

Received Date : 02-Jun-2016

Revised Date : 30-Aug-2016

Accepted Date : 02-Sep-2016

Article type : Review

Colour in Digital Pathology: A Review

Authors

Emily L Clarke^{1,2}, Darren Treanor^{1,2}

Institutions

1. Department of Histopathology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
2. Section of Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK

Corresponding Author

Dr Emily Clarke

Leeds Institute of Cancer and Pathology

University of Leeds

Beckett Street

Leeds, UK

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/his.13079

This article is protected by copyright. All rights reserved.

LS9 7TF

Tel +44(0)1133 438509

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

Running Title: Colour calibration in digital microscopy

Keywords: Colour calibration; colour standardisation; digital pathology; whole slide imaging

Conflict of Interest Statement: We declare the following interests: Author 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 will receive no personal remuneration.

Introduction

Colour is central and integral to the practice of pathology; pathologists use coloured histochemical and immunohistochemical stains to identify structures and reach diagnoses. Given the magnitude of these diagnoses to patient management and outcome, it is imperative that pathologists are able to make accurate and reliable assertions.

Over recent years, new technology has emerged to enable pathologists to carry out their diagnostic work digitally, rather than with the analogue optical microscope. This technology is often referred to as a 'digital microscope' in which whole slide images (WSIs) of tissue slides are scanned, transmitted and displayed on a computer monitor. Whilst the method of making diagnoses may be gradually evolving, the need for diagnostic reliability remains a constant. Concerns that WSI may not be an accurate representation of the pathology, may, in part explain why digital pathology has not yet become part of routine clinical practice. To try to resolve these issues, there has been a recent surge in research in the field of WSI validation (1), however the impact of color remains unknown.

The aim of this review is to provide a summary of the issues surrounding colour control in digital microscopy for pathologists. This necessitates an initial brief introduction to the physics of colour and its interpretation by the human visual system, but much of the vast field of colour science is beyond the scope of this work.

Colour Basics

Colour is 'an attribute of things that results from the light they reflect, transmit, or emit in so far as this light causes a visual sensation that depends on its wavelengths' (2), in that light with different wavelengths are perceived as different colours. Those colours that are formed by a single wavelength of light, are termed 'monochromatic colours' (e.g. red, orange, violet, green, blue, yellow). These make up our colour spectrum including wavelengths from 390nm to 700nm. Light with wavelengths outside of this spectrum are not visible to humans and therefore do not form part of

our colour spectrum. These include infrared with wavelengths greater than 1000 nm and ultraviolet with wavelengths less than 300nm.

Figure 1 – The spectrum of light

Source:

https://upload.wikimedia.org/wikipedia/commons/2/22/Spectrum_of_light.png

The colour of an object is the result of our interpretation of the object's surface, transmission and emission properties. There are three main ways in which objects can affect light. Firstly, opaque objects reflect light; they can do so 'specularly' like a mirror, or 'scattered' with diffuse scattering, which eludes to a 'roughened' surface. Secondly, objects that transmit light appear transparent (no scattering) or translucent (with scattering). Thirdly, those that emit light have excited electrons secondary to a chemical reaction (chemoilluminescence), an elevated temperature (incandescence), or from absorbing light at other frequencies (fluorescence) (3,4). Pathology tissue slides are translucent when being viewed using light microscopy – in other words, the colours they exhibit are due to transmission of light with scattering.

In 1730, Newton demonstrated that white light, for example from the sun, can be dispersed with a prism into all visible wavelengths. This is a perfect demonstration of additive colour mixing; colours can be added together to make white. Additive colours are produced by objects which themselves create light; for example, computer monitors, rainbows, fireworks. By contrast, subtractive colours combine to make black and are produced by objects that do not emit light, e.g. colour printing, photographs or fine art. It therefore follows that glass microscope slides are

comprised of subtractive colours, whereas the digital microscope displays additive colours. It is important to ensure that diagnostic information contained within the specimen is unaffected by the conversion from a subtractive colour system to an additive colour system (5).

Figure 2 – Demonstration of additive and subtractive colour mixing. Representation of additive colours are generally portrayed using the RGB colourspace. Representation of subtractive colours are normally demonstrated using the CMYK colourspace.

