

This is the accepted manuscript version of the following article:

Randell R, Ruddle RA, Thomas RG, Mello-Thoms C, Treanor D. (2014) Diagnosis of major cancer resection specimens with virtual slides: Impact of a novel digital pathology workstation. In press. DOI: 10.1016/j.humpath.2014.06.017

Title: Diagnosis of major cancer resection specimens with virtual slides: Impact of a novel digital pathology workstation

Authors: Rebecca Randell PhD BSc(Hons)^a, Roy A. Ruddle PhD BSc(Hons)^b, Rhys G. Thomas PhD BSc(Hons)^b, Claudia Mello-Thoms MS PhD^c, Darren Treanor PhD BSc(Hons) MB BCh FRCPath^{d+e}

Affiliations and addresses:

^aSchool of Healthcare, University of Leeds, Leeds LS2 9UT, UK

^bSchool of Computing, University of Leeds, Leeds LS2 9JT, UK

^cMedical Radiation Sciences, The University of Sydney, Sydney NSW 2150, Australia

^dSt. James's University Hospital, Leeds Teaching Hospital NHS Trust, Leeds LS9 7TF, UK

^eLeeds Institute of Cancer & Pathology, University of Leeds, Wellcome Trust Brenner Building, St. James's University Hospital, Leeds LS9 7TF, UK

Corresponding author: Rebecca Randell, Email: r.randell@leeds.ac.uk, Telephone: 0113 3431337

Key words: digital pathology, virtual slides, whole slide imaging, telepathology, time to diagnosis

Running title: Impact of a novel workstation on diagnosis

Conflict of interest and funding disclosures: No conflicts of interest to declare. This research was funded by the National Institute for Health Research (NIHR) under the New and Emerging Applications of Technology (NEAT) programme.

Summary: Digital pathology promises a number of benefits in efficiency in surgical pathology, yet the longer time required to review a virtual slide than a glass slide currently represents a significant barrier to the routine use of digital pathology. We aimed to create a novel workstation that enables pathologists to view a case as quickly as on the conventional microscope. The Leeds Virtual Microscope (LVM) was evaluated using a mixed factorial experimental design. Twelve consultant pathologists took part, each viewing one long cancer case (12-25 slides) on the LVM and one on a conventional microscope. Total time taken and diagnostic confidence were similar for the microscope and LVM, as was the mean slide viewing time. On the LVM, participants spent a significantly greater proportion of the total task time viewing slides and revisited slides more often. The unique design of the LVM, enabling real time rendering of virtual slides while providing users with a quick and intuitive way to navigate within and between slides, makes use of digital pathology in routine practice a realistic possibility. With further practice with the system, diagnostic efficiency on the LVM is likely to increase yet more.

1. Introduction

Digital pathology is promoted as offering many potential benefits, one of which is increased efficiency of pathologists and laboratories. One time and motion study suggests that digital pathology, by removing the need for pathologists to organise glass and paper, could result in a 13.4% direct efficiency improvement [1]. This requires viewing the slides to be no less efficient than on a conventional microscope, but currently the time required to read a virtual slide is a significant barrier to the introduction of digital pathology [2-6].

Five studies have systematically compared the time taken to read glass slides and virtual slides (as opposed to studies that rely on self-report data, e.g. [7, 8]). Early work by our group compared the time taken by two pathologists across seven diagnostic tasks. Average time per task was 67% longer with the virtual slides than the glass slides, but was over 300% longer for certain tasks [9]. Velez et al. compared the time taken by three pathologists to examine dermatopathology cases [10]. Average time spent per slide was 23 seconds on the microscope, 34 seconds with the in-house slide viewing software, and 38 seconds with vendor supplied slide viewing software. Some evidence has been presented to suggest that, using a slide tray view for cases containing approximately eight slides, viewing virtual slides takes moderately experienced users a similar amount of time as it would to view glass slides [11]. Our work has demonstrated how combining ultra-high resolution displays with virtual reality (VR) technology, techniques for fast interaction and high usability, first in a wall-sized display [12] and then a three-screen digital pathology workstation [13], allowed pathologists to make diagnoses from individual virtual slides as quickly as they could using glass slides.

Most cases have more than one slide, and the time for diagnosis is linearly proportional to the number of slides in a case [14], so any speed penalty of virtual slide viewing software is

multiplied . However, this is counteracted by an advantage of virtual slides, which is that the time spent changing slides is almost zero. In this paper, we present: (1) a new version of the Leeds Virtual Microscope (LVM), where pathologists can view large cases as a seamless series of slides, rather than many individual ones; and (2) the results of a controlled experimental evaluation of the LVM, comparing time to view a cancer resection case on the LVM to the time taken on the conventional microscope, using larger cases than in previous studies.

