Authors: Rebecca Randell[a], Roy A. Ruddle[b], Claudia Mello-Thoms[c], Rhys G. Thomas[b], Phil Quirke[a] and Darren Treanor[ad]

Title: Virtual reality microscope versus conventional microscope on time to diagnosis: an experimental study

Affiliations and addresses:

[a]Leeds Institute of Molecular Medicine, University of Leeds, Wellcome Trust Brenner Building, St. James's University Hospital, Leeds LS9 7TF, UK

[b]School of Computing, University of Leeds, Leeds LS2 9JT, UK

[c]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh PA 15232, USA

[d]St. James's University Hospital, Leeds Teaching Hospitals NHS Trust, Leeds LS9 7TF, UK

Corresponding author:

Rebecca Randell

Email: r.randell@leeds.ac.uk

Telephone: 0113 3438509

Running title: Virtual reality microscope evaluation

Key words: histopathology, digital pathology, whole slide imaging, virtual slide

Aims: Digital slides promise a number of benefits in terms of efficiency. However, currently it can take up to 60% longer to perform diagnoses using digital slides. The aim was to create and evaluate a virtual reality (VR) microscope that is as efficient as the conventional microscope, to support the introduction of digital slides into routine practice.

Methods and Results: A VR microscope was designed and implemented by combining ultra-high resolution displays with VR technology, techniques for fast interaction, and high usability. It was evaluated using a mixed factorial experimental design with technology and task as within-participant variables and grade of histopathologist as a between-participant variable. Time to diagnosis was similar for the conventional and VR microscope. However, there was a significant difference in the mean magnification used between the two technologies, with participants working at a higher level of magnification on the VR microscope.

Conclusions: The results suggest that, with the right technology, efficient use of digital pathology for routine practice is a realistic possibility. Further work is required to explore what magnification is required on the VR microscope for histopathologists to identify diagnostic features, and the effect on this of the digital slide production process.

Introduction

Digital slides promise a number of benefits in terms of efficiency. However, currently, it takes significantly longer to perform diagnoses using digital slides,[1, 2] in some cases up to 60% longer,[3] with the result that use of such slides is currently largely restricted to education and external quality assessment schemes. We found that a very high resolution (54-million pixel) wall-sized display allows diagnoses to be made as fast as with a microscope,[4] but a desktop solution that enables this is required for use of digital slides to be feasible for routine practice.

Viewing digital slides on a computer takes longer than viewing glass slides on a conventional microscope for a number of reasons:

1. A standard computer monitor gives a much smaller field than that of a conventional microscope.[5]

2. With most slide viewing software the user clicks and drags the slide. Because of the size of the images, navigating across a slide and switching back and forth between regions of interest takes a long time. Most software provides a small thumbnail of the whole slide and speed can be increased by use of the thumbnail to navigate the slide. However, histopathologists complain the thumbnails are not large enough to determine which areas of the slide they wish to view at higher magnification.[6]

3. Large file sizes result in a rendering lag as the user navigates around the slide,[7] leading to users complaining of 'pixelation' as pieces of the area of the slide they want to view upload onto the screen.[8]

With the aim of enabling diagnoses to be made as quickly on a digital slide as on a glass slide on a conventional microscope, we have developed a VR microscope, which addresses these issues:

1. Three high resolution desktop publishing grade 27" screens are combined to provide an 11 mega-pixel display (see Figure 1), giving a field of 0.07 mm² at 40×, compared to the field on a microscope at the same magnification of 0.03 mm².

2. All navigation is done in a giant 'diagnostic' thumbnail, 1440 pixels wide, in the top right-hand corner of the screen. When a slide is first opened, the magnification automatically adjusts so that the whole slide is fitted to the screen.

3. Underlying this, VR technology enables digital slides to be rendered in real-time, significantly reducing the problem of 'pixelation'.

[Figure 1 should be placed approximately here]

Materials and methods

A controlled user experiment was run using a mixed factorial experimental design. Each participant performed six diagnostic tasks on both the VR microscope and on the conventional microscope. The order of the technology used, the task order and the slide set used with each technology were counterbalanced.

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.

Experimental task

On both the VR microscope and on the conventional microscope, participants were given three gastrointestinal (GI) biopsy cases and three skin specimen cases to diagnose. Each case comprised a single glass or digital slide and a one line clinical history (see Table 1 for details of the cases). Participants were asked to record their diagnosis on an answer sheet 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).