Source: <https://commons.wikimedia.org/wiki/File:Subtractive-Additive-Colour-Mixing.jpg>

In order to describe, organise and categorise colours we can employ the use of a colourspace. There are various colourspaces, with one of the commonest being the HSV colour space (Figure 3) which comprises three qualities for each colour; hue, saturation and value. Each colour will have a defined value in each of these three qualities and therefore can be accurately measured and replicated. A colour's 'hue' is the main discernible attribute of a colour e.g. red. 'Saturation' is the intensity of a colour and 'value' is the lightness of a colour. An alternative to the HSV colourspace, is the RGB (Red, Green, Blue) colourspace (Figure 2). In this colourspace, instead of each colour being defined by hue, value and saturation, each is given a value for red, green and blue (normally from 0 to 256 in an 8 bit-display) (5).

Figure 3 – Graphical representation of the HSV colourspace. Assigning a value to each of the qualities, enables accurate categorisation and replication of each colour.

Source:

https://upload.wikimedia.org/wikipedia/commons/0/0d/HSV_color_solid_cylinder_alpha_lowgamma.png

Colour and the Human Visual System

As previously highlighted, the colour of an object is significantly influenced by human perception and interpretation; indeed, it is likely that individuals view and interpret colour differently (6). As a consequence, the role of colour in medical image interpretation is a very complex topic spanning physics, engineering, physiology and psychology (7,8). However, a brief summary of how the human visual system interprets images is provided below.

The human eye can identify up to 10 million different colours (8), yet the sensitivity for some hues and saturation is stronger than others. This was demonstrated by MacAdam's ellipses in the CIELAB colourspace back in 1942. His psycho-physical experiments into Just Noticeable Differences (JNDs) highlighted that the Human Visual System (HVS) appears more sensitive to changes in purple than other colours (5,9). It is therefore perhaps fortuitous (and perhaps no coincidence) that the most common stain, haematoxylin and eosin, is purple and pink.

Whilst the HVS very sensitive to change to hue and saturation, it is very adaptive to brightness. It can adapt to 14 decades (orders of magnitude) of brightness, making use of two mechanisms; light adaption and dark adaption. Light adaption occurs when entering a brightly lit environment from a dark one, and takes around 5 minutes

to lower sensitivity to illumination levels. Conversely, dark adaption takes around 30 minutes, increases visual sensitivity, and enables some degree of night vision. It is this change in illumination sensitivity that results in a lit candle appearing to be much brighter in a dark room than within a brightly lit environment.

The HVS is also capable of chromatic adaptation. This is the biological equivalence of “automatic white balance”, whereby the surrounding illumination will have an effect on the perceived colour (10). This can be observed by examining a white piece of paper under many different illuminations e.g. in daylight, fluorescent and incandescent and observing the difference. This adaptation allows a phenomenon called ‘colour constancy’ to occur – this is the way in which the brain adjusts perceived colours based on their surroundings. For example, a banana looks yellow on a bright sunny day and yellow during a candlelit meal, even though the illumination is significantly different (11). Conversely, a famous example of this phenomenon is the controversy of the ‘black and blue dress’ photo, in which 30% of the population view the colour as white and gold (Figure 4).

Figure 4 – The ‘black and blue’ dress photograph. It is seen by 30% population as white and gold due to differences in interpretation by the Human Visual System.

Source: <https://i.ytimg.com/vi/AEz9wQVHiYA/hqdefault.jpg>

This photograph accentuates the effect of the surroundings on the perceived colours and so does not afford colour constancy. Instead, some people attribute the ‘blue’ within the dress as due to the surroundings and illumination of the photo, thereby seeing white and gold, whereas others assign the ‘blue’ to the dress itself and therefore see the dress as blue and black (12). Beyond artificial chromatic

adaptation, a recent study has provided evidence of this occurring in the natural setting. Perception of colour changes with the seasons; during the summer when we are surrounded by 'green', we compensate for this by shifting our 'unique yellow' settings (13). It is currently unknown how chromatic adaptation influences the perceived colour of whole slide images, but in order to attempt to minimise variation of interpretation, it may be advisable to standardise the illuminance of digital microscope scanners and displays.