1.1 The Leeds Virtual Microscope

This new version of the LVM is novel for the following reasons:

1. It is the only digital pathology workstation that provides the pathologist with a similar visual field as the conventional microscope. The LVM combines two high-resolution medical grade screens, a Barco 6.7 megapixel Coronis Fusion and a 3.1 megapixel Nio screen (Barco N.V., Kortrijk, Belgium) (see Figure 1), to provide an almost 10 megapixel display. This is driven from a standard workstation PC containing two professional grade Nvidia Quadro graphics cards (Nvidia Corporation, Santa Clara, California, USA).
2. It provides a unique three view layout. The entirety of the 6.7 megapixel screen is occupied by the detail view of the case. The 3.1 megapixel screen is composed of two separate areas. The upper half provides a gigantic diagnostic ‘thumbnail’ overview of the current slide (the slide at the centre of the detail view), shown at a fixed magnification. The lower half provides glass-sized images of all slides in the case, with the slide metadata (slide number, stain, block description) displayed at all times.
3. The LVM was designed to enable the pathologist to seamlessly pan within and between slides and zoom in, to see the detail of an individual slide, and out, to see the whole case at low magnification. When a case is first opened, the magnification automatically adjusts

so that the whole case is fitted to the detail view. All interaction with the LVM is performed using a standard keyboard and a SteelSeries Xai high-resolution gaming mouse (SteelSeries ApS, Valby, Denmark). Zooming is performed using the scroll wheel on the mouse or by pressing keys on the keyboard.

4. The interface has custom algorithms which let the pathologist pan precisely, quickly and with minimal fatigue just using the case overview. The case overview allows continuous panning with the mouse by dragging, while a click of the mouse or the press of a key on the keyboard allows the user to quickly jump between distant locations on a given slide and in the case. A minimum and maximum zoom magnification limit was imposed so the user could neither view an image that was too pixellated or zoomed out so far that no detail could be viewed on any of the slides.
5. The high resolution display allows the pathologist to make rapid decisions about tissue and disease processes at low magnification, reducing the need to zoom in to inspect features or confirm findings, facilitating a 'Gestalt'-driven diagnostic process [15].

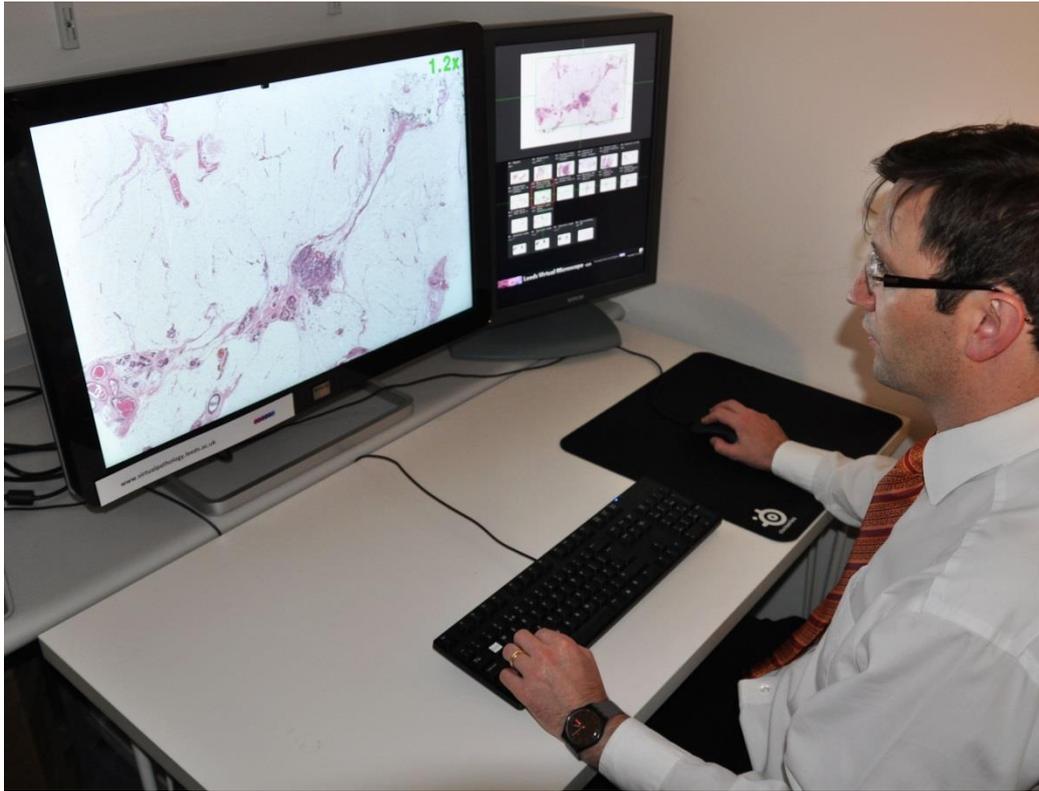


Figure 1: The Leeds Virtual Microscope. The two screens provide almost 10 megapixels. Main screen provides detail view while right-hand screen provides slide overview and case overview.

