Big City. *Internaltional Journal of Artificial Intelligence in Education,* 8, 30-43.

Krieger, N., Hiatt, R. A., Sagebiel, R. W., Clark, W. H., Jr. & Mihm, M. C., Jr. (1994) Inter-observer variability among pathologists' evaluation of malignant melanoma: effects upon an analytic study. *Journal of Clinical Epidemiology,* 47, 897-902.

Kristula, D. *The History of the Internet.*(2001) Available at: http://www.davesite.com/webstation/net-history.shtml [Accessed 8-8-5].

Kronz, J., Silberman, M. & Epstein, J. *The WHO/ISUP Consensus Classification of Urothelial (Transitional Cell) Neoplasms tutorial and proficiency test.*(2004) Available at: http://162.129.103.34/bladder [Accessed 10-8-5].

Kronz, J. D., Silberman, M. A., Allsbrook, W. C., Jr., Bastacky, S. I., Burks, R. T., Cina, S. J., Mills, S. E., Ross, J. S., Sakr, W. A., Tomaszewski, J. E., True, L. D., Ulbright, T. M., Weinstein, M. W., Yantiss, R. K., Young, R. H. & Epstein, J. I. (2000) Pathology residents' use of a Web-based tutorial to improve Gleason grading of prostate carcinoma on needle biopsies. *Human Pathology,* 31, 1044-50.

Kronz, J. D., Silberman, M. A., Allsbrook, W. C., Jr., Bastacky, S. I., Burks, R. T., Cina, S. J., Mills, S. E., Ross, J. S., Sakr, W. A., Tomaszewski, J. E., True, L.

D., Ulbright, T. M., Weinstein, M. W., Yantiss, R. K., Young, R. H. & Epstein, J. I. (2000) Pathology residents' use of a Web-based tutorial to improve Gleason grading of prostate carcinoma on needle biopsies. *Hum Pathol,* 31, 1044-50.

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

Langer, I., Schewe, S., Haedecke, C., Puppe, F. & Rheinhardt, T. (1998) Learning at the computer: evaluation of an intelligent tutoring system. *Eur J Med Res,* 3, 119-26.

Langer, I., Schewe, S., Haedecke, C., Puppe, F. & Rheinhardt, T. (1998) Learning at the computer: evaluation of an intelligent tutoring system. *European journal of medical research.,* 3, 119-26.

Lee, E. S., Kim, I. S., Choi, J. S., Yeom, B. W., Kim, H. K., Han, J. H., Lee, M. S. & Leong, A. S. (2003) Accuracy and reproducibility of telecytology diagnosis of cervical smears. A tool for quality assurance programs. *American Journal of Clinical Pathology,* 119, 356-60.

Legg, E. F. & Hurrell, A. E. (1984) External quality assessment of quantitative urinary analysis. *Annals of clinical biochemistry.,* 21 (Pt 6), 491-3.

Lehmann, C. & Cohen, B. (2002) Dermatlas:An Online Collaborative Education Tool. *The Internet Journal of Dermatology,* 1.

Lehmann, H. P., Lehmann, C. U. & Freedman, J. A. (1997) The use of simulations in computer-aided learning over the World Wide Web. *JAMA: the journal of the American Medical Association.,* 278, 1788.

Lemaire, J. B., Schaefer, J. P., Martin, L. A., Faris, P., Ainslie, M. D. & Hull, R. D. (1999) Effectiveness of the Quick Medical Reference as a diagnostic tool. *journal de l'Association medicale canadienne.,* 161, 725-8.

Leong, F.-M. (2002) Diagnostic Quantitative Pathology. *8th Congress of the European Society for Analytical Cellular Pathology, 6th Congress of European Group of Telepathology.* Heraclion, Crete, Greece.

Leong, F. J. & McGee, J. O. (2001) Automated complete slide digitization: a medium for simultaneous viewing by multiple pathologists. *Journal of Pathology,* 195, 508-14.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., LKlopfer, D. & Wang, Y. (1988) *Expertise in a complex skill: diagnosing x-ray pictures. In: The Nature of Expertise,* Hillsdale, NJ, Lawrence Erlbaum Associates.