Colour appearance is also affected by various visual cognitive mechanisms including, 'memory colour' (14). Memory colour is a phenomenon whereby people associate a particular colour with a recognizable object. However, interestingly, their 'memory' of the colour is often not based in reality as people often remember colours as more saturated than in actuality. This may have a bearing on pathologists' ability to accurately compare colours from glass slides with the colours of WSIs.

The HVS also changes as part of the natural aging process, which acts to decrease colour vision. The cones within the retina (the cells responsible for colour vision) decline in sensitivity, rendering images less bright with reduced colour discrimination. Most commonly, blue colours appear more 'washed out' than others (15), resulting in the majority of colour defects in the elderly population being of blue-yellow type (16). Whilst the effect of colour-blindness amongst pathologists has been evaluated in a handful of studies (17,18), the effect of aging on interpreting slides is unknown, and is likely to be more complex than simple physical visual changes.

So Does Colour Matter in Pathology?

This question is hotly debated, primarily due to a lack of definitive evidence either way. It is understandable that given our ability to adapt well to variations in colour, that some conclude that colour does not matter in the field of digital pathology. However, evidence of adaptation is not a reason against control of colour variation, since this argument overlooks key points. Firstly, the implications of adaptation to colour variation have not been evaluated, and so lack of colour standardisation may be a hindrance to diagnostic process. Secondly, whilst we are certainly adept at dealing with variation in colour, our ability to adapt has not been well studied or quantified and as such, we may only be able to compensate for colour variation to a certain degree before it begins to impact diagnosis.

A further argument against the need for accurate colour representation in digital pathology is the existence of colour blind pathologists. This idea is reinforced by a recent small study by Campbell et al 2015 (19) indicating a 92.7% concordance between whole slide images converted to grayscale and the original glass diagnoses in breast biopsy cases. Whilst this may seem a logical argument, a study by Poole et al 1997 (18) indicated that colour blind pathologists had a lower mean score (94% vs. 99%) as compared to their colleagues with normal colour vision when trying to identify pathological features. This finding is supported – amusingly - by a recent paper demonstrating that pigeons (*Columba livia*) had a reduced accuracy for detecting breast cancer in using monochrome images as compared to full colour (20). Furthermore, a survey of pathologists by Akman et al 2015 (17) reported that 61.3% of pathologists interviewed felt that colour blind pathologists should not perform pathology – thereby indicating that many pathologists do feel that colour interpretation is important in the diagnostic process.

A further argument for the importance of colour in digital pathology, is the common laboratory practice of re-cutting and re-staining referred to as 'foreign' slides indicating that, at the very least, pathologists prefer working with familiar colours and possibly that this may impact on their ability, confidence or speed of diagnosis. This perceived preference for certain colours amongst pathologists has also been demonstrated in our experience with WSIs after a recent EQA raised concerns from pathologists that stain irregularities would compromise their diagnostic ability (21). Also, few pathologists would argue that colour is non-essential in the realm of immunostains.

Regardless of this divide in professional opinion, in the field of colour science there is agreement that colour control is a necessary step in digital imaging, and the US Food and Drug Administration (FDA) (22) have released guidance stating that digital microscope images should be displayed in a consistent and reliable fashion.

Colour Variation in Pathology

Unfortunately, colour variation is substantial in pathology; this has anecdotally been acknowledged for some time, but has only recently been formally quantified. Disparities in routine H&E staining was evaluated by Gray et al (23)(24), through staining the same piece of tissue, scanning it into a digital slide scanner and performing image deconvolution. Whilst there was high repeatability in H&E ratio when staining on the same day (mean difference 0.47%), the H&E ratio varied considerably when stained on different days (mean difference 8.32%).

Figure 5 – A selection of slides from Gray et al 2015 (24) demonstrating routine wide variation in staining. These four different slides of the same piece of appendix were stained on different days. The wide range of different colour stains in routine practice is clearly evident.