2. Materials and Methods

2.1 Participants and study protocol

A controlled user experiment was run using a mixed factorial experimental design. Twelve consultant pathologists participated, four breast specialists, four GI specialists, and four gynaecology specialists. There was variation in experience of pathology and use of virtual slides according to subspecialty (see Table 1).

		Breast	GI	Gynae
Experience of pathology (yrs)	Median	19	20	26
	IQR	(14-24)	(19-21)	(19-27)
Number of virtual slides seen before	Median	200	120	50
	IQR	(150-600)	(40-560)	(43-81)

Table 1: Participants' experience of pathology and digital pathology according to subspecialty. GI, gastrointestinal; Gynae, gynaecology; IQR, interquartile range.

Each participant was given training on how to use the LVM, lasting between 30 minutes and one hour. During the experiment, participants were given two long cancer cases to report, reporting one on the LVM and one on the conventional microscope. These cases would be typical of the daily work of a general pathologist, requiring them to assess the slides to make a diagnosis of cancer, stage and grade it. To make the tasks as similar to the participants' routine work as possible, participants were given cases that were relevant for their subspecialty. For each subspecialty, two cases were used and the case used in each condition was counterbalanced, so that each case was viewed an equal number of times in each condition. The breast cases consisted of 18 and 19 slides, both GI cases consisted of 12 slides, and the gynaecology cases consisted of 18 and 25 slides. All slides were three micron sections stained with routine haematoxylin and eosin. The slides were scanned on an Aperio T3 scanner with a 40× objective lens and a spatial sampling period of 0.25 µm/pixel. Images were compressed with conventional JPEG compression, quality 70. For the microscope tasks, a Leica DMRB microscope with 2.5×, 5×, 10×, 20× and 40× objectives was used. The order of the technology used was also counterbalanced, removing systematic bias caused by

practice effects or boredom effects [16].

While undertaking the experimental task on the LVM, log files recorded their interaction with the technology. A video camera was placed behind the participant to capture non-slide activities like reading the clinical details and filling in the proforma. When using the conventional microscope, one video camera recorded the microscope stage and one video camera was placed behind the participant, again to capture non-slide activities, as well as assisting in determining the order in which slides were viewed. For each case, participants were asked to complete a proforma and rate their confidence in their diagnosis on a 7-point Likert scale (from 1 to 7, where 1 = not confident at all and 7 = very confident). The proformas captured standard information that would have to be recorded when reporting such a case. To reduce the impact of fatigue, participants were given a short break between the tasks on the LVM and the conventional microscope. Following the experiment, participants were asked to complete a System Usability Scale questionnaire to capture their perception of the usability of the LVM [17]. Participation in the experiment took between 45 minutes and one hour. Local Research Ethics Committee approval for this research was obtained (Multicentre Research Ethics Committee 10/H1307/12) and written consent was gained from all participants.

2.2 Analysis

The total time taken for each task was measured from the video recordings. The video data was analysed to capture how time was spent across different activities (viewing the slide, changing slide, reading clinical details, making notes and completing the proforma) and, for trials on the microscope, the mean magnification used, the number of times a slide was returned to, and the order in which slides were viewed. For trials on the LVM, the mean

magnification used, the number of times a slide was returned to, and the order in which slides were viewed was captured from the log files.

Initial analyses indicated significant variation in total time taken according to the case so a normalised total time taken (total time taken expressed as a percentage of mean total time taken for all trials for that case) was used, an approach that we have used in previous studies [13]. We hypothesised there would be more revisiting of slides with the LVM than the microscope because the cost of switching slides is lower. In any given case, pathologists would be expected to revisit a small number of slides, and that number may not vary much with the total number of slides in a case because only a few will be critical. Therefore, the number of additional slide views (number of views minus number of slides) was analysed, rather than the average number of views per slide. Confidence in diagnosis, normalised total time taken, normalised number of additional slide views, and mean magnification were analysed using analyses of variance (ANOVAs), with technology as a within-participant factor and subspecialty as a between-participant factor. This analysis was undertaken using IBM SPSS Statistics 19. System Usability Scale scores were calculated for each participant using the standard methodology for the questionnaire [18].