[Table 1 should be placed approximately here]

We chose to focus on GI biopsies and skin specimens because together these make up 51% of the slides produced at Leeds Teaching Hospitals National Health Service (NHS) Trust where this work was undertaken. 31% of cases contain just one slide and this is typical of many GI biopsy and skin specimen cases. The cases were selected by one of the authors (DT), a consultant histopathologist, and were considered to be straightforward cases from a diagnostic point of view.

All slides were stained with haematoxylin and eosin. The slides were scanned on an Aperio T3 scanner (Aperio Technologies, San Diego, CA, USA) 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 (Leica Microsystems GmbH, Wetzlar, Germany) with 2.5×, 5×, 10×, 20× and 40× objectives was used.

Participants

Eight specialist trainee histopathologists and eight consultant histopathologists participated in the experiment. Four of the consultants were GI specialists and four were skin specialists. Median experience of pathology was 4 years (range 0.5-5.5 years) for the trainees and 17.5 years (range 6-28 years) for the consultants. The median number of digital slides that participants estimated they had previously used was 35 (range 0-2000) for trainees and 35 (range 10-50) for consultants.

Procedure

Before the experiment, on a separate occasion, each participant was given one-to-one training on how to use the VR microscope, lasting a maximum of 30 minutes. The researcher (RR) demonstrated how to use the technology and then the participant completed a series of basic navigation tasks, e.g. scanning the epidermis on a skin specimen slide. Participants then diagnosed two previously unseen slides, one GI biopsy and one skin specimen, under supervision on the VR microscope. If necessary, the researcher gave feedback on how the participant could have made better use of the functionality of the VR microscope, to make their use of the technology more efficient.

Before undertaking the experimental tasks using the VR microscope, participants were given the opportunity to refamiliarise themselves with the technology by doing two practice tasks, one GI biopsy and one skin specimen, and to ask the researcher any questions about the technology. While undertaking the experimental tasks on the VR microscope, a video camera placed behind the participant recorded the session and log files recorded the participants' interaction with the technology. When using the conventional microscope, one video camera recorded down the microscope and one video camera recorded the stage. Participants were given a pen with which they could annotate the slides if they wished. Following the experiment, a brief interview was conducted with participants to explore their experience of using the VR microscope. This interview was audio recorded.

To reduce the impact of fatigue, participants were invited to take a short break between the VR microscope and the conventional microscope. Participation in the experiment took between 45 minutes and one hour.

Analysis

Diagnoses given by participants were categorised as correct, a minor error (one with no consequences for patient care, e.g. normal tissue misdiagnosed as showing nonspecific changes), or a major error (one with consequences for patient care, e.g. compound naevus misdiagnosed as

pericytoma). Trials where the participant recorded that they did not know the diagnosis or where the diagnosis given was classified as a 'major error' were excluded from analysis on the basis that the participant did not adequately complete the task (for details, see Results).

Quantitative data were analysed using analyses of variance (ANOVA), with technology and task as within-participant factors and, where relevant, grade of histopathologist (consultant versus trainee) as a between-participant factor. The interviews were transcribed and analysed using thematic content analysis.[9]

Subjective assessment of magnification required on both the conventional microscope and the VR microscope to identify relevant features was undertaken by one of the histopathologist authors (DT) for those slides where one or more participants made an error.

Results

A total of eight trials (4%) were excluded, five on the VR microscope and 3 on the conventional microscope. On the VR microscope, there were two GI trials where a consultant recorded 'Don't know' (both by the same consultant who is a skin specialist) and one GI trial where a trainee recorded 'Don't know'. On the conventional microscope, there was one GI trial where a consultant recorded 'Don't know' and two GI trials where trainees recorded 'Don't know'. There were two major errors in the VR microscope trials, both made by trainees and both made on the same slide, a skin slide in set 2. In addition, although not excluded from the analysis, there were six minor errors in the VR microscope trials, one made by a consultant on a GI slide in set 1 and the others made by trainees, four on GI slides and one on a skin slide, all in set 1. No major or minor errors were made in the conventional microscope trials.

Confidence in diagnosis

Mean confidence in diagnosis amongst consultants was 6.29 on the conventional microscope and 6.21 on the VR microscope for the GI trials and 6.63 on the conventional microscope and 6.50 on the VR microscope for the skin trials (see Figure 2). Mean confidence in diagnosis amongst trainees was 5.58 on the conventional microscope and 5.08 on the VR microscope for the GI trials and 5.98 on the conventional microscope and 5.54 on the VR microscope for the skin trials. An ANOVA showed a significant main effect of technology ($F(1, 14) = 5.18$, $p < .05$). The main effect of experience was non-significant ($F(1, 14) = 4.30$, $p > .05$), as was the main effect of task ($F(1, 14) = 4.11$, $p > .05$). There were no significant interaction effects.