Unfortunately, even the advent of automated staining does not resolve the issue – further work has demonstrated H&E ratio differences of over 100% between four different automated staining instruments from the same manufacturer (24)(23). Despite these known inconsistencies, anecdotally it seems pathologists are able to make successful diagnoses from slides with a wide variety of staining variation and using a range of optical microscopes.

Colour consistency is much more of an issue with the introduction of the digital microscope, since it does not enable the pathologist to view the tissue directly and so colour variation is not confined to irregularities within the staining process. Therefore, digitisation of the slide introduces further lack of colour control compounding the variation already introduced through the slide staining process.

There are multiple stages involved in making a digital image, each of which may substantially influence slide colour. These stages include; sample illumination, magnification, image capture, compression, storage, transmission and finally reproduction on the computer display (Figure 6). To the best of our knowledge, the effect on the slide colour from each of these individual steps in the digitisation process has not been quantified. However, the overall effect of colour variation due to scanning the slide was investigated by Gray et al (24), with a mean difference of 7% in H&E ratio when scanning the same slide into the differing scanners on the same day. It should be noted though, that this measurement does not take into

account the effect of colour variation due to the display, so the overall impact of digitisation on colour variation from scanning to display is likely to be greater.

Figure 6 – The stages required in capturing and displaying a digital slide. Each of the stages shown have the capacity to substantially influence digital slide colour.

Colour constancy in digital imaging is measured in a unit called Delta-E (dE) introduced by the International Commission on Illumination (CIE). This is a number quantifying the change in visual perception between two colours. The smallest perceptual change in colour is generally regarded as 1 dE, with 2-10 dE being perceptible at a glance, 11-49 dE colours are more similar than opposite and 100 dE representing exact opposite colours (25). A study by Shrestha et al 2014 (26), indicated an average colour discrepancy of 10 dE in uncalibrated scanners alone, so it is easy to imagine the wide variation in colours that are derived from inconsistencies at every stage in digitisation of slides.

The effect of colour differences on human interpretation of digitised pathology images has not been widely studied, however pathologists themselves have raised concerns that colour variability may negatively impact on their diagnostic performance. This was formally noted in feedback from two recent national external quality assurance programmes (EQA) in the UK. In one EQA, 14 of 84 comments expressed discontent with the WSI colour and 4 participants expressly mentioned concern that the variation in colour may negatively impact their diagnostic performance (21). However, whether subjective comments reflect diagnostic ability remains unknown (27).

Differences in slide colour may also have serious implications for the reproducibility of image analysis algorithms. There are numerous papers suggesting algorithms for digital analysis of immunostains (28–32), but only a handful consider the huge implications for colour variation in WSI (33,34). Gavrielides et al (35) reported variation in colour between three different WSI instruments from two manufacturers leading to variation in performance of image analysis algorithm for HER2/neu. Approximately 20-30% of cases scored as 2+ expression on one scanner were re-scored to a different class on another instrument.

Colour Calibration

Colour calibration is an imaging process which seeks to match colours between two or more devices. End-to-end calibration describes the process of controlling colour from source to output through each step of the imaging pathway. Colour calibration is a routine process in the print and photography industry and has been adopted by most digital systems. Even smart phones can be colour calibrated since the advent of the SpyderGALLERY app from Datacolour (36).

Colour calibration works by comparing ‘known’ colours from a set of colour patches, with the colours of those same patches when an image is taken of them with the device. The differences between the ‘known’ colour values for each of the patches and the colour values acquired through imaging the patches allow for numerical identification of the deficiencies within that specific imaging system and allow for necessary adjustments to be made. Compensation for those differences is afforded with the use of an ICC (International Colour Consortium) colour profile, which can then be used to calibrate subsequent images.

It is important that the coloured patches are representative of the colours encountered by the device, otherwise colour calibration may not be accurate. Within photography, digital photographs are often calibrated using a Macbeth colour checker (Figure 7) (37), which includes 'memory colours'. Memory colours are patches of critical colours often encountered in photographs that would result in a very objectionable outcome if incorrectly captured e.g. sky or grass.

Figure 7 – A Macbeth colour checker. This device is primarily used to colour calibrate digital cameras. It includes patches for grayscale values, a range of different colour hues (saturated and pastel) and a selection of important 'memory colours' such as the colour of the sky.