Likert, R. (1932) *A Technique for the Measurement of Attitudes,* New York, McGraw-Hill.

Lincoln, M. J., Turner, C. W., Haug, P. J., Warner, H. R., Williamson, J. W., Bouhaddou, O., Jessen, S. G., Sorenson, D., Cundick, R. C. & Grant, M. (1991) Iliad training enhances medical students' diagnostic skills. *Journal of medical systems.,* 15, 93-110.

Loyola University Chicago. *Liver Cirrhosis.*(2005) Available at: http://www.meddean.luc.edu/lumen/MedEd/orfpath/cirhosis.htm [Accessed 2005].

Lundin, M., Lundin, J., Helin, H. & Isola, J. (2004) A digital atlas of breast histopathology: an application of web based virtual microscopy. *J Clin Pathol,* 57, 1288-91.

Lundin, M., Lundin, J., Helin, H. & Isola, J. (2004) A digital atlas of breast histopathology: an application of web based virtual microscopy. *Journal of Clinical Pathology,* 57, 1288-91.

Lundin, M., Lundin, J. & Isola, J. (2004) Virtual microscopy. *Journal of Clinical Pathology,* 57, 1250-1.

Macromedia Inc. *Macromedia.*(2005) Available at: http://www.macromedia.com/ [Accessed 20-8-5].

Mairinger, T., Netzer, T. T., Schoner, W. & Gschwendtner, A. (1998) Pathologists' attitudes to implementing telepathology. *J Telemed Telecare,* 4, 41-6.

Martens, A., Bernauer, J., Illmann, T. & Seitz, A. (2001) "Docs 'n drugs--the virtual polyclinic": an intelligent tutoring system for web-based and case-oriented training in medicine. *Proc AMIA Symp*, 433-7.

Martens, A., Bernauer, J., Illmann, T. & Seitz, A. (2001) "Docs 'n drugs--the virtual polyclinic": an intelligent tutoring system for web-based and case-oriented training in medicine. *Proceedings / AMIA. Annual Symposium. AMIA Symposium.*, 433-7.

Media Cybernetics. *Media Cybernetics. From images to answers.*(2005) Available at: http://www.mediacy.com/ [Accessed 12-12-4].

Medical Solutions PLC. *PathScope.*(2005) Available at: http://www.pathscope-eqa.co.uk/ [Accessed 12-12-4].

Melton, J. & Swanson, J. *Loyola University Dermatology Medical Education Website.*(1997) Available at: http://www.lumen.luc.edu/lumen/MedEd/medicine/dermatology/melton/title.htm [Accessed 7-7-5].

Metavir group (1994) Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. *Hepatology,* 20, 15-20.

Microsoft corporation. *visual filters.*(2004) Available at: http://msdn.microsoft.com/library/default.asp?url=/workshop/author/filter/filters.asp [Accessed 12-12-4].

Miller, G. A. (1956) The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological review.,* 63, 81-97.

Miller, G. A. (1994) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review.,* 101, 343-52.

Miller, R., Masarie, F. E. & Myers, J. D. (1986) Quick medical reference (QMR) for diagnostic assistance. *M.D. computing: computers in medical practice.,* 3, 34-48.

Mitros, F. (1996) The virtual hospital atlas of liver pathology.

Montironi, R., Whimster, W. F., Collan, Y., Hamilton, P. W., Thompson, D. & Bartels, P. H. (1996) How to develop and use a Bayesian Belief Network. *J Clin Pathol,* 49, 194-201.

Morrison, M. L., McCluggage, W. G., Price, G. J., Diamond, J., Sheeran, M. R., Mulholland, K. M., Walsh, M. Y., Montironi, R., Bartels, P. H., Thompson, D. & Hamilton, P. W. (2002) Expert system support using a Bayesian belief network for the classification of endometrial hyperplasia. *Journal of Pathology,* 197, 403-14.

Mullick, F. G., Fontelo, P. & Pemble, C. (1996) Telemedicine and telepathology at the Armed Forces Institute of Pathology: history and current mission. *Telemedicine Journal,* 2, 187-93.