[Figure 2 should be placed approximately here]

Time to diagnosis

Initial analyses indicated much variation in time to diagnosis according to the particular slide so a normalised time to diagnosis (time to diagnosis expressed as a percentage of mean time to come to a diagnosis for all consultant trials for that slide where the diagnosis was classified as correct or minor error) was used. Mean normalised time to diagnosis amongst consultants was 0.88 on the conventional microscope and 1.06 on the VR microscope for the GI trials and 0.91 on the conventional microscope and 1.13 on the VR microscope for the skin trials (see Figure 3). Mean normalised time to diagnosis amongst trainees was 1.52 on the conventional microscope and 1.43 on the VR microscope for the GI trials and 1.33 on the conventional microscope and 1.39 on the VR microscope for the skin trials. An ANOVA showed no significant difference in time to diagnosis between the VR microscope and the conventional microscope ($F(1, 14) = 0.87$, $p > .05$), but the effect of experience was significant ($F(1, 14) = 5.33$, $p < .05$). The main effect of task was non-significant ($F(1, 14) = 0.13$, $p > .05$).

[Figure 3 should be placed approximately here]

Navigation and interaction

Looking at the navigation path used on the VR microscope and reviewing the videos of tasks undertaken on the VR microscope, we found that on GI slides, participants navigated in a similar way as on the conventional microscope, viewing each level of tissue in turn, despite the fact that the VR microscope would allow them to easily move between levels to compare pieces of tissue on each level (see Figure 4). On reviewing the videos, it also seemed that participants were making small adjustments to the position of the digital slide to ensure that pieces of tissue were not crossing the bezels.

[Figure 4 should be placed approximately here]

Analyses were conducted to investigate consultants' mean magnification and mean number of magnification changes. Mean magnification was 7.97 on the conventional microscope and 14.93 on the VR microscope for the GI trials and 8.62 on the conventional microscope and 12.37 on the VR microscope for the skin trials (see Figure 5). An ANOVA showed a significant main effect of

technology (F(1, 7) = 40.18, p < .01), but the main effect of task was non-significant (F(1, 7) = 1.71, p > .05).

[Figure 5 should be placed approximately here]

Mean number of magnification was 5.21 on the conventional microscope and 14.67 on the VR microscope for the GI trials and 3.67 on the conventional microscope and 7.38 on the VR microscope for the skin trials (see Figure 6). An ANOVA showed a significant main effect of both technology (F(1, 7) = 15.17, p < .01) and task (F(1, 7) = 8.22, p < .05) on number of magnification changes. The VR microscope did not indicate to participants what level of magnification they were currently at, so we assume participants were zooming in to whatever magnification level they felt was necessary in order to see the relevant diagnostic features. However, it is also likely that the ease with which the VR microscope magnification could be changed (the press of a button, or movement of the mouse scroll wheel) contributed to the number of changes that took place, and overshoot when using the scroll wheel may have increased the mean magnification.

[Figure 6 should be placed approximately here]

Diagnostic accuracy and image quality

For all diagnostic features on those slides where one or more participants had made an error, the VR microscope required at least double the magnification of that on the conventional microscope in order to make a confident diagnosis. For example, both of the major errors were made on one of the skin specimen slides which showed a benign intradermal naevus. A low power scan to evaluate the tissue can be done on 2.5× on the conventional microscope. However, naevus cells which appeared light and fine textured on the conventional microscope appeared dark and coarse textured on the VR microscope so they looked more suspicious until the magnification is increased to 5×. Ruling out epidermal involvement can be done at 5× on the conventional microscope but 10× was required on the VR microscope. Determining if there is melanocytic atypia can be undertaken at 10× on the conventional microscope but on the VR microscope the nuclear detail was unclear at the same magnification. Closer examination of the naevus cells at 20× was required to ensure that premalignant atypia is not present. The underlying causes of these differences between conventional microscopes and the VR microscope are unknown but are likely to include the imaging process itself (including acquisition and compression),the computer displays used (off the shelf displays rather than medical grade monitors) and the familiarity of the pathologists with the

technology.

Histopathologist perceptions

All participants reported that they felt the training adequately prepared them for undertaking the tasks on the VR microscope. Half of the participants, five consultants and three trainees, reported experiencing no difficulty in using the VR microscope. However, despite participants' apparent satisfaction with the training, a recurrent issue theme in the interviews related to controlling the speed of panning. One consultant reported panning too fast so that she 'overshot'. One trainee commented that he found the bezels between the screens distracting and so spent additional time positioning the slide so that the pieces of tissue did not cross screens. None of the participants commented that image quality was inadequate. During the interviews however two consultants did describe feeling less confident in their diagnosis with the digital slides, with one expressing a desire to see how participants would perform with more difficult cases.