Source:

<http://www.xrite.com/ResourceRoom/category.aspx?CategoryID=16&PartID=1257>

Colour Calibration in Pathology

As previously mentioned, the creation of digital slides involves many stages (Figure 6), and so the process of colour calibration in digital pathology is broadly categorised into two main areas: internal and external colour calibration. Internal colour calibration involves standardisation and correction of the scanning process itself, whilst external colour calibration is concerned with standardisation of the display, accounting for the monitor's effect on perceived colour as well as the viewing environment.

External colour calibration is more straightforward as a spectrophotometer or colorimeter (Figure 8) is used to externally colour calibrate through measuring the appearance of colours of the display, drawing comparisons with the original image and making necessary adjustments with the use of a monitor ICC profile.

Figure 8 – Xrite i1Display Pro Colorimeter (A), a simple spectrophotometer in (B). Colorimeters, like in A are small devices, which use red, green and blue filters to measure the absorbance of different colours (38). The colorimeter is hung in front of the screen, while colour patches are displayed on specialist software for use with the colorimeter. The colorimeter measures the colour of the patch as it is displayed and this is compared to the 'known' colour for that patch. By displaying numerous patches, a colour profile for the display can be created, taking into account the deficiencies of the display. Spectrophotometers (B) (39), operate slightly differently to colorimeters; they have a self-contained light source used to measure the spectral data (transmittance/ reflectance). They have a wide range of functions, including colorimetry.

Source: <https://www.flickr.com/photos/seeminglee/8287570888>

https://upload.wikimedia.org/wikipedia/commons/0/0d/Spectrophotometer_small_for_8_samples-03.jpg

Only a couple of studies have investigated the clinical impact of external colour calibration and variation of the display characteristics on clinical performance in WSI. Krupinski et al 2012 (40) compared a colour calibrated monitor with one that was uncalibrated and demonstrated that whilst there was no benefit in colour calibration

in terms of diagnostic accuracy, there was a statistically significant improvement in time to diagnosis (mean time to diagnosis calibrated = 4.895 seconds vs. uncalibrated = 6.304 seconds, $p = 0.0460$). Similarly, a small series of experiments by a vendor (41) has indicated that colour and luminance stability increases diagnostic accuracy and inter-pathologist agreement, whilst decreasing reading time. However, Hanna et al 2015 (42) did not demonstrate an effect of display colour calibration on diagnostic accuracy.

Internal colour calibration however presents more difficulties. A review of the issues surrounding colour consistency in digital pathology was discussed at the Summit on Colour in Medical Imaging in May 2013 (43), which highlighted the need for an established 'Gold Standard' in slide colour. It was proposed that since the optical microscope is the gold standard viewing modality, virtual slides should mimic the colour of the glass slide as viewed down the microscope. Subsequently, the FDA have recommended in their recent guidance (22) that 'colour calibration should be calibrated with a target slide. The test object should contain a set of measurable and representative colour patches', similar to the Macbeth colour chart used in photography and ideally have similar spectral characteristics to stained tissue.

To meet this need, Bautista et al 2014 (44) made a 9-patch colour calibration slide made with plastic colour filters and demonstrated a statistically significant reduction in CIELAB variation by 3.42 units between WSIs produced from different scanners of the same tissue slide. A further study by Wei et al. 2014 (45) presented an alternative colour calibration slide for H&E stain only, with promising laboratory evaluation results. Leica Biosystems (46) have adopted a different approach through creating an ICC profile for use with their AT2 scanners using stain transmittance spectra from published literature, without scanning a colour calibration target. The

effect of this colour calibration profile can be activated or deactivated using their 'colour management' button within ImageScope® (Figure 9).

Figure 9 –A & B are the same case of a tubulovillous adenoma with high grade dysplasia and pseudoinvasion as part of the UK Bowel Cancer Screening Programme using the same desktop display. A is viewed using ImageScope® from Leica Biosystems (46) without colour management applied and B is with colour management using Spectrum Webscope® also from Leica Biosystems (46). Some participants were concerned that the images were 'too dark', 'too blue' or 'too intensely stained', which suggests that they may not have been using colour managed slides to complete the EQA.