Nayar, R. & Solomon, D. (2004) Second edition of 'The Bethesda System for reporting cervical cytology' - atlas, website, and Bethesda interobserver reproducibility project. *Cytojournal,* 1, 4.

Nayar, R. & Solomon, D. *Bethesda System Website Atlas.*(2005) Available at: http://www.cytopathology.org/NIH/ [Accessed 11-7-5].

NHS Cancer Screening Programmes. *Non-operative diagnosis Subgroup of the National Coordinating Group for Breast Screening Pathology. Guidelines for non-operative diagnostic prcedures and reporting in breast cancer screening.*(2001) Available at: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp50.pdf [Accessed 17-8-5].

Nieder, G. L. & Nagy, F. (2002) Analysis of medical students' use of web-based resources for a gross anatomy and embryology course. *Clinical Anatomy,* 15, 409-18.

Nieder, G. L., Scott, J. N. & Anderson, M. D. (2000) Using QuickTime virtual reality objects in computer-assisted instruction of gross anatomy: Yorick--the VR Skull. *Clinical Anatomy,* 13, 287-93.

Nordrum, I., Engum, B., Rinde, E., Finseth, A., Ericsson, H., Kearney, M., Stalsberg, H. & Eide, T. J. (1991) Remote frozen section service: a telepathology project in northern Norway. *Human Pathology,* 22, 514-8.

Nordrum, I., Johansen, M., Amin, A., Isaksen, V. & Ludvigsen, J. A. (2004) Diagnostic accuracy of second-opinion diagnoses based on still images. *Human Pathology,* 35, 129-35.

Norman, D. (1982) *Learning and Memory,* San Francisco, WH Freeman.

Norman, G. R. & Schmidt, H. G. (2000) Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Medical Education.,* 34, 721-8.

O'Briain, S. (2005) Department of Pathology, St. James hospital, Dublin

O'Brien, M. J., Keating, N. M., Elderiny, S., Cerda, S., Keaveny, A. P., Afdhal, N. H. & Nunes, D. P. (2000) An assessment of digital image analysis to measure fibrosis in liver biopsy specimens of patients with chronic hepatitis C. *Am J Clin Pathol,* 114, 712-8.

O'Shea, D. G. (2005) School of Biotechnology, Dublin City University

OmniMD. *Electronic Medical Record.*(2004) Available at: http://www.omnimd.com/ [Accessed 8-8-5].

Ong, J. & Ramachandran, S. *Intelligent Tutoring Systems: The What and the How.*(2000) Available at: http://www.learningcircuits.org/2000/feb2000/ong.htm [Accessed 7-6-5].

Oxford Centre for Staff Development (1992) Report of the Council for National Academic Awards Improving Student
Learning Project. Improving student learning. Oxford, Oxford Centre for Staff Development.

Patel, V. L., Arocha, J. F. & Kaufman, D. R. (2001) A primer on aspects of cognition for medical informatics. *Journal of the American Medical Informatics Association,* 8, 324-43.

Petersen, I., Wolf, G., Roth, K. & Schluns, K. (2000) Telepathology by the Internet. *J Pathol,* 191, 8-14.

Psybox. *Learn the intelligent way.Psybox.*(2003) Available at: http://www.psybox.com/web_dictionary/schema1.htm [Accessed 7-8-5].

Rakovitch, E., Mihai, A., Pignol, J. P., Hanna, W., Kwinter, J., Chartier, C., Ackerman, I., Kim, J., Pritchard, K. & Paszat, L. (2004) Is expert breast pathology assessment necessary for the management of ductal carcinoma in situ? *Breast Cancer Research and Treatment,* 87, 265-72.

Ramsay, A. D. (1999) Errors in histopathology reporting: detection and avoidance. *Histopathology,* 34, 481-90.

Reid, W. A., Harvey, J., Watson, G. R., Luqmani, R., Harkin, P. J. & Arends, M. J. (2000) Medical student appraisal of interactive computer-assisted learning programs embedded in a general pathology course. *Journal of Pathology,* 191, 462-5.

Riggs, R., Purtillo, D. & Connor, D. (1974) Medical consultation via telecommunications. *JAMA,* 228, 600-602.

