Article:
Clarke, EL, Munnings, C, Williams, B et al. (2 more authors) (2020) Display evaluation for primary diagnosis using digital pathology. Journal of Medical Imaging, 7 (02). p. 1. ISSN 2329-4302

https://doi.org/10.1117/1.jmi.7.2.027501

Title Page

Title

Display Evaluation for Primary Diagnosis using Digital Pathology

Running Title

Display evaluation for digital pathology

List of Authors

Emily L Clarke[1, 2], Craig Munnings[3], Bethany Williams[1,2], David Brettle[3], Darren Treanor[1,2,4]

1. Division of Pathology and Data Analytics, University of Leeds, Leeds, UK
2. Histopathology Department, Leeds Teaching Hospitals NHS Trust, Leeds, UK
3. Medical Physics Department, Leeds Teaching Hospitals NHS Trust, Leeds, UK
4. Division of Neurobiology, Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Department of Clinical Pathology, Centre for Diagnostics, Region Östergötland, Sweden

Corresponding Author

Dr Emily Clarke

Clinical Research Training Fellow in Histopathology

University of Leeds and Leeds Teaching Hospitals NHS Trust

Beckett Street

Leeds

LS9 7TF

UK

Tel +44(0)7775 065648

Email: e.l.clarke@leeds.ac.uk

Abstract

Purpose

As pathology departments around the world contemplate digital microscopy for primary diagnosis, making an informed choice regarding display procurement is very challenging in the absence of defined minimum standards. In order to help inform the decision we aimed to conduct an evaluation of displays with a range of technical specifications and sizes.

Approach

We invited histopathologists within our institution to take part in a survey evaluation of 8 short-listed displays.

Pathologists reviewed a single haematoxylin and eosin slide of a benign nevus on each display and gave a single score to indicate their preference in terms of image quality and size of the display.

Results

Thirty-four pathologists took part in the display evaluation experiment. The preferred display was the largest and had the highest technical specifications (11.8MP resolution, 2100 cd/m^2 maximum luminance). The least preferred display had the lowest technical specifications (2.3MP resolution, 300 cd/m^2 maximum luminance). A trend was observed towards an increased preference for displays with increased luminance and resolution.

Conclusions

This experiment demonstrates a preference for large medical grade displays with the high luminance and high resolution. As cost becomes implicated in procurement, significantly less expensive medical grade displays with slightly lower technical specifications may be the most cost-effective option.

Keywords

Digital pathology; display; whole slide image; monitor

Text

Background

For over a decade, digital microscopy has been an essential tool in pathological research. However, the transition to use within the clinical setting for primary diagnosis has been slow, largely due to patient safety concerns. As research into digital pathology for primary diagnosis has increased over the past few years alleviating many concerns, many centres around the world are now aspiring to become digital. Digital pathology is now being viewed as an essential clinical tool for modern pathology services.

The histopathology department within Leeds Teaching Hospitals NHS Trust is in the process of undergoing full adoption of digital pathology across all subspecialties(1). A key issue when undertaking full adoption of digital pathology, is the need to decide appropriate displays for pathologists. Unfortunately, there is little research on the topic, resulting in an absence of published guidelines outlining minimum display standards from relevant government bodies. As far as we are aware, there is only one research paper evaluating different computer displays for primary diagnosis in digital pathology, which is authored by our group(2). This paper concludes that screens with greater resolution speeds up low power assessment of WSIs. In this current paper, as well as in (2), resolution is referred to as the number of vertical and horizontal pixels displayed by a monitor.

In addition, we have conducted a number of experiments in the course of developing the Leeds Virtual Microscope which culminated in the use a 6.7MP Barco Coronis Fusion medical grade display alongside a 3.1MP Barco Nio display (Barco, Kortrijk, Belgium)(3). Despite there being a paucity of research, initial guidance has recently been released from the Royal College of Pathologists, highlighting the importance of the display when reporting digitally as well as the need for pathologists to be aware that displays with differing technical specifications can affect the appearance of the WSI(4).

By contrast, there are extensive guidelines and minimum standards for displays in digital radiology, outlined in the Summary of Guidance from the Institute of Physics and Engineering in Medicine 2005 (IPEM) (5) and Picture Archiving and Communication Systems (PACS) guidance from the Royal College of Radiologists(6,7). Within the subset of physical parameters alone within the IPEM guidelines, there are minimum requirements for; image display monitor condition, greyscale contrast ratio, distance and angle calibration, resolution, greyscale drift, DICOM greyscale calibration, uniformity, variation between monitors and room illumination. Within the PACS guidelines, there are minimum requirements for; screen resolution, screen size, maximum luminance, luminance ratio, greyscale calibration, greyscale bit depth and video display interface.

In the absence of defined minimum standards for displays for primary diagnosis using digital pathology, decisions regarding display purchase are challenging. There are many factors which require due consideration. Firstly, the image quality, encompassing resolution, contrast ratio and luminance needs to be considered. A high-resolution display, e.g. one with a large number of horizontal and vertical pixels will result in a clearer or crisper image containing more detail, than one with a lower number of horizontal and vertical pixels, provided the physical size remaining approximately the same. This can raise confusion if the pixel numbers remain the same, yet the physical size differs, as this alters the dots per inch which is a better measure of resolution. Contrast ratio is defined as a ratio of the darkest color a display can produce (black) and the lightest color (white). A higher contrast ratio display will afford better subtle detail discrimination. Luminance is defined as the luminance intensity per unit area of light travelling in a given angle and may loosely be considered the ‘brightness’ of the display.

Given that many lessons can be learnt from digital radiology(8), the hypothesis that image quality of the display is likely to impact user performance in digital pathology(2) would be reasonable, since this has been found to be true in digital radiology. This has been supported by research in digital radiology that higher display resolution results in greater diagnostic accuracy(9,10). As such, displays for digital radiological diagnosis require a minimum

resolution of 2 mega pixels (MP) for most diagnostic tasks, as stipulated in the IPEM guidelines(5) and guidance from The Royal College of Radiologists(6). The exception being plain film x-ray, necessitating a 3MP display and mammography, which requires the use of 5MP displays with a maximum pixel pitch of 0.17mm. Moreover, the recommendation is that the display matrix size should be as close to the raw image data as possible. If this recommendation were to be extrapolated to digital pathology, then it is arguable that only very high-resolution displays should be considered(11).