Sources:

[http://www.virtualpathology.leeds.ac.uk/eqa/specialist/nbcs/bcsp/bcspcirs.php?circ=N%20\(b15a\)](http://www.virtualpathology.leeds.ac.uk/eqa/specialist/nbcs/bcsp/bcspcirs.php?circ=N%20(b15a))

Building upon this work, our group developed a unique colour calibration test object (Figure 10). By using histochemical stains and a tissue mimicking substrate, our target is able to provide a more accurate colour representation than film based targets with an estimated 60% reduction in colour error, and can be used with scanners from any manufacturer. Pilot work evaluating its effectiveness in the clinical setting, has indicated that colour calibration of virtual slides is preferred by pathologists and results in improved diagnostic confidence (47–50).

Figure 10a – A prototype colour calibration slide produced in collaboration with the University of Leeds and FFEI Limited, UK (47). The colour patches are created using a histochemical stains and a biopolymer which uptakes stains like real tissue. Each of the patches have a ‘known’ colour. When the colour calibration slide is scanned into a virtual slide scanner, the resultant colour patches on the whole slide image can be used to compare with the ‘known’ values in much the same way as a MacBeth colour checker (Figure 7).

Figure 10b – The effect of our unique colour calibration test object on the appearance of WSIs. The first case is a Ziehl-Neelsen stain of a lung biopsy with mycobacterium with and without calibration. The second is a normal duodenal biopsy stained with H&E, again with and without colour calibration.

Guidance and Recommendations

In digital radiology, where most images are greyscale, the need for image standardization is recognized as essential to ensure diagnostic consistency and has become a standard part of all radiology imaging workflows. There are clear guidelines regarding calibration requirements and minimum technical standards of displays used for diagnosis (51–53). The method used for radiological display calibration is the Digital Imaging and Communications in Medicine (DICOM) Part 14 Grey Scale Display Function (GSDF) (54). This method ensures that each shade grey in the transition from white to grey is perceptually linear by taking into account the ability of the human visual system for light, dark and chromatic adaption, preventing two shades of grey from being too close together perceptually. Studies have indicated that calibration using DICOM Part 14 GSDF for radiology displays results in better observer performance (55,56) and much of the research in digital

radiology is based around compliance to these standards (57).

As previously stated, the FDA have produced guidance stating that it is essential to control colour in digital pathology, and have recommended the use of a target slide ideally with similar spectral characteristics to stained tissue (22). Furthermore guidance from the ICC White Paper 44 (58) regarding displays for diagnostic digital pathology, indicates that all medical grade displays should be colour calibrated and the display should be checked for compliance every 50 days, as the displays can change over time. They also state that ambient light must be stable, as calibration of the display takes into account the ambient lighting. Unfortunately, only a few studies have addressed this topic in digital pathology (40,41), and so a lack of primary research has prevented the production of guidance concerning minimum display requirements for diagnostic digital pathology to date.

The Future of Colour in Digital Pathology

It is clear that an integral part of digital pathology that has yet to be solved is colour standardisation; in order to do so, further work is needed focussing on fine tuning colour calibration methods in relation to the effect on diagnosis. Looking to the future, we envisage several ways in which colour accuracy in pathology may be used to improve healthcare:

1. Once a global baseline for colour has been established, this will optimise diagnostic accuracy/ reliability as well as facilitating worldwide collaboration. As pathologists get better at using WSIs for routine work, they will re-adjust to the appearance of the calibrated slides, but they will still likely adapt these colours to best suit themselves and the tissue, activating personal colour profiles.

2. Secondly, digital image analysis algorithms for immunostaining will supersede human semi-quantitative analysis. This will be due to improvements in algorithm reliability, facilitated by colour calibration. This might be used alongside colour normalisation techniques to account for pre-processing variation in staining.
3. Given that tissue staining is a shackle of light microscopy, digital pathology opens new doors with regards to pseudostaining. This involves digital superposition of colour to the WSI as opposed to staining the actual tissue. Some work has already been undertaken in this field, including an interesting paper by Kather et al 2015(59), which presented altered colours for histology based upon manipulation of colour maps to enhance perceptual contrast. Further work is needed to establish if and how such approaches might improve diagnostic quality.