Trainees expressed most enthusiasm for the VR microscope. One trainee, in the first six months of specialist training, described the VR microscope as being easier to use than the conventional microscope, one trainee who is a regular user of digital slides described the VR microscope as 'faster and less cumbersome' than the microscope, and another trainee described the VR microscope as being fun to use.

Perceptions of the speed with which a diagnosis could be made on the VR microscope compared to the conventional microscope varied between consultants and trainees. Five of the eight consultants felt that they worked slower on the VR microscope with only one consultant saying that they felt that the VR microscope was faster. In contrast, three trainees felt that they were able to work faster on the VR microscope and only two felt that they were slower on the VR microscope. Overall, diagnosis took a similar amount of time on the conventional and VR microscopes (see above).

Discussion

Implications for design

One motivation for this study was to design a digital microscope that performs as well as a conventional microscope. The results indicate that participants were able to quickly learn how to use the VR microscope and were able to complete the tasks successfully by only navigating in the thumbnail.

The higher number of magnification changes on the VR microscope, in comparison to the conventional microscope, could be due to limitations of the digital imaging creating a need for a higher level of magnification to make a diagnosis. However, looking at the pattern of magnification changes suggests that participants frequently zoomed in or out too far and then immediately zoomed in the other direction to correct for this. This could be due to motor control, in terms of lack of familiarity with zooming by using the scroll wheel on the mouse. However, it could also be because participants were learning what level of magnification they needed in order to be able to see the necessary detail, for example, zooming in a number of levels and then realising they did not need or want to look at such a high magnification.

On GI biopsy slides, participants would zoom in to look at a piece of tissue and then zoom out to navigate as they would on a conventional microscope, hence the higher number of magnification changes on the GI biopsy slides. This is despite the fact that the VR microscope should enable them to navigate quickly and easily on a high magnification by clicking in the thumbnail view, thereby reducing the need to zoom in and out. The persistence of this pattern of navigation may be due to the need for contextual information that is provided at lower magnification, or limited experience of digital slides meaning participants were reluctant to change their strategy. Additionally, histopathologists have to cognitively interpret what they were seeing and so, by navigating in the way that they are used to, they free up cognitive resources that otherwise would be engaged in learning and monitoring a new navigation strategy.

Implications for practice
The results presented here are based on a small study and a much larger study would be required to demonstrate equivalence with certainty. However, these initial findings and the largely positive response, achieved with only minimal training, suggest that the VR microscope represents a significant step towards making use of digital pathology for routine clinical practice a realistic possibility.

The results highlight the importance of training. It is not just about learning to use the technology *per se* but learning how to make diagnoses with that technology. Histopathologists need to understand that the digital slide is a new technology that does not replicate the conventional microscope, so they have to relearn what they should expect to see at each magnification level, as it may be that it is necessary to view the slide at a higher level of magnification than would typically be required on a conventional microscope. While we have begun to explore the issue of what level of

magnification is required, further work with multiple observers is needed.

Although this study was not designed to assess diagnostic accuracy, the diagnostic errors seen when using the VR microscope but not the conventional microscope are concerning and indicate that we cannot take for granted the direct translation of information from the conventional microscope to the VR microscope. This suggests the need for large scale trials of the technology in routine diagnostic practice to understand the significance of this finding. A barrier to undertaking such trials has been the time required to make a diagnosis using digital slides. By reducing the time to diagnosis, the VR microscope makes this a more readily achievable goal.

Competing interests

None to declare

References

1. Li X, Liu J, Xu H *et al.* A feasibility study of virtual slides in surgical pathology in China. *Human Pathology* 2007;**38**;1842-1848.
2. Velez N, Jukic D, Ho J. Evaluation of 2 whole-slide imaging applications in dermatopathology. *Human Pathology* 2008;**39**;1341-1349.
3. Treanor D, Quirke P. The virtual slide and conventional microscope - a direct comparison of their diagnostic efficiency. *Annual Meeting of the Pathological Society of Great Britain and Ireland*. Glasgow, 2007.
4. 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.
5. Treanor D. Virtual slides: an introduction. *Diagnostic Histopathology* 2009;**15**;99-103.
6. Gilbertson J, Ho J, Anthony L, Jukic D, Yagi Y, Parwani A. Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clinical Pathology* 2006;**6**;4.