Secondly, physical size and logistical positioning of the display within the pre-existing pathologist office should not be overlooked. The challenges with digital pathology are that the image datasets have a higher native resolution than can be physically displayed 1:1, unlike in radiology where the native image size and physical screen size can allow 1:1 pixel display. Having a larger display in digital pathology allows easier magnification and consequently fewer segments required to cover the whole image at a native 1:1 pixel display. Having lower resolution displays will require more panning of the dataset in order to realise the same physical coverage at native resolution. In digital radiology, there are guidelines and minimum requirements of the physical screen size with PACS and the Imaging Informatics Group stating a minimum size of 17” and a recommendation of equal or greater than 20”(7). Of note, there is no upper limit for size of the display in digital radiology.

Thirdly, the cost of the displays is also very important in modern healthcare, where there is growing demand and limited capacity. There is a significant difference in the cost of displays, from very cheap consumer grade desktop displays (approximately £200/ $260) to recently released medical grade displays costing up to £30,000 (approximately $39,300). It is inevitable that cost will be a key factor when deciding which displays to purchase. The issue of cost is also discussed in the Royal College of Radiologists guidance(6), appreciating that medical grade displays are considerably more expensive than their consumer grade alternatives. However, due to inferior lifetime display characteristics with consumer grade displays (increased luminance and contrast ratio deterioration with time), alongside self-calibration and

quality control for the expected display lifetime with medical grade displays, use of consumer grade displays should be “carefully considered” when used for primary diagnosis.

In the absence of defined minimum standards, we aimed to conduct a survey evaluation of short-listed displays by pathologists within our department, to inform the purchase of displays for primary diagnosis in digital pathology within Leeds Teaching Hospitals NHS Trust.

Methods

We performed a thorough search of the available displays for purchase in May 2017. After reviewing the specifications of displays, we chose a short-list of 8 displays covering a range of technical specifications in terms of resolution, luminance and color contrast ratio and physical size. Five displays were medical grade, two were professional grade, and one was consumer grade. Medical grade displays are those which are marketed and manufactured as a medical device and must conform to appropriate guidelines/ approvals e.g. CE medical mark, Medicines and Healthcare Products Regulatory Agency (MHRA), or Food and Drug Administration (FDA)(12). Professional displays are those which are designed for specialist use (e.g. gaming/ photography etc.), and consumer grade are those which are readily available and are designed for general use.

The published technical specifications of the displays short-listed for inclusion in the study can be seen in Table 1.

Monitor	Category	Panel type	Screen size (inches)	Resolution (MP)	Max Luminance (cd/m²)	Contrast Ratio
A	Medical	LCD	31.1	8.8	850	1450:1
B	Professional	LCD	31.1	8.8	350	1500:1
C	Professional	LCD	32	8.3	350	1000:1

D	Medical	LCD	33.6	11.8	2100	1200:1
E	Medical	LCD	30.4	6.7	1050	1500:1
F	Medical	LCD	21.3	5.8	1000	1400:1
G	Consumer	LCD	24	2.3	300	1000:1
H	Medical	OLED	24.5	2.1	275 (measured)	Infinite

Table 1 *– Published technical specifications of the short-listed displays.*

Each display was set-up within the same window-less room, to remove the effect of natural light. Where possible, computer displays were angled to ensure no significant reflections from ceiling lights. Artificial lighting remained constant throughout the experiment, at normal light levels within a pathology department (approximately 300 lux). The display set-up can be seen in Figure 1.

Figure 1 *– The evaluation set-up. Each of the displays were set up within a window-less room with fixed artificial lighting. The same whole slide image was shown on each display for side-by side comparison.*

All vendors were invited to provide guidance on optimal display settings for this evaluation, however, no responses were obtained perhaps due to lack of evidence in this area. Consequently, each display was adjusted in order to optimise the presentation of the SMPTE test pattern(12). An SMPTE pattern was created using bespoke in-house software that can generate the pattern to a specified resolution. For each monitor, the SMPTE was generated to display at the native resolution of the monitor at a 1:1 pixel ratio. Once the displays were adequately adjusted, measurements of luminance, contrast ratio and uniformity were made.

The range of parameters that could be adjusted on each monitor varied significantly between brand, category and price bracket, with some monitors offering a wider range of adjustments than others. The settings that were adjusted included built-in display curve, set luminance, colour temperature, brightness and contrast. The display settings were adjusted to be as consistent as possible by two medical physicists (authors CM and DB) who are experienced in display calibration. The variation of color between displays after optimization can be seen in Figure 2.