Rind, D. M., Kohane, I. S., Szolovits, P., Safran, C., Chueh, H. C. & Barnett, G. O. (1997) Maintaining the confidentiality of medical records shared over the Internet and the World Wide Web. *Ann Intern Med,* 127, 138-41.

Rous, B. *Pathnet. histopathology training school.*(2005) Available at: http://www.pathnet.org.uk/moodle/ [Accessed 20-9-5].

Rous, B., Skilbeck, A., Danson, S., Bailey, W., Thottempudi, M. & Ji, W. *TAPIR. the teaching archive of pathology images resources.*(2005) Available at: http://tapir.caret.cam.ac.uk/tapir/jsp/menu.jsp [Accessed 20-9-5].

Rousselet, M. C., Michalak, S., Dupre, F., Croue, A., Bedossa, P., Saint-Andre, J. P. & Cales, P. (2005) Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology,* 41, 257-64.

Royal College of Physicians of Ireland. *Royal College of Physicians of Ireland.*(2004) Available at: www.rcpi.ie [Accessed 14-7-5].

Russell, W. O. (1966) Continuing education--a challenge for pathologists and medical technologists. *The American journal of medical technology.,* 32, 57-9.

Sacile, R., Montaldo, E., Ruggiero, C., Nieburgs, H. E. & Nicolo, G. (2003) A decision support system to detect morphologic changes of chromatin arrangement in normal-appearing cells. *IEEE transactions on nanobioscience.*, **2**, 118-23.

Saeger, K., Schlüns, K., Scharder, T. & Hufnagl, P. (2003) The virtual miscroscope for routine pathology based on a PACS system for 6 Gb images. *International Congress Series,* 1256, 299-304.

Scheuer, P. J. (1991) Classification of chronic viral hepatitis: a need for reassessment. *Journal of Hepatology,* 13, 372-4.

Scheuer, P. J. (2003) Liver biopsy size matters in chronic hepatitis: bigger is better. *Hepatology,* 38, 1356-8.

Schmidt, H. G. (1983) Problem-based learning: rationale and description. *Med Educ,* 17, 11-6.

Schmidt, H. G., Norman, G. R. & Boshuizen, H. P. (1990) A cognitive perspective on medical expertise: theory and implication. *Academic medicine: journal of the Association of American Medical Colleges,* 65, 611-21.

Schoonjans, F. *MedCalc.*(2005) Available at: http://www.medcalc.be/download.php [Accessed].

Shek, T. W. (1996) Bayesian Belief Network in histopathology. *Journal of Clinical Pathology,* 49, 864.

Shortliffe, E. & Perreault, L. (Eds.) (2001) *Medical Informatics. Computer Applications in Health Care and Biomedicine*, Springer.

Shotts, W. *Cjpeg.*(1998) Available at: http://linuxcommand.org/man_pages/cjpeg1.html [Accessed 7-6-5].

Sim, J. & Wright, C. C. (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther,* 85, 257-68.

Singh, N., Akbar, N., Sowter, C., Lea, K. G. & Wells, C. A. (2002) Telepathology in a routine clinical environment: implementation and accuracy of diagnosis by robotic microscopy in a one-stop breast clinic. *Journal of Pathology,* 196, 351-5.

SlidePath Ltd. *Digital SlideBox.*(2004) Available at: www.slidepath.com [Accessed 4-4-5].

Smith, A. E., Sherman, M. E., Scott, D. R., Tabbara, S. O., Dworkin, L., Olson, J., Thompson, J., Faser, C., Snell, J. & Schiffman, M. (2000) Review of the Bethesda System atlas does not improve reproducibility or accuracy in the classification of atypical squamous cells of undetermined significance smears. *Cancer,* 90, 201-6.

Snell, J. J. & Brown, D. F. (2001) External quality assessment of antimicrobial susceptibility testing in Europe. *The Journal of antimicrobial chemotherapy,* 47, 801-10.

Spencer, J. A. & Jordan, R. K. (1999) Learner centred approaches in medical education. *British Medical Journal,* 318, 1280-3.

Spring, K. & Davidson, M. *Basic Concepts and Formulas in Microscopy.*(2005) Available at: http://www.microscopyu.com/articles/formulas/formulasindex.html [Accessed 7-7-5].