7.   Jeong W, Schneider J, Turney SG *et al.* Interactive histology of large-scale biomedical image stacks. *IEEE Transactions on Visualization and Computer Graphics* 2010;**16**;1386-1395.

8.   Rojo MG, Garcia GB, Mateos CP, Garcia JG, Vicente MC. Critical Comparison of 31 Commercially Available Digital Slide Systems in Pathology. *International Journal of Surgical Pathology* 2006;**14**;285-305.

9.   Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook*. 2nd ed. ed. Thousand Oaks, California: SAGE Publications, 1994.

Set	Task	Clinical details provided	Diagnosis
1	Skin	Female, 49 yrs, skin lesion on nose, excised	Intradermal naevus
		Male, 63 yrs, skin lesion right ear excised	Basal cell carcinoma
		Female, 16 yrs, papilloma left axilla, excised	Seborrhoeic keratosis
	GI	Female, 77 yrs, diarrhoea, colonic biopsies	Normal
		Male, 53 yrs, constipation with rectal bleeding, tumour is seen in rectum on endoscopy	Adenocarcinoma
		Female, 59 yrs, anaemia query coeliac disease, duodenal biopsies	Normal
2	Skin	Male, 27 yrs, nodule left lower lip, 4 mm, excised	Haemangioma
		Male, 39 yrs, skin lesion left groin, excised	Compound naevus
		Male, 63 yrs, skin lesion left ear, query basal cell carcinoma, excised	Basal cell carcinoma
	GI	Male, 70 yrs, diarrhoea, colonic biopsies	Normal
		Female, 83 yrs, melena, tumour seen on oesophago-gastroscopy, biopsied	Squamous carcinoma
		Female, 78 yrs, thickening of caecum on CT scan, biopsies taken	Adenocarcinoma

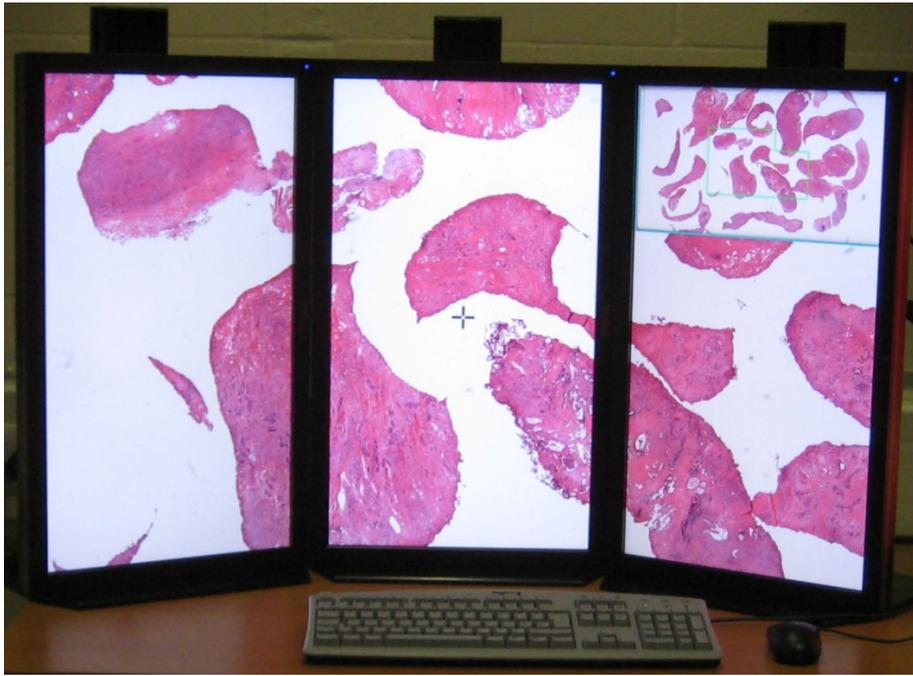Table 1: Details of cases and diagnoses used for experimental task.
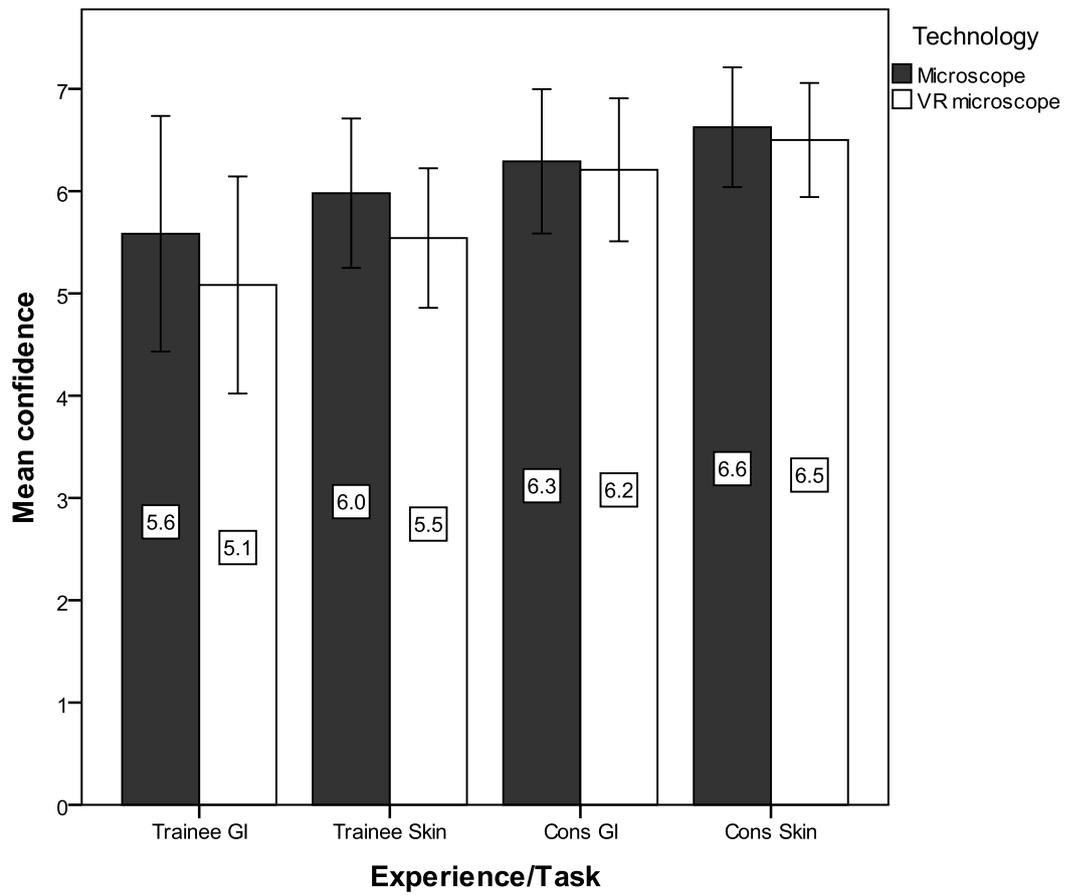
Figure 1: The VR microscope.