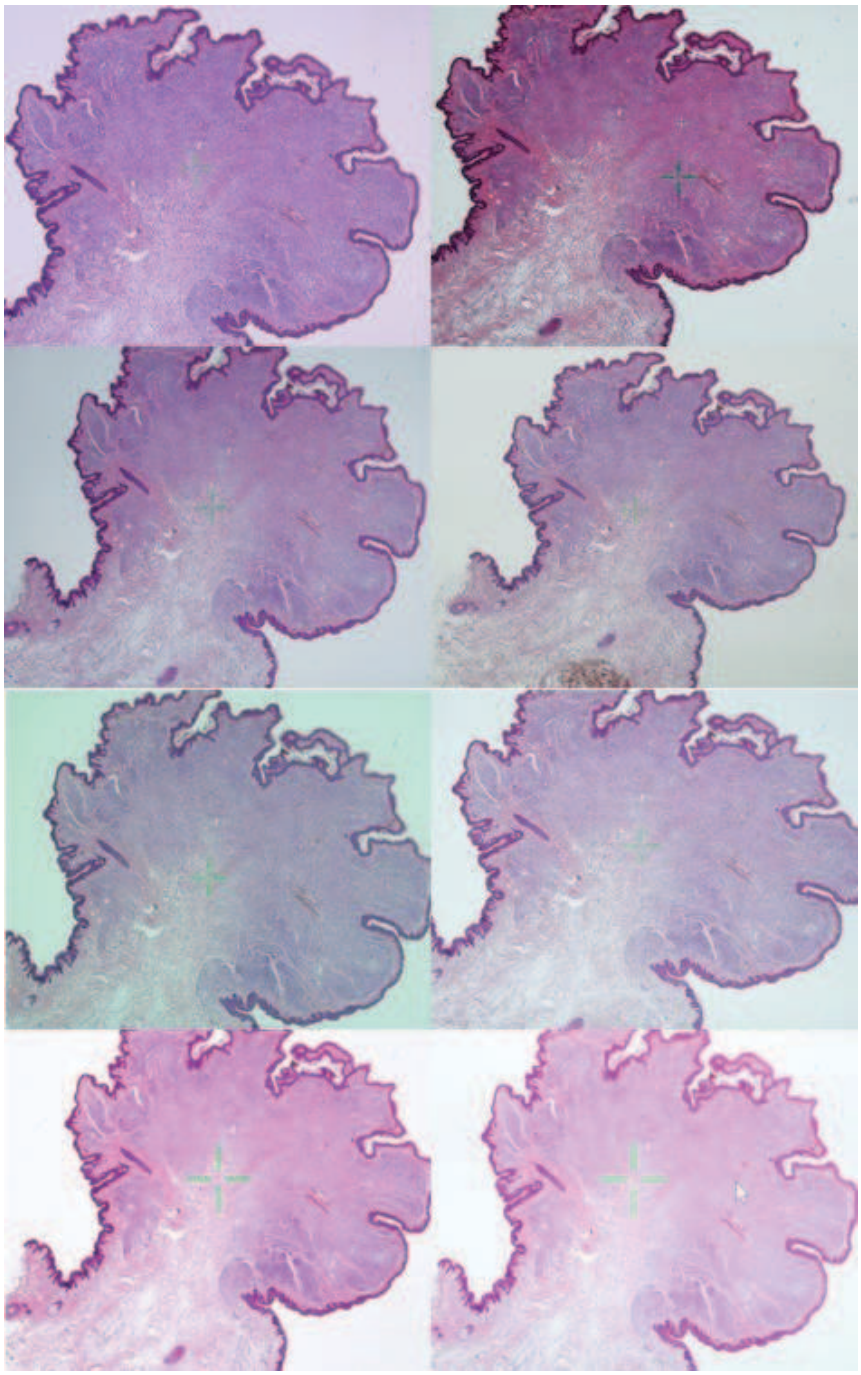

Figure 2 *– The appearance of the WSI of the benign intra-dermal nevus on each of the displays. The top left is Monitor A, top right is Monitor B etc. Despite attempts to standardise appearance of the WSI in terms of color, the ability to do so was limited by the technical specifications of each display. In particular, the color of the WSI of Monitor B was much darker than the other displays. To ensure accurate comparison between images of different displays, these photographs were taken with a fixed International Organisation for Standardisation (ISO), aperture and shutter speed.*

In addition to the SMPTE test pattern, a series of test patterns from the American Association of Physicists in Medicine (AAPM)(13) were used to carry out quantitative measurements of display luminance, contract ratio and display uniformity.

All luminance measurements were taken using a calibrated Unfors Xi lightmeter after all the monitors were given a minimum of 30 minutes to warm up sufficiently.

The TG18-LN01 and TG18-LN18 patterns were used to measure peak black and peak white values respectively, in order to calculate Measured Peak Luminance. The peak white value (measured peak luminance (cd/m^2)) is shown in Table 2. The ratio of the values measured for peak black and peak white provided the contrast ratio of the monitor, also shown in Table 2.

Monitor	Measured Peak Luminance (cd/m^2)	Measured Contrast Ratio	Measured Display Uniformity (10% luminance)
A	624	1299:1	8.7
B	232	1219.1	5.18
C	310	911:1	15.72
D	896	1400:1	6.96
E	670	1290:1	4.2
F	545	1515:1	12.43
G	329	632:1	9.14
H	275	3441:1	2.99

Table 2 *– Measured luminance, contrast ratio and display uniformity using a Unfors Xi light meter, for each of the displays after experiment configuration.*

The TG18-UN10 and TG18-UN80 patterns were used to carry out a subjective appraisal of display uniformity. These were visually assessed to check the displays for any gross artefacts (banding, light bleeding, pixel dropout etc.). None of the displays exhibited any of these artefacts.

A quantitative assessment of display uniformity was carried out by making measurements of the TG18-UNL10 test pattern. This is a uniform greyscale pattern displayed at a 10% luminance level. Measurements were made at the centre and the periphery (5 measurements in total) as indicated on the test pattern and the maximum (L_{max}) and minimum (L_{min}) luminance values are used to calculate the percentage uniformity (U%) using the following formula(5):

$$U(\%) = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \times 200$$

It was noted there was some disparity with the measured values for peak luminance and contrast ratio versus the claimed specifications of the manufacturers. All monitors showed a lower peak luminance compared with the claimed specification except for Monitor G, which surprisingly produced a measured luminance value higher than that specified in the technical documentation (329cd/m^2 vs. 300cd/m^2). Furthermore, two of the monitors on trial (Monitors D & F) produced contrast ratios that exceeded the stated specification. The differences between the claimed and measured values are not entirely unexpected however as often some technical aspects of a monitor's performance are derived in controlled conditions, or using specific display curves which do not necessarily reflect how they would be used in 'real-life' scenarios, or indeed in this trial. For instance, a monitor may achieve a maximum luminance of 1000cd/m^2 in lab conditions but in practice it would not be used in this manner as image contrast would be compromised. Overall it was observed that monitors which claimed to have higher luminance/contrast ratios were shown to have higher measured

luminance/contrast ratios than the lower specification monitors, after they were configured for the experiment by the two medical physicists.