The incentives to digitize pathology workflow are significant; enhancing working efficiency, global collaboration, teaching/ training benefits and improved quality and patient safety. Inherent to the widespread introduction of digital pathology is appropriate colour management of WSIs. However, colour management in this field is made difficult primarily due to the absence of a 'Gold Standard'; future research should focus on trying to solve this pivotal issue.

Acknowledgements

Authors EC and DT wrote the review.

We wish to take this opportunity to thank Dr Nick Mapstone, Lead Pathologist for the North West UK Bowel Cancer Screening Programme, for the EQA participant feedback included within this review.

References

1. Gaucher E. The Diagnostic Accuracy of Digital Microscopy: A Systematic Review. Press.
2. Collins Dictionary Online [Internet]. [cited January 5, 2016]. Retrieved from: <http://www.collinsdictionary.com/>
3. Physics of Color [Internet]. [cited April 14, 2016]. Retrieved from: <https://en.wikipedia.org/wiki/Color>
4. Sinn P, Wirtz R, Keller M, Waldburger N, Kriegsmann M, Veltrup E, et al. Comparison of IHC image analysis, IHC visual scoring, and quantitative real time PCR data for prediction of residual cancer burden after neoadjuvant chemotherapy. Vol. 50, European Journal of Cancer. 2014. p. S198–9.
5. Berns R. Billmeyer and Saltzman's Principles of Colour Technology, 3rd Edition. Third. Wiley; 2000.
6. Robertson D. The woman with superhuman vision [Internet]. 2014. Retrieved from: <http://www.bbc.com/future/story/20140905-the-women-with-super-human-vision>
7. Krupinski EA. Role of Color in Telemedicine Applications.

- Accepted Article
8. Krupinski EA. Current perspectives in medical image perception. *Atten Percept Psychophys* [Internet]. 2010;72(5):1205–17. Retrieved from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3881280&tool=pmc&rendertype=abstract>
 9. Macadam DL. Visual Sensitivities to Color Differences in Daylight. *J Opt Soc Am*. 1942;32:247.
 10. Kang HR. Chromatic Adaptation. In: *Computational Color Technology* [Internet]. p. 43–56. Retrieved from: <http://ebooks.spiedigitallibrary.org/content.aspx?bookid=122§ionid=31560914>
 11. Fairchild M. *Color Appearance Models, Second Edition*. John Wiley & Sons Ltd; 2005. 1-34 p.
 12. Winkler A, Spillman L, Werner J, Webster M. Asymmetries in blue-yellow color perception and in the color of “the dress”. *Curr Biol*. 2015;15:547–8.
 13. Wellbourne LE, Morland A, Wade A. Human colour perception changes between seasons. *Curr Biol*. 2015;25:646–7.
 14. Bartels A, Zeki S. The architecture of the human color center. *Eur J Neurosci*. 2000;12:172–93.
 15. Heiting G. How Your Vision Changes as You Age [Internet]. [cited April 14, 2016]. Retrieved from: <http://www.allaboutvision.com/over60/vision-changes.htm>
 16. Schneck M, Haegerstrom-Portnoy G, Lott L, Brabyn J. Comparison of panel D-