Figure 2: Mean confidence with error bars showing 95% confidence interval (CI).
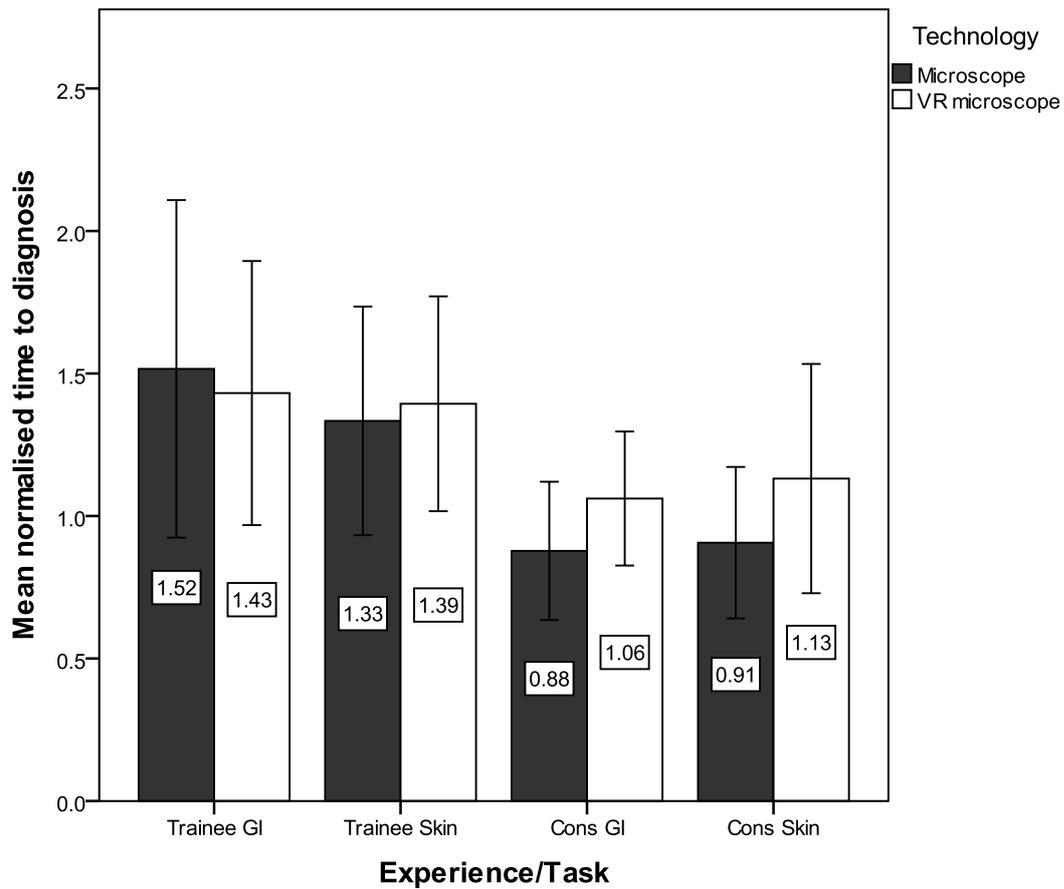
Figure 3: Mean normalised time to diagnosis with error bars showing 95% confidence interval (CI).
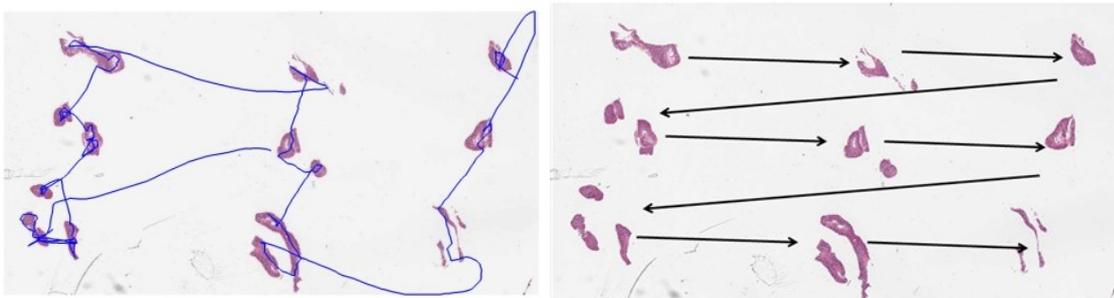