We designed the evaluation to include a quick subjective assessment by pathologists of each of the eight screens according to their preference for image quality and physical size during dedicated 1 hour 'open sessions' at the participant's convenience. We chose the assessment to include one haematoxylin and eosin (H&E) stained whole slide image (WSI) of average quality scanned on a Leica Aperio AT2 (Milton Keynes, UK) digital slide scanner at x20 objective, showing a benign intradermal nevus.

We also designed a separate, second evaluation for selected participants. These participants assessed two H&E stained slides including the benign intradermal naevus (as in the general evaluation), as well as a further H&E stained slide of a micrometastasis of breast ductal carcinoma within an axillary lymph node, one human epidermal growth factor receptor 2 (HER2) haematoxylin -diaminobenzidine (H-DAB) stained slide of control breast tissue, and 1 papanicolaou (PAP) stained slide of a cervical spin showing severe dyskaryosis.

The survey questionnaires used visual analogue scales from 0-100. Each participant was asked to score using a straight line on each of the scales for each of the displays to indicate their preference (0 = the worst possible screen for digital pathology; 100 = the best possible screen for digital pathology). We decided to use the metric of 'professional preference' for this evaluation, as it may be considered a surrogate for the 'fitness for purpose' of a display. There was also the option of writing comments for the study authors. An example of the visual analogue scale used can be seen in Figure 3. Participants were able to pan and zoom as they so desired. Time limits were not imposed.

Figure 3 *– An example of the 10cm Visual Analogue Scale (VAS) used for preference scoring by participants (not to scale). The far-left hand of the scale indicated 'the worst possible screen for diagnostic digital pathology' and the far-right hand of the scale indicated 'the best possible screen for diagnostic digital pathology'. Participants were asked to indicate where on the scale their preference lay for each monitor (separate scale per monitor) by drawing a vertical line. Scores were calculated by measuring with a ruler from the far-left hand side of the scale to the point the participant had drawn.*

All logos on the displays were concealed from the participants and anonymised monitor letters were used in order to remove the effect of prior preference, as can be seen in Figure 1. The users were also not aware of the technical specifications or cost of the displays.

An email was sent to all histopathology trainees and consultants within Leeds Teaching Hospitals NHS Trust, inviting them to take part in the evaluation. Consultant pathologists were those who have completed their specialist training in histopathology in the United Kingdom (5-year minimum duration) and were registered as a fully qualified histopathologist with the Royal College of Pathologists. Trainee pathologists were junior doctors, who were currently undertaking their specialist training in histopathology. The evaluation took place over a period of 1 week in May 2017, over four separate 1 hour 'open sessions' during which participants were able to come and complete the evaluation.

Understandably, there is a wide variation in cost of the displays included in this experiment, with the most expensive display having a published cost of 60 times that of the cheapest display. To evaluate the benefit of the displays with respect to the cost, we conducted a cost: benefit analysis. Given the substantial variation in cost between displays and a preference score of 0-100, we decided to use the cost ranking (1-8) and preference ranking (1-8) in this analysis. We used the following formula to calculate the cost-preference score:

*Cost-preference score = (cost rank * cost rank weighting) + (preference rank * preference rank weighting)*

Ethical approval for this work was obtained from Leeds West LREC 10-H1307-12.

Data were analysed using STATA version 15.1. Significance was set at less than 0.05. Means and 95% confidence intervals are presented. Correlations were performed using Pearson's correlation co-efficient. Means were compared using one-way ANOVA with Bonferroni correction for multiple comparisons. Linear regression was used to estimate associations between technical parameters and score.

Results

A total of thirty-four pathologists took place in the general evaluation of the single H&E slide (21 consultants and 13 trainees). Four participants took part in the selected evaluation (2 pathologists and 2 medical physicists). The medical physicists were expert scientists in medical imaging.

For the H&E evaluation, the mean score for each display is shown in Table 3. Overall, the preferred display was Monitor D (preferred by 22/34 of participants), with a large display size (33.6"), high resolution (11.8 MP), high maximum luminance (2100 cd/m^2) and high contrast ratio (1200:1).

	General evaluation (n=34)			Selected evaluation (n=4)		
Monitor	**All pathologists (n=34)**	**Consultants (n=21)**	**Trainees (n=13)**	**H&E**	**H-DAB**	**PAP**
D	81	82	80	95	82	89
A	68	69	74	82	78	71
C	66	62	72	57	51	53
E	64	61	67	79	70	69
F	55	55	52	62	62	60
H	47	44	47	38	61	56
B	45	39	56	21	25	23
G	41	38	43	34	36	18

Table 3 *– Overall mean scores for each display for both the general and selected evaluation. The general evaluation involved a single H&E slide. The monitors have been ranked according to preference within the general evaluation. The selected evaluation involved two haematoxylin and eosin (H&E) stained slides, one human epidermal growth factor receptor 2 haematoxylin -diaminobenzidine (H-DAB) stained slide of control breast tissue, and 1 papanicolaou stained slide.*

The overall mean scores and associated confidence intervals for all pathologists in the general evaluation can be seen in Figure 4.