3. Results

All reported diagnoses were considered to be within the usual variability seen in pathology diagnosis and therefore all trials were included in the subsequent analysis.

3.1 Time taken and interaction

Mean normalised time taken for each combination of technology and subspecialty is shown in Figure 2. An ANOVA showed no significant difference in normalised total time taken

between the LVM and the conventional microscope ($F(1, 9) = 0.18, p > .05$). The effect of subspecialty was non-significant ($F(2, 9) = 0.00, p > .05$). The mean viewing time per slide was also similar for the LVM and the microscope (37.5 and 35.7 seconds respectively).

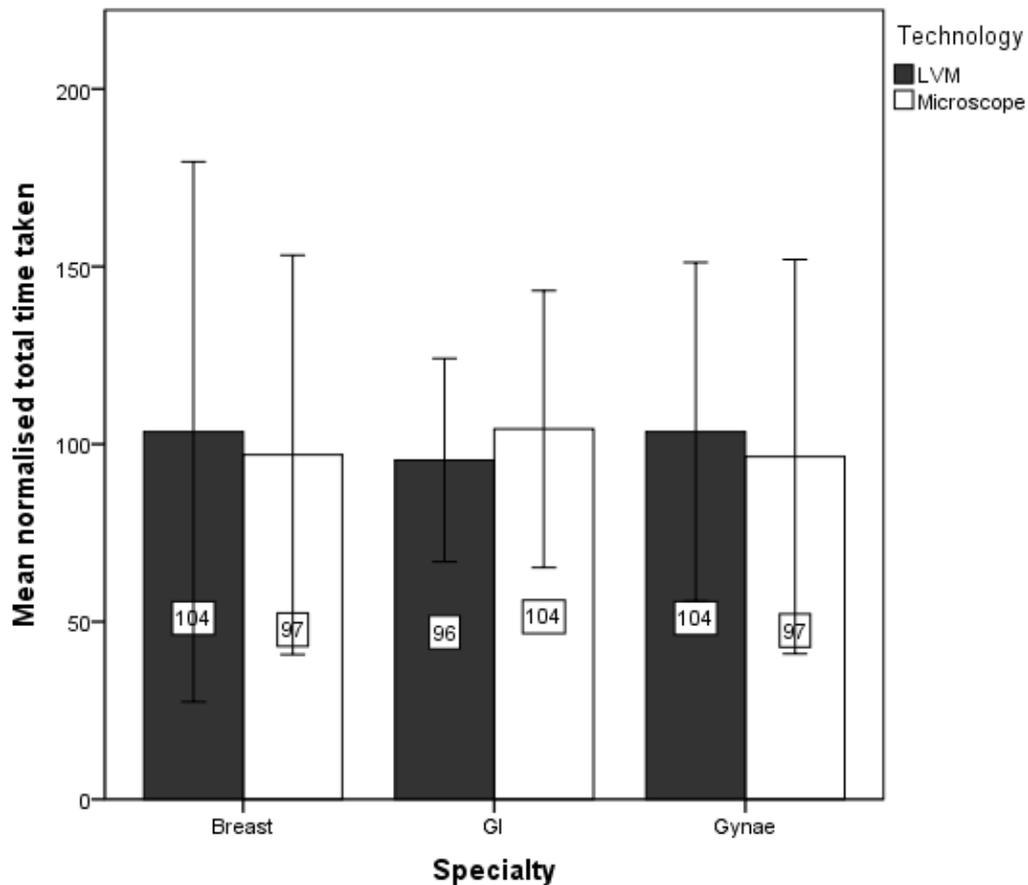


Figure 2: Mean normalised total time taken (total time taken expressed as a percentage of mean total time taken for all trials for that case) with error bars showing 95% confidence interval (CI). GI, gastrointestinal; Gynae, gynaecology; LVM, Leeds Virtual Microscope.

Figure 3 shows how time was spent on the microscope and the LVM. The LVM removes the need for physical loading and unloading of glass slides, which took up 16% of time on the microscope (15% for breast and GI cases and 18% for gynaecology cases). The percentage of time spent viewing slides was greater than on the microscope and an ANOVA showed this to

be a significant difference ($F(1, 9) = 112.33, p < .05$). The effect of subspecialty on the percentage of time spent viewing slides was non-significant ($F(2, 9) = 2.82, p > .05$).