Steinberg, D. M. & Ali, S. Z. (2001) Application of virtual microscopy in clinical cytopathology. *Diagnostic Cytopathology,* 25, 389-96.

Suebnukarn, S. & Haddawy, P. (2004) A Collaborative Intelligent Tutoring System for Medical Problem-Based Learning. *Procedings of the International Conference on Intelligent User Interfaces.* Madeira, Portugal.

Swedish Society of Pathology. *Svensk Förening för Patologi.*(2004) Available at: http://www.svls.se/sektioner/pa/index.html [Accessed 8-8-4].

Szolovits, P. & Pauker, S. G. (1978) Categorical and Probabilistic Reasoning in Medical Diagnosis. *Artificial Intelligence,* 11, 115-144.

Tang, Z. Y., Savino, A., Wong, E. K., Koss, L. G. & Shaw, L. G. (1988) An expert system designed as a tutoring tool in cervical cytology: TTCC-1 system. *Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology,* 10, 417-22.

Teasdale, S. (1996) The future of clinical audit: learning to work together. *British Medical Journal,* 313, 574.

The Irish Committee on Higher Medical Training. *Curriculum for Higher Specialist Training in Histopathology.*(2004) Available at:

http://www.rcpi.ie/CONTENT/N000000164/C000_R001_DLD/ICHMT_Histopathology_200105.pdf [Accessed 14-7-5].

The Royal College of Pathologists. *RCPath.*(2005) Available at: http://www.rcpath.org [Accessed 8-7-5].

The Royal College of Pathologists of Australasia. *Pathology - Discipline by Discipline.*(2004) Available at: http://www.rcpa.edu.au/public/pathology/discipline.cfm#anatomical [Accessed 6-6-5].

Tiersma, E. S., Peters, A. A., Mooij, H. A. & Fleuren, G. J. (2003) Visualising scanning patterns of pathologists in the grading of cervical intraepithelial neoplasia. *Journal of Clinical Pathology,* 56, 677-80.

Tranberg, H. A., Rous, B. A. & Rashbass, J. (2003) Legal and ethical issues in the use of anonymous images in pathology teaching and research. *Histopathology,* 42, 104-9.

Trestle Corporation. *Trestle corporation. Connect. Collaborate. Communicate.*(2005) Available at: http://www.trestlecorp.com/home.asp [Accessed 6-7-5].

Turner, C. W., Lincoln, M. J., Haug, P., Williamson, J. W., Jessen, S., Cundick, K. & Warner, H. (1991) Iliad training effects: a cognitive model and empirical findings. *Proceedings of the Annual Symposium on Computer Application in Medical Care.*, 68-72.

Tversky, A., and Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science,* 185, 1124-1131.

UKAS. *United Kingdom Accreditation Service.*(2005) Available at: http://www.ukas.com [Accessed 21-8-5].

UKNEQAS. *United Kingdom National External Quality Assessment Service.*(2001) Available at: http://www.ukneqas.org.uk/ [Accessed 7-8-5].

University of Alabama at Birmingham. *IPLAB.Net, The online interactive pathology laboratory.*(2005) Available at: http://iplab.net/ [Accessed 13-7-5].

University of Florida College of Medicine. *Online Testing Center.*(2000) Available at: http://testing.medinfo.ufl.edu/ [Accessed 13-7-5].

University of Florida College of Medicine. *Extensible Assessment Machine (XAM).*(2003) Available at: http://evalsuite.medinfo.ufl.edu/docs/xam/ [Accessed 13-7-5].

University of Kansas Medical Center (1996) Basic Histopathology website.

University of Leeds. *Liver EQA 2005 circulation Q.*(2005) Available at: http://www.virtualpathology.leeds.ac.uk/eqa/liver.php [Accessed 8-8-5].

University of Udine. *eSlides.*(2004) Available at: http://www.telemed.uniud.it/eslides/ [Accessed 7-6-5].

Vincent, C., Neale, G. & Woloshynowych, M. (2001) Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal,* 322, 517-9.

Viosport. *Glossary.*(2004) Available at: http://www.viosport.com/support/index.php?page=glossary [Accessed 7-8-5].