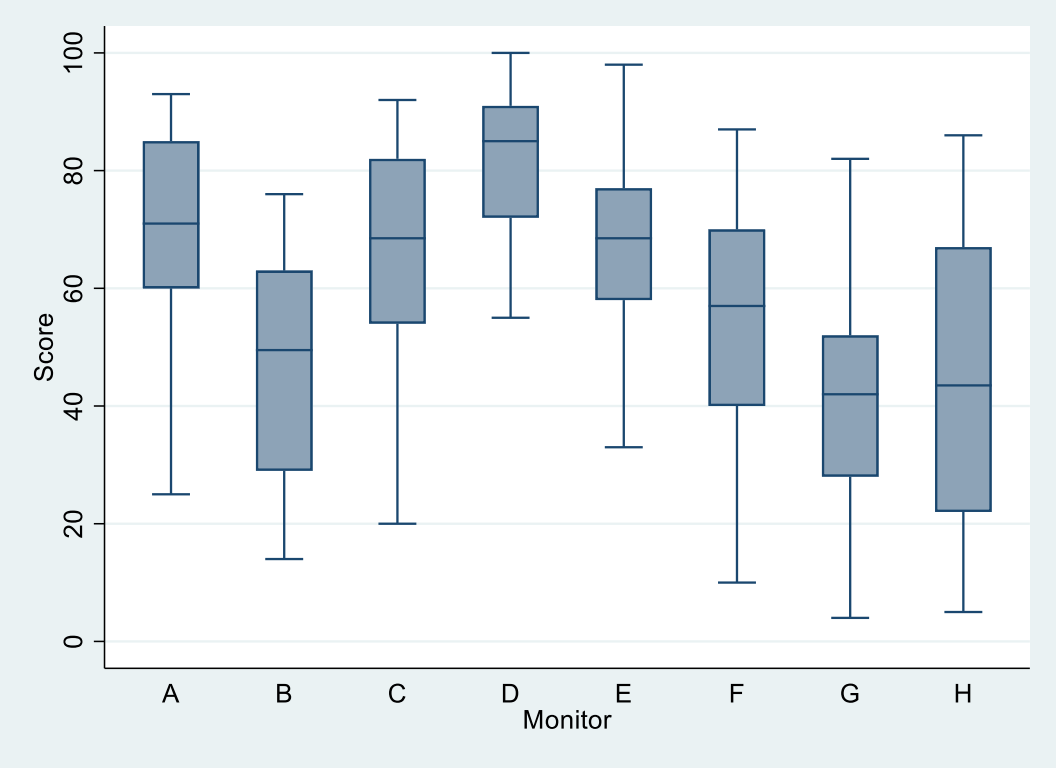

Figure 4 *– Overall scores by display, including all pathologists in the general evaluation.*

There were no statistical differences in preference scores between consultants and trainees, except for Monitor B, which was preferred more by trainees than consultants ($p = 0.028$). The consultants and trainee scores were positively correlated, with a Pearson's correlation co-efficient of 0.904 ($p=0.002$).

Monitor D was also the most preferred display for H-DAB and PAP slides as determined by the selected participants. The least preferred display for H-DAB was Monitor B, whereas Monitor G was the least preferred display for the PAP assessment. The results of the H-DAB and PAP evaluation were positively correlated with the results of the H&E assessment (Pearson's correlation co-efficient 0.923 and 0.915, respectively).

There is a trend for an increase in score as measured peak luminance rises. This can be seen in Figure 5 below, where the monitors have been sorted according to peak luminance.

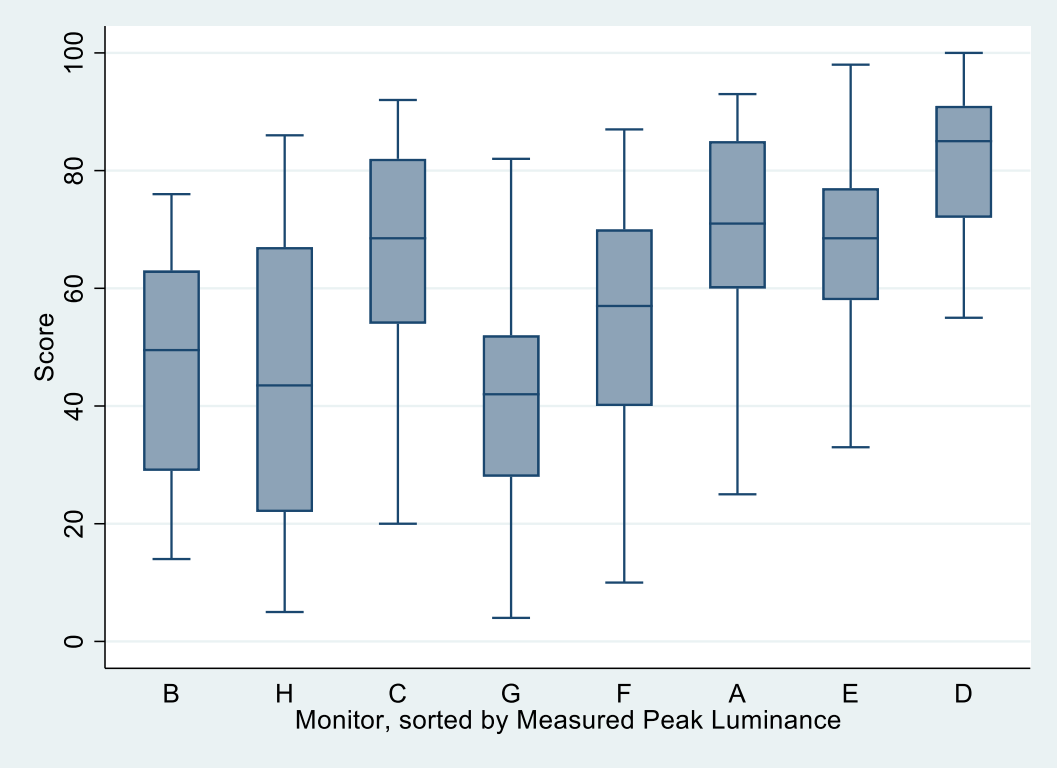

Figure 5 *– Graphical representation of the scores for each monitor, sorted by measured peak luminance, low to high, left to right. The monitor with the lowest peak luminance (Monitor B) is at the far left, whereas the monitor with the highest luminance (Monitor D) is on the far right. This indicates that as measured peak luminance increases, the score generally increases.*

Additionally, as resolution increased, so did the score. This can be seen in Figure 6 below, where the monitors have been sorted according to peak resolution.