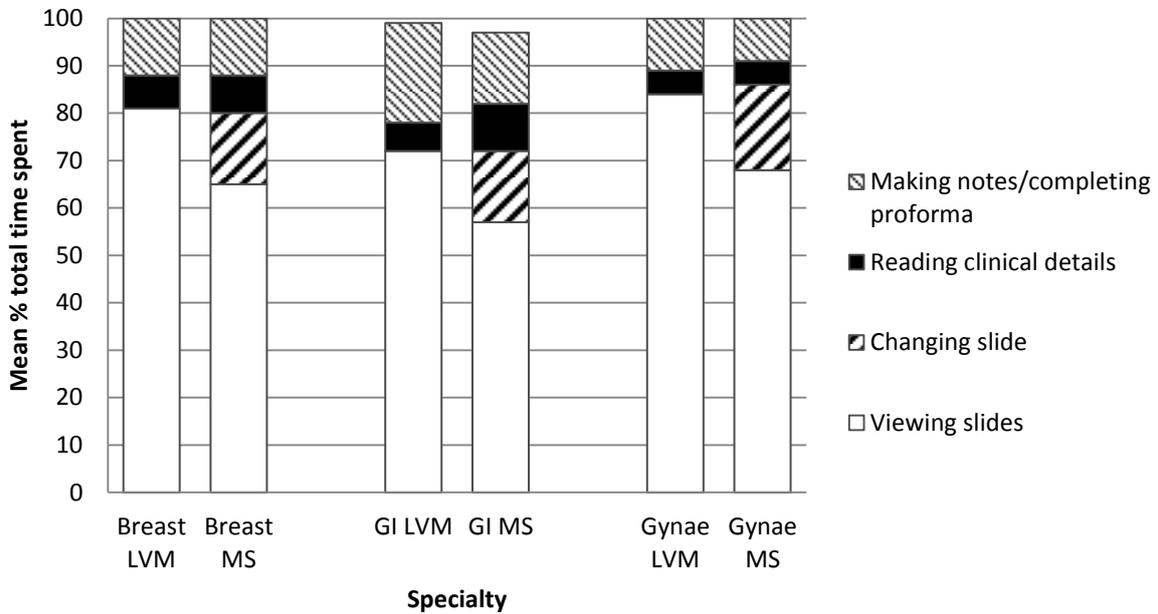


Figure 3: Mean percentage of time spent in different activities. Note that the percentage of time on key activities does not add up to 100% in all columns because of time spent in activities that could not be coded meaningfully, e.g. brief pauses. GI, gastrointestinal; Gynae, gynaecology; LVM, Leeds Virtual Microscope; MS, microscope.

After normalising the distribution of the number of additional slide views data by applying a square root transformation, an ANOVA showed that participants made significantly more additional slide views with the LVM than the microscope ($F(1, 9) = 6.38, p < .05$, see Figure 4), confirming our hypothesis. In absolute terms the difference was large, with participants making an average of 6.6 times more additional slide views with the LVM than the microscope. The effect of subspecialty was non-significant ($F(2, 9) = 0.36, p > .05$).