Figure 4: Navigation path on a GI biopsy task on VR microscope. Track of a participant's navigation path on a GI biopsy task on the VR microscope (left) and an alternative navigation path, which the VR microscope allows, where pieces of tissue on each level are compared (right).
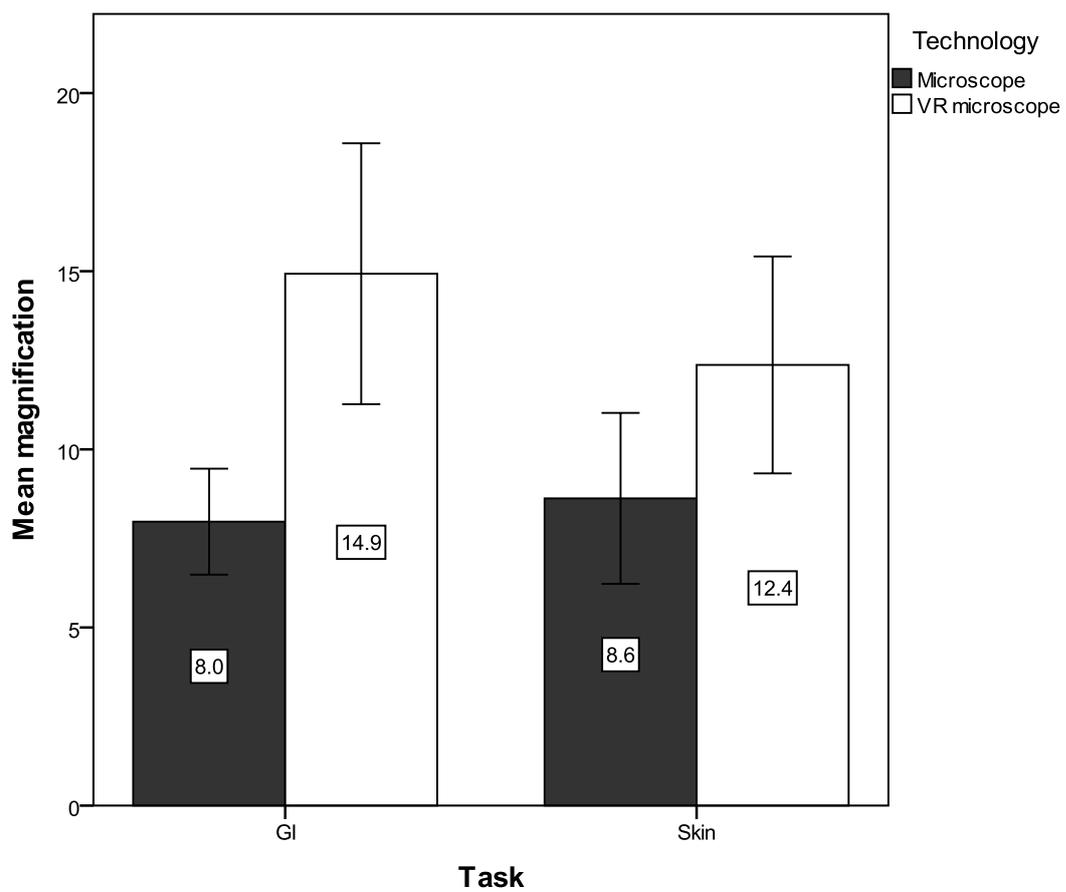
Figure 5: Mean magnification for consultants with error bars showing 95% confidence interval (CI).
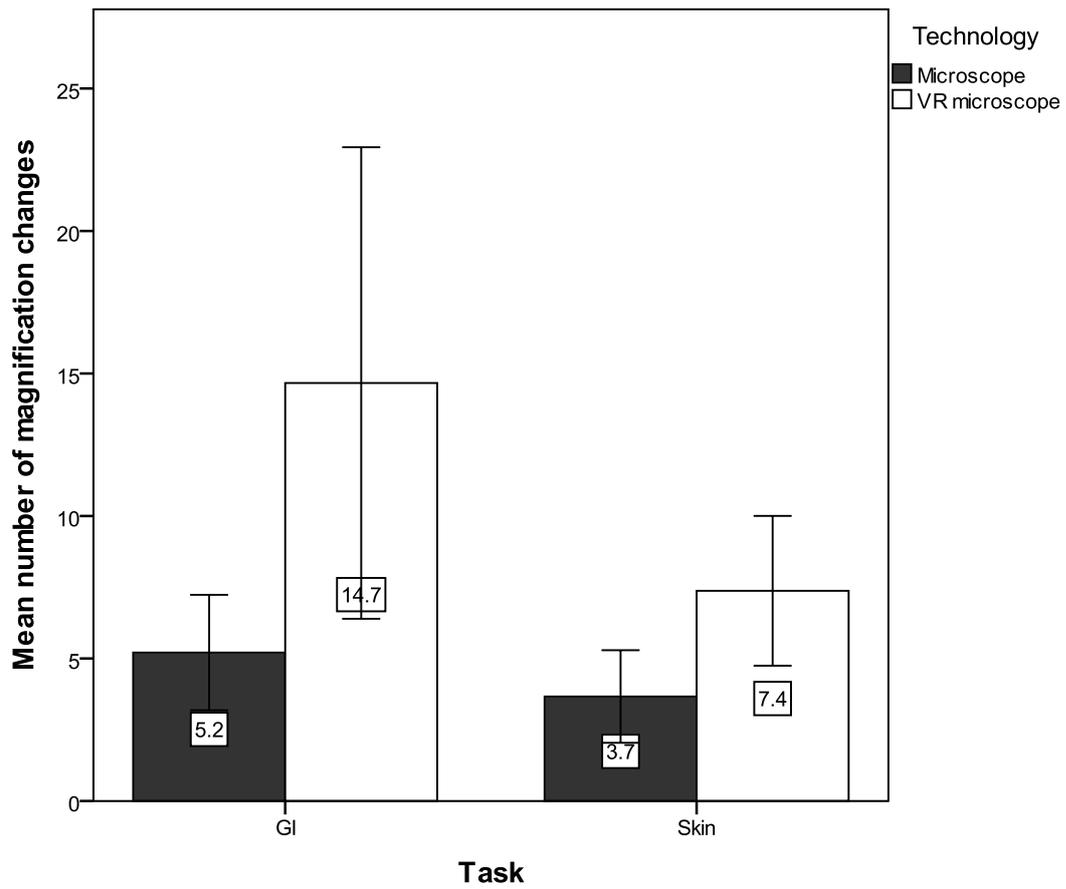
Figure 6: Mean number of magnification changes for consultants with error bars showing 95% confidence interval (CI).