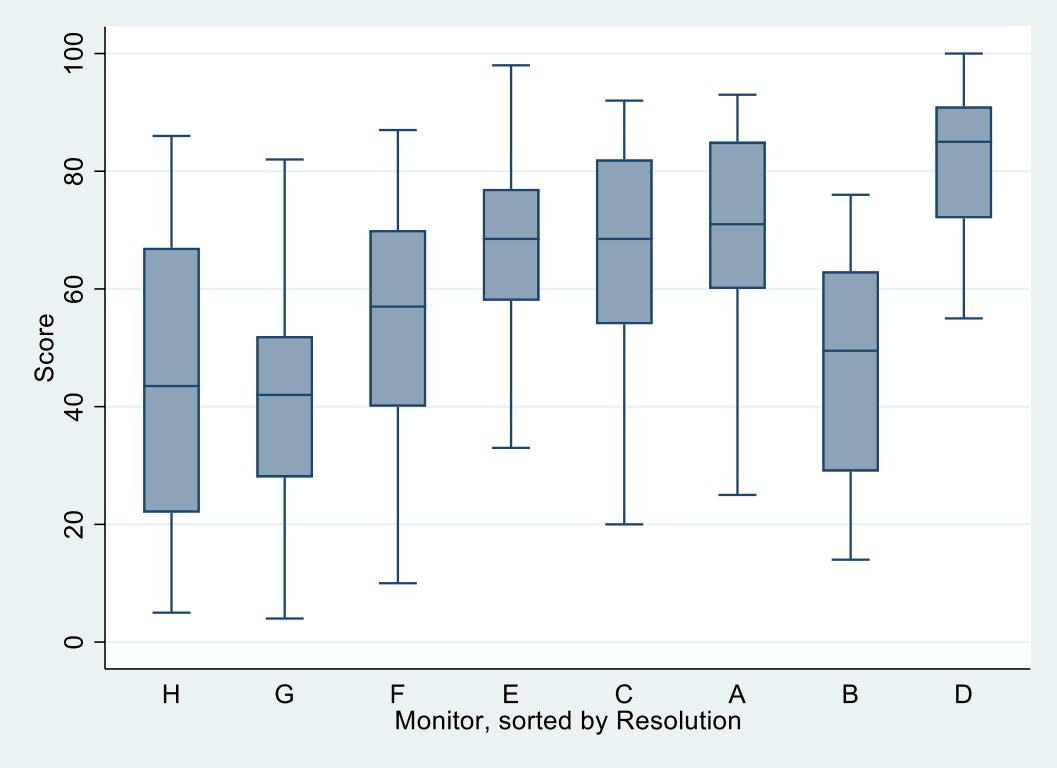

Figure 6 *– Graphical representation of the scores for each monitor, sorted by the reported resolution in megapixels, low to high, left to right. The monitor with the lowest resolution (Monitor H) is at the far left, whereas the monitor with the highest resolution (Monitor D) is on the far right. This indicates that as measured peak resolution increases, the score generally increases.*

In terms of contrast ratio, scores did not seem to increase with increases in contrast ratio, as can be seen in Figure 7.

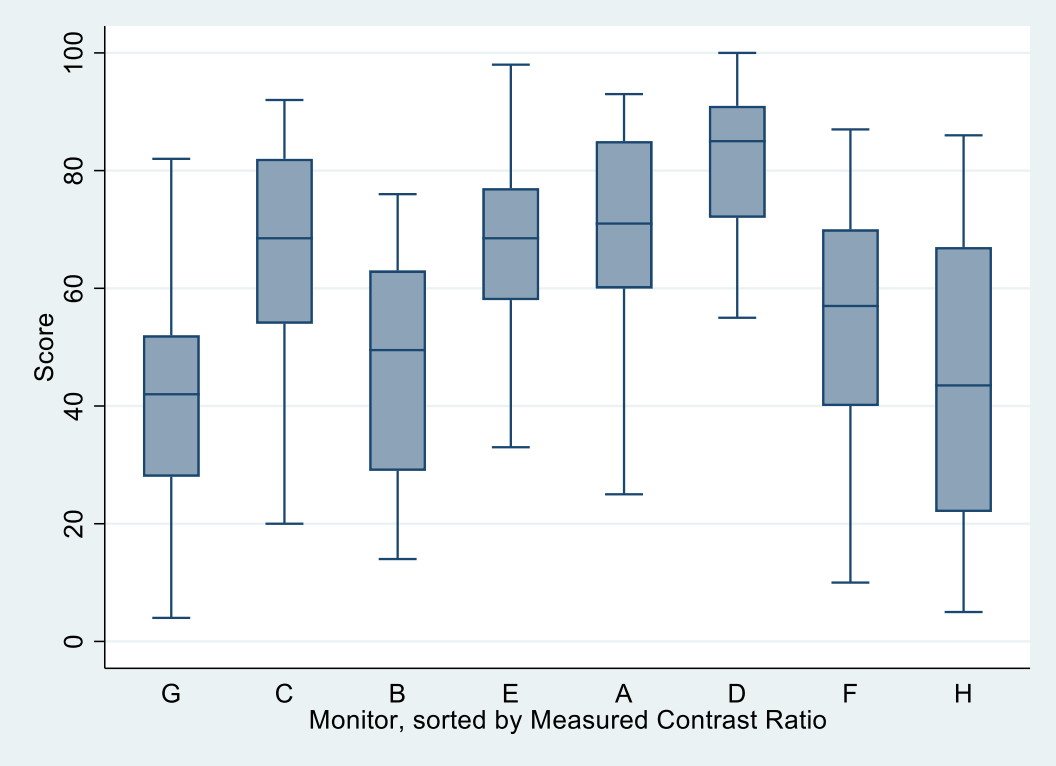

Figure 7 *– Graphical representation of the scores for each monitor, sorted by the measured contrast ratio, low to high, left to right. The monitor with the lowest contrast ratio (Monitor G) is at the far left, whereas the monitor with the highest contrast ratio (Monitor H) is on the far right. This indicates that as measured contrast ratio increases, this does not seem to increase the score.*

The results of the cost:benefit analysis can be seen in Figure 8, with the analyses based on the equation below.

*Weighted score = (cost rank * cost rank weighting) + (preference rank * preference rank weighting)*

The Cost Weighting on the x-axis is a measure of the 'importance of cost', as compared to a fixed preference weighting of 1. At a cost weighting of '0', cost is not implicated in the score and instead reflects pure preference. At a weighting of '1' cost is as important as preference.