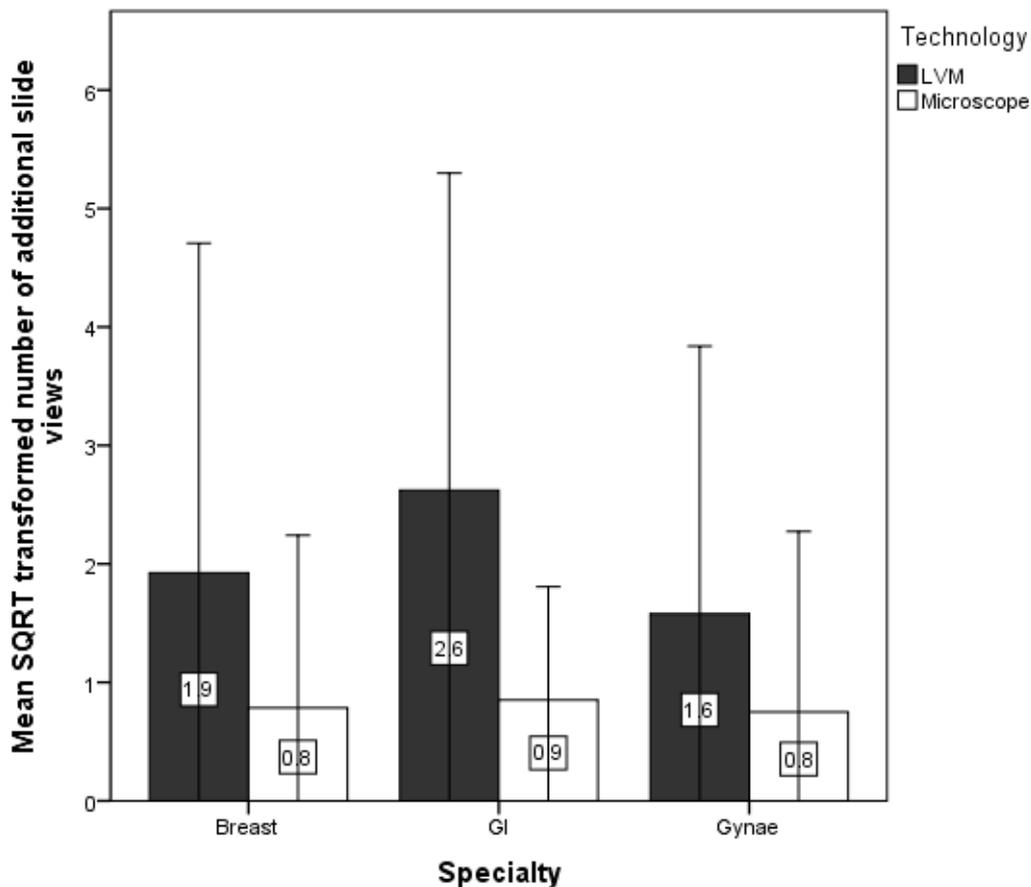


Figure 4: Mean number of additional slide views (number of views minus number of slides), normalised by applying a square root transformation, with error bars showing 95% confidence interval (CI). GI, gastrointestinal; Gynae, gynaecology; LVM, Leeds Virtual Microscope; SQRT, square root.

Mean magnification was 6.8 (SD 3.3) on the LVM and 6.2 (SD 3.3) on the microscope. An ANOVA showed no significant difference according to technology ($F(1, 7) = 0.15, p > .05$) or subspecialty ($F(2, 7) = 3.18, p > .05$).

3.2 Pathologist perceptions

Although confidence in diagnosis was slightly lower on the LVM (see Figure 5), an ANOVA showed no statistically significant difference ($F(1, 9) = 3.95, p > .05$). The effect of

subspecialty was non-significant ($F(2, 9) = 0.81, p > .05$). System Usability Scale scores can range from 0 to 100, where higher scores indicate better usability. Scores of 70 or above are considered to be acceptable, with better products scoring in the high 70s or above [17]. The mean System Usability Scale score for the LVM was 78 (range 58-93).

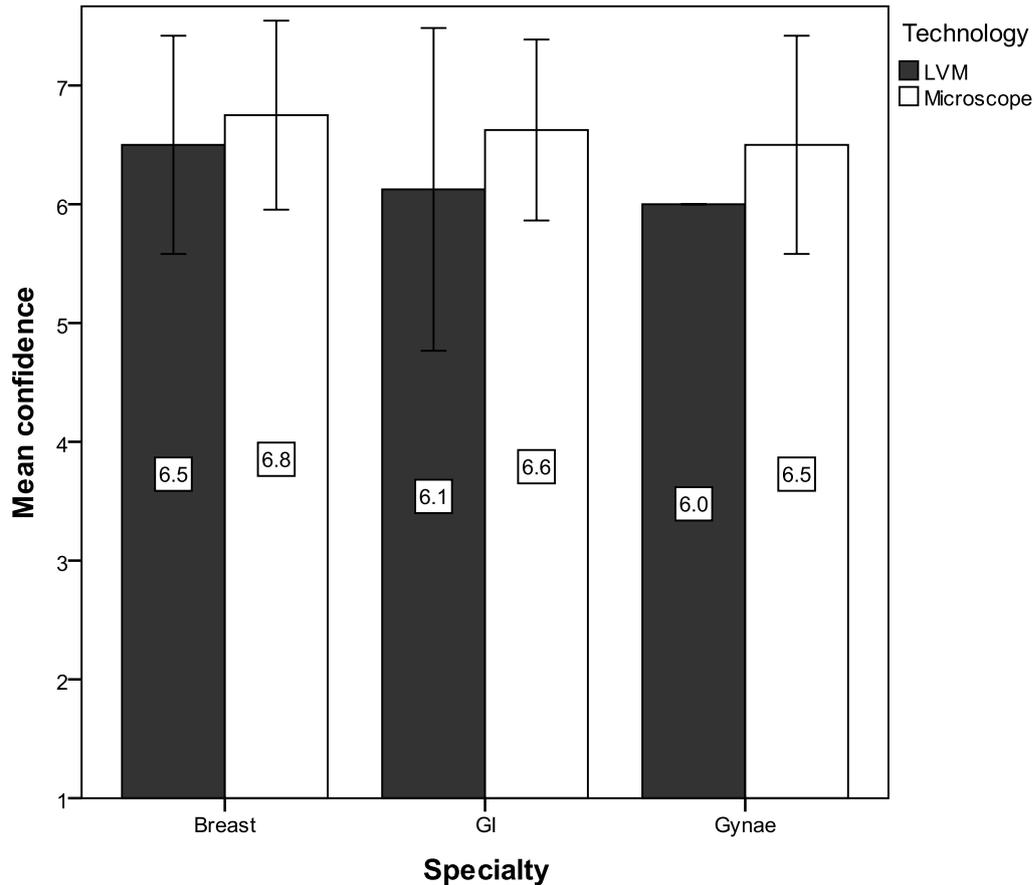


Figure 5: Mean confidence with error bars showing 95% confidence interval (CI). GI, gastrointestinal; Gynae, gynaecology; LVM, Leeds Virtual Microscope.