Wang, X. H., Zheng, B., Good, W. F., King, J. L. & Chang, Y. H. (1999) Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network. *International journal of medical informatics,* 54, 115-26.

WebSideStory. *WebSideStory. Active Insight. Active Marketing.*(2004) Available at: http://www.websidestory.com [Accessed 12-12-4].

Weinberg, D. S., Allaert, F. A., Dusserre, P., Drouot, F., Retailliau, B., Welch, W. R., Longtine, J., Brodsky, G., Folkerth, R. & Doolittle, M. (1996) Telepathology diagnosis by means of digital still images: an international validation study. *Human Pathology,* 27, 111-8.

Weinstein, L. J., Epstein, J. I., Edlow, D. & Westra, W. H. (1997) Static image analysis of skin specimens: the application of telepathology to frozen section evaluation. *Human Pathology,* 28, 30-5.

Weinstein, R. S. (1986) Prospects for telepathology. *Human Pathology,* 17, 433-4.

Weinstein, R. S., Bloom, K. J. & Rozek, L. S. (1989) Telepathology. Long-distance diagnosis. *American Journal of Clinical Pathology,* 91, S39-42.

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. *Human Pathology,* 35, 1303-14.

Weinstein, R. S., Descour, M. R., Liang, C., Bhattacharyya, A. K., Graham, A. R., Davis, J. R., Scott, K. M., Richter, L., Krupinski, E. A., Szymus, J., Kayser, K. & Dunn, B. E. (2001) Telepathology overview: from concept to implementation. *Human Pathology,* 32, 1283-99.

Wejstal, R., Alaeus, A., Fischler, B., Reichard, O., Uhnoo, I. & Weiland, O. (2003) Chronic hepatitis C: updated Swedish consensus. *Scand J Infect Dis,* 35, 445-51.

Wejstal, R., Fischler, B., Glaumann, H., Norkrans, G., Reichard, O., Sonnerbor, A., Uhnoo, I. & Weiland, O. (2000) Chronic hepatitis C--Swedish experts' meeting recommends combination treatment. *Scand J Infect Dis,* 32, 465-70.

Westin, J., Lindh, M., Lagging, L. M., Norkrans, G. & Wejstal, R. (1999) Chronic hepatitis C in Sweden: genotype distribution over time in different epidemiological settings. *Scandinavian journal of infectious diseases.,* 31, 355-8.

Whimster, W. F., Hamilton, P. W., Anderson, N. A., Humphreys, S., Boyle, M., Sundaresan, M., Rainey, A., Giles, A., Hopster, D. & Bartels, P. H. (1996) Reproducibility of Bayesian belief network assessment of breast fine needle aspirates. *Anal Quant Cytol Histol,* 18, 267-74.

WHO (1997) Hepatitis C: global prevalence. *Wkly Epidemiol Rec,* 72, 341-4.

WHO (2003) *Pathology and Genetics of Tumours of the Breast and Female Genital Organs*, IARCPress.

Winkfield, B., Aube, C., Burtin, P. & Cales, P. (2003) Inter-observer and intra-observer variability in hepatology. *European journal of gastroenterology & hepatology,* 15, 959-66.

Winokur, T. S., McClellan, S., Siegal, G. P., Redden, D., Gore, P., Lazenby, A., Reddy, V., Listinsky, C. M., Conner, D. A., Goldman, J., Grimes, G., Vaughn, G. & McDonald, J. M. (2000) A prospective trial of telepathology for intraoperative consultation (frozen sections). *Human Pathology,* 31, 781-5.

WinZip International LLC. *WinZip.*(2005) Available at: http://www.winzip.com/ [Accessed 8-7-5].

World Wide Web Consortium. *Leading the Web to its full Potential.*(2005) Available at: http://www.w3.org [Accessed 7-8-5].

Yamauchi, K. & Fukatsu, T. (1995) A decision support system for diagnostic consultation in laboratory tests. *Medinfo,* 8 Pt 2, 1034.

Zoomify Inc. *Zoomify.*(2004) Available at: http://www.zoomify.com/ [Accessed 7-8-5].