At a cost weighting of ‘2’ cost is twice as important as preference. Monitor D only achieves the highest score, when cost is not included in the evaluation. Monitor C becomes the highest scoring display when cost is as important as preference.

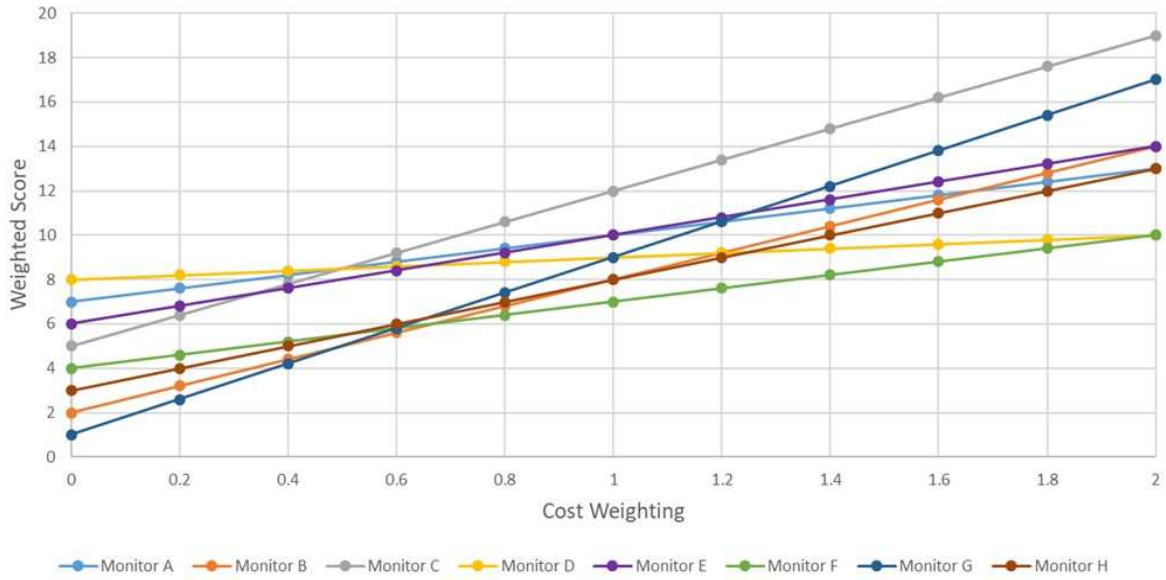

Figure 8 – *‘Importance of Cost’ and its impact on cost-preference score. Cost weightings are varying weightings of cost rank, as compared to a fixed preference rank of 1. A cost weighting of ‘0’ reflects pure preference scores without cost. At a weighting of ‘1’, cost is as important as preference. Weightings >1 are relative cost weightings as compared to a preference of 1, i.e. when cost becomes more important than preference, up to a maximum of cost being twice as important as cost.*

Discussion

As pathology departments strive to digitise for primary diagnosis, display choice becomes important. The absence of defined minimum display standards may have resulted in departments who have already ‘gone digital’ purchasing displays possibly without a full appreciation for the impact of display choice on the end user.

Philips was the first vendor to achieve FDA approval in 2018, which stipulates the use of a 4MP medical grade Barco display(14). Arguably, this appreciation by FDA of the influence of the display on primary diagnosis, highlights the need for minimum requirements, in a similar fashion to the detailed minimum display requirements in digital radiology.

As far as we are aware, this experiment represents the first attempt in digital microscopy to evaluate displays of varying specification on user preference. We were unsurprised that there was a substantial difference in user preference between displays of differing specifications. Despite users being blinded to make, model, technical specification and cost of the displays, Monitor D, the most expensive display with the highest technical specifications, was the most preferred. Additionally, the display that scored the lowest for preference for H&E and PAP (Monitor G), also had the lowest technical specifications in terms of luminance and contrast ratio and second lowest resolution. This display was also the least expensive display.

Sorting the monitors and their scores by their technical specifications, indicate that both increasing resolution and increasing luminance seem to increase the preference. Surprisingly, an increase in contrast ratio did not seem to increase the preference; Monitor H, with an organic light emitting diode (OLED) panel and therefore infinite contrast ratio (OLED pixels emit light directly and therefore individual pixels can be completely turned off and emit no light at all), did not perform as well as expected. We had anticipated that this display's high contrast ratio would have resulted in a substantial improvement on user preference, due to the improved ability to distinguish subtle color differences. However, it appears that contrast ratio does not compensate for low resolution. This is further supported by Monitor C which has a lower contrast ratio than Monitor B but otherwise similar specifications, scoring a much better preference score.

Twelve users voiced their preference for the larger displays (>30"), with no users highlighting their preference for the smaller screens (<30"). However, 2 participants did suggest that Monitor D (33.6") was too large with the most preferable size being that of Monitor A or E (31.1"). This preference for larger displays over the smaller displays is in accordance with

PACS and Imaging Informatics Group recommendations(6) which do not provide an upper limit for display size.

The main strengths of the work are with regards to the stringent methodology to try and minimise the effect of numerous confounders on the results of this experiment. This is particularly with regards to the set-up of the environment and displays, blinding of the participants and the choice of participant task. A further strength was the number of participants in an experiment of this type. It should also be possible to test new monitors and add their results to the portfolio already created by replicating the methodology.

However, there are limitations to this work. It is important to remember that preference for a display may not neatly translate into improved clinical performance; whether in terms of speed of diagnosis, accuracy of diagnosis or less user fatigue/ eye strain. Therefore, future work will involve evaluation of displays with respect to their quantitative impact on user performance.

For ease of participant involvement, we asked most participants to evaluate only one H&E slide and the selected participants only evaluated one PAP and one H-DAB slide. The preference scores may be influenced by the choice of slides and their specific stain properties.

Finally, it was not possible to truly isolate the effect of the display size from image quality; ideally it would be preferable to vary each independently from the other to fully appreciate the influence of each variable on preference.