4. Discussion

Digital pathology has the potential to significantly impact the efficiency of pathology. These changes are already being seen in the areas of providing second opinions and remote diagnosis, but uptake in primary diagnosis has been slow. The usability and efficiency of the

pathologists' workstation is likely to be a significant contributor. Our group has spent seven years studying the role of digital pathology in the efficiency of reading slides, beginning with work to confirm the suspected inefficiency of the systems [9], through detailed qualitative and quantitative studies of how pathologists work [14, 19], to the pathologist-centred design and evaluation of a novel digital pathology workstation [13].

This paper reports the evaluation of the latest result of that work. It is the first study to objectively show no significant difference between virtual and glass slides in diagnostic speed for multi-slide cases, using larger cases than previous studies, reflective of the sort of work in a typical surgical pathology workload. The impressive performance of the LVM can be credited to its unique design, enabling real time rendering of virtual slides while providing users with a quick and intuitive way to move around and between slides. It is noteworthy that pathologists spent 16% more time productively on diagnostic work (looking at slides) with the LVM than with a conventional microscope, using that to revisit slides, probably to confirm or check aspects of the diagnosis. The absence of significant difference in mean magnification according to technology is in contrast to our previous work, where we found pathologists used a significantly higher magnification when viewing virtual slides [13], and this is possibly due to the use of medical grade screens in the current study.

The mean viewing time per slide was similar for the LVM and the microscope. Because changing slides on the LVM is almost instantaneous, it could be expected that, if all other aspects of the task were equivalent in time across the two conditions, total time taken would be faster on the LVM. However, slides were returned to more frequently on the LVM, adding additional slide viewing time. The results suggest that, because it was so easy and fast to revisit slides in the LVM, being able to return to a slide with just one click, participants chose

to do so instead of holding large amounts of diagnostic information in memory. This fits with existing research that suggests people adopt least-effort tradeoffs which mean they rely on imperfect ‘knowledge in-the-head’ (memory) when perfect ‘knowledge in-the-world’ (in this case, the information contained within the slides) requires effort to access [20], as is the case with the microscope; with the LVM it seems that the reduced effort to access the slides led to greater use of perfect ‘knowledge in-the-world.’ As such, we can hypothesise that use of this technology would decrease cognitive load and possibly pathologists’ fatigue in long and complex cancer cases. While our evaluation did not explore cognitive load or fatigue, we plan to investigate these issues in future work.

Alternatively, the additional slide views on the LVM may have been a result of participants exploring the technology or have been influenced by their expectations of how long a task should take, motivating them to use the ‘extra time’ to continue viewing the slides. As revisiting slides is a behavior seen more often in trainee pathologists who are having difficulty developing strategies for collecting and summarising the information in a case, it could be that the behavior reflected a lack of familiarity, either with the LVM or virtual slides more generally. What we do not currently know is whether this behavior would persist outside of the experimental situation or as participants become more familiar with the technology. Although confidence in diagnosis was similar between the LVM and the conventional microscope, the signal to noise ratio is likely to be higher on the LVM because participants were less familiar with the technology and this would motivate participants to confirm their diagnosis by returning to the slides.

It is important to note that the lack of significant difference reported in this study does not demonstrate that the LVM and microscope are equivalent in terms of time to diagnosis.

Demonstrating equivalence would need a much larger study and, for that reason, this study should be seen as exploratory research. However, given that participants had 10-20 years' experience of using a microscope, and only a single short training session with the LVM, the overall time to diagnosis on the LVM is encouraging. Even if viewing a case on the LVM takes the same time as on the microscope, other workflow benefits around sorting paper and glass, distributing, assembling or reconciling cases, and on-screen synoptic reporting will likely lead to global efficiency benefits of 5-10% (conservatively) in diagnosis. With further practice with the system, diagnostic efficiency on the LVM is likely to increase yet more. With such faster diagnosis, pathologist acceptance is likely to increase, which will in turn drive greater adoption, providing the promised productivity and cost benefits for healthcare organisations. There is a need for large scale trials of digital pathology to assess diagnostic accuracy and impact on efficiency; by reducing the time to diagnosis, the LVM makes such large scale trials more achievable and therefore more likely.