Conclusions

To conclude, we have shown that pathologists demonstrate a preference for medical grade displays with the highest technical specifications, with a trend towards an increased preference for displays with increased luminance and resolution. As cost becomes implicated in the decision over display procurement, medical grade displays with a slightly lower price point become preferable.

We hypothesise, that most cases could be diagnosed using any display. However, there will be specific, challenging cases (e.g. assessment of dysplasia or finding small objects such as a micro-metastasis) that high technical specifications (particularly high-resolution displays) will prove advantageous in terms of user performance.

Disclosures/ Acknowledgements

Disclosures

Author Darren Treanor (DT) is on the advisory board of Sectra and Leica/ Aperio. He receives no personal remuneration for these boards. DT has had a collaborative research project with FFEI, where technical staff were funded by them. He received no personal remuneration. DT is a co-inventor on a digital pathology patent which has been assigned to Roche-Ventana on behalf of his employer. He received no personal remuneration.

Acknowledgements/ Funding Sources

We would like to take this opportunity to thank all the vendors for taking part in this evaluation, without whom this work would not have been possible.

We would also like to thank our pathologist colleagues at Leeds Teaching Hospitals NHS Trust for giving up their valuable time to help us in this assessment. We are also very appreciative of the help of technical support staff from the University of Leeds and Leeds Teaching Hospitals, particularly Dave Turner, Steve Toms and Michael Todd.

The authors of this study are members of the Northern Pathology Imaging Co-operative, NPIC (Project no. 104687) which is supported by a £50m investment from the Data to Early Diagnosis and Precision Medicine strand of the government's Industrial Strategy Challenge Fund, managed and delivered by UK Research and Innovation (UKRI). However no specific funding was received for this study.

Authors' Contributions

Authors EC, DT and DB designed the research study.

Authors EC, CM and BW performed the research.

Author EC analysed the data.

Author EC wrote the manuscript which was revised by all the authors.

Author Biographies

Dr Emily Clarke is a Histopathology Registrar and Clinical Research Training Fellow, based at the University of Leeds and Leeds Teaching Hospitals NHS Trust. She has specific research interests in digital pathology and dermatopathology.

Mr Craig Munnings is a Specialist Technical Officer within the Medical Physics Department of Leeds Teaching Hospitals NHS Trust.

Dr Bethany Jill Williams is a Histopathology registrar, PhD Fellow in Digital Pathology and lead for digital pathology training and validation at Leeds Teaching Hospitals NHS Trust and the Northern Pathology Imaging Co-Operative.

Professor David Brettle is the Head of Medical Physics at Leeds Teaching Hospitals NHS Trust. His research interests are concerned with optimising imaging technologies for health care provision. He is an honorary professor at the University of Salford, UK and a visiting senior research fellow at the University of Leeds, UK

Dr Darren Treanor is a Consultant Histopathologist and Honorary Clinical Associate Professor at Leeds Teaching Hospitals NHS Trust and the University of Leeds. He is also a Guest Professor at the University of Linköping, Sweden.

References

1. Leica Biosystems. Leica Biosystems and Leeds Hospital establish strategic partnership to provide quantifiable benefits of digital pathology [Internet]. 2017. Available from: https://www.leicabiosystems.com/news-events/news-details/article/leica-biosystems-and-leeds-hospital-establish-strategic-partnership-to-provide-quantifiable-benefits/News/detail/

2. Randell R, Ambepitiya T, Mello-Thoms C, Ruddle R, Brettle D, Thomas R, et al. Effect of Display Resolution on Time to Diagnosis with Virtual Pathology Slides in a Systematic Search Task. J Digit Imaging. 2015;28(1):68–76.

3. Randell R, Ruddle RA, Thomas RG, Mello-Thoms C, Treanor D. Diagnosis of major cancer resection specimens with virtual slides: impact of a novel digital pathology workstation. Hum Pathol. 2014;45(10):2101–6.

4. Cross S, Furness P, Igali L, Snead D, Treanor D. Best practice recommendations for implementing digital pathology Unique document number G162 Document name Best practice recommendations for implementing digital pathology. 2018;1–38. Available from: https://www.rcpath.org/uploads/assets/uploaded/d6b14330-a8b9-4f5e-bbe443f0d56de24a.pdf

5. Institute of Physics and Engineering in Medicine. Report 91 - Recommended standards for the Routine Performance Testing of Diagnostic X-Ray Imaging Systems. 2005.

6. The Royal College of Radiologists. Picture archiving and communication systems (PACS) and guidelines on diagnostic display devices Third edition. 2019;(February). Available from: https://www.rcr.ac.uk/system/files/publication/field_publication_files/bfcr192_pacs-

diagnostic-display.pdf

7. The Royal College of Radiologists. Picture archiving and communication systems (PACS) and guidelines on diagnostic display devices Second edition Board of the Faculty of Clinical Radiology. R Coll Radiol. 2012;

8. Krupsinski E. Virtual slide telepathology workstation of the future: lessons learned from teleradiology. Hum Pathol. 2009;40(8):1100–11.

9. Krupinski EA. Medical grade vs off-the-shelf color displays: Influence on observer performance and visual search. J Digit Imaging. 2009;22(4):363–8.

10. Chen Y, James J, Turnbull A, Gale A. The use of lower resolution viewing devices for mammographic interpretation: implications for education and training. Eur Radiol. 2015;25(10):3003–8.

11. Krupinski E a, Kallergi M. Choosing a radiology workstation: technical and clinical considerations. Radiology. 2007;242(3):671–82.

12. U.S. Department of Health and Human Services. Display Devices for Diagnostic Radiology Guidance for Industry andFood and Drug Administration Staff. Food Drug Adm. 2017;1–14.

13. American Association of Physicists in Medicine. Display Quality Assurance [Internet]. 2019. Available from: https://www.aapm.org/pubs/reports/RPT_270.pdf

14. FDA. FDA allows marketing of first whole slide imaging system for digital pathology [Internet]. 2017. Available from: https://www.fda.gov/news-events/press-announcements/fda-allows-marketing-first-whole-slide-imaging-system-digital-pathology