Acknowledgements

We are very grateful to the histopathologists at Leeds Teaching Hospitals NHS Trust who participated in this experiment. We would like to thank our project advisory board for their input into the interpretation of the results, particularly Dr David Brett, Professor Julia Brown, and Professor Jenny Hewison. We would also like to thank Mr Mike Hale and Mr Dave Turner for their assistance in setting up various aspects of this evaluation. This report is independent research commissioned by the National Institute for Health Research (NIHR) under the New and Emerging Applications of Technology (NEAT) programme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors acknowledge the support of the NIHR, through the Comprehensive Clinical Research Network.

References

- [1] Stratman C, Drogowski L, Ho J. Digital Pathology in the Clinical Workflow: A Time & Motion Study [conference presentation]. Pathology Visions; October 24-27; San Diego, CA; 2010. Available from: <https://digitalpathologyassociation.org/pathology-visions-2010>. Last accessed: June 6 2014.
- [2] Evans AJ. Transitioning to Primary Diagnosis by WSI Telepathology: Time to Go Live at UHN [conference presentation]. Pathology Visions; September 29 - October 1; San Antonio, TX; 2013.
- [3] Lele SM, Bhaduri A, Madiwale C, et al. Whole Slide Scans and Web-Based Conferencing Tools Used in Combination Provide Accurate Assessment of Cases for Real-Time Global Clinical Consultation [conference presentation]. Pathology Visions; October 24-27; San Diego, CA; 2010. Available from: <https://digitalpathologyassociation.org/pathology-visions-2010>. Last accessed: June 6 2014.
- [4] Têtu B, Evans A. Canadian Licensure for the Use of Digital Pathology for Routine Diagnoses. Arch Pathol Lab Med 2014;138:302-304.
- [5] Treanor D. Virtual slides: an introduction. Diagn Histopathol (Oxf) 2009;15:99-103.
- [6] Wienert S, Beil M, Saeger K, Hufnagl P, Schrader T. Integration and acceleration of virtual microscopy as the key to successful implementation into the routine diagnostic process. Diagn Pathol 2009;4:3.
- [7] Furness P. A randomized controlled trial of the diagnostic accuracy of internet-based telepathology compared with conventional microscopy. Histopathology 2007;50:266-273.
- [8] Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR. Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies. Hum Pathol 2006;37:322-331.
- [9] Treanor D, Quirke P. The Virtual Slide and Conventional Microscope - a Direct Comparison of Their Diagnostic Efficiency. J Pathol 2007;213:7a.
- [10] Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications in dermatopathology. Hum Pathol 2008;39:1341-1349.
- [11] Stratman C, Drogowski L, Ho J. The Impact of Digital Pathology on Pathologists' Time: Continued Time & Motion Studies [conference presentation]. Pathology Visions; San Diego, CA; 2011.
- [12] Treanor D, Jordan-Owers N, Hodrien J, Wood J, Quirke P, Ruddle RA. Virtual reality Powerwall versus conventional microscope for viewing pathology slides: an experimental comparison. Histopathology 2009;55:294-300.
- [13] Randell R, Ruddle RA, Mello-Thoms C, Thomas RG, Quirke P, Treanor D. Virtual reality microscope versus conventional microscope regarding time to diagnosis: an experimental study. Histopathology 2013;62:351-358.
- [14] Randell R, Ruddle RA, Quirke P, Thomas RG, Treanor D. Working at the microscope: analysis of the activities involved in diagnostic pathology. Histopathology 2012;60:504-510.
- [15] Nodine CF, Kundel HL. The cognitive side of visual search in Radiology. In: O'Regan JK, Levy-Schoen A, editors. Eye Movements: From Physiology to Cognition, Amsterdam: Elsevier Science Publishers BV; 1987, p. 573-582.
- [16] Field A. Discovering statistics using SPSS. London: Sage 2009.
- [17] Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. Int J Hum-Comput Int 2008;24:574-594.
- [18] Tullis T, Albert B. Measuring the user experience. Burlington, MA: Morgan Kaufmann; 2008.
- [19] Randell R, Ruddle RA, Thomas R, Treanor D. Diagnosis at the microscope: a workplace study of histopathology. Cogn Tech Work 2012;14:319-335.

[20] Gray WD, Fu WT. Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. Proceedings of the SIGCHI conference on Human factors in computing systems; ACM; 2001. p. 112-119. Available from: <http://dl.acm.org/citation.cfm?id=365061>. Last accessed: June 6 2014.