- 15 tests in a large older population. *Optom Vis Sci*. 2014;91:284–90.
17. Akman O, Kosemehmetoglu K. Ocular diseases among pathologists and pathologists' perceptions on ocular diseases: a survey study. *Turkish J Pathol* [Internet]. 2015;194–9. Retrieved from: <http://www.turkjpath.org/doi.php?doi=10.5146/tjpath.2015.01326>
18. Poole CJM, Hill DJ, Christie JL, Birch J. Deficient Colour Vision and Interpretation of Histopathology Slides: Cross Sectional Study. *Source BMJ Br Med J* [Internet]. 1997;315:1279–81. Retrieved from: <http://www.jstor.org/stable/25176231> <http://www.jstor.org/page/>
19. Campbell WS, Talmon GA, Foster KW, Lele SM, Kozel JA, West WW. Sixty-five thousand shades of gray: Importance of color in surgical pathology diagnoses. *Hum Pathol* [Internet]. Elsevier Inc.; 2015;46:1945–50. Retrieved from: <http://dx.doi.org/10.1016/j.humpath.2015.08.016>
20. Levenson RM, Krupinski EA, Navarro VM, Wasserman EA. Pigeons (Columba livia) as Trainable Observers of Pathology and Radiology Breast Cancer Images. *PLoS One*. 2015;10:e0141357.
21. Bowel Cancer Screening Programme External Quality Assurance Feedback from Dr Nic Mapstone [Internet]. Retrieved from: <http://www.gieqa.org.uk/b15a/>
22. FDA. Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices; Guidance for Industry and Food and Drug Administration Staff. 2016;
23. Gray A, Wright A, Jackson P, Hale M, Treanor D. Quality control of histochemical stains: repeatable and objective quantification using virtual

slides [Oral presentation]. 2013.

24. Gray A, Wright A, Jackson P, Hale M, Treanor D, James S. Quantification of histochemical stains using whole slide imaging: development of a method and demonstration of its usefulness in laboratory quality control. *J Clin Pathol.* 2015;68:192–9.
25. Schuessler Z. Delta E - Color Difference Algorithms [Internet]. [cited April 14, 2016]. Retrieved from: <http://zschuessler.github.io/DeltaE/learn/>
26. Shrestha P, Hulsken B. Colour accuracy and reproducibility in whole slide imaging scanners. *SPIE Digit Libr.* 2014;9041.
27. Ng CS, Palmer CR. Analysis of diagnostic confidence and diagnostic accuracy: a unified framework. *Br J Radiol.* 2007;80:1550–62.
28. Ali HR, Irwin M, Morris L, Dawson S-J, Blows FM, Provenzano E, et al. Astronomical algorithms for automated analysis of tissue protein expression in breast cancer. *Br J Cancer* [Internet]. Nature Publishing Group; 2013;108:602–12. Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/23329232>
29. Ayad E, Mansy M, Elwi D, Salem M, Salama M, Kayser K. Comparative Study between Quantitative Digital Image Analysis and Fluorescence In Situ Hybridization of Breast Cancer Equivocal HER2 2+ Cases. *J Pathol Inform.* 2015;6(31).
30. Helin HO, Tuominen VJ, Ylinen O, Helin HJ, Isola J. Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry [Internet]. 2nd ed. Vol. 468. Germany: Springer Verlag; 2016. p. 191–8. Retrieved from:

<http://link.springer.de/link/service/journals/00428/index.htm>

31. Holten-Rossing H, Moller Talman M-L, Kristensson M, Vainer B, Ben-Chetrit N, Chetrit D, et al. Quantitative immunohistochemical analysis reveals association between sodium iodide symporter and estrogen receptor expression in breast cancer. *Am J Clin Pathol* [Internet]. United States; 2013;8:2123–31. Retrieved from:
<http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med3&NEWS=N&AN=1967930>
32. Laurinaviciene A, Dasevicius D, Ostapenko V, Jarmalaite S, Lazutka J. Membrane connectivity estimated by digital image analysis of HER2 immunohistochemistry is concordant with visual scoring and fluorescence in situ hybridization results: algorithm evaluation on breast cancer tissue microarrays. Vol. 6, *Diagnostic pathology*. 2011. p. 87.
33. Magee D, Treanor D, Crellin D, Shires M, Smith K, Mohee K, et al. Colour Normalisation in Digital Histopathology Images. *Opt Tissue Image Anal Microsc Histopathol Endosc*. 2009;100–11.
34. Reinhard E, Ashikhmin M, Gooch B, Shirley P. Color transfer between images. *IEEE Comput Graph Appl*. 2001;21:34–41.
35. Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. *Arch Pathol Lab Med*. 2011;135:233–42.
36. SpyderGALLERY. DATACOLOR. Lawrenceville, US;

37. McCamy C, Marcus H, Davidson J. A color-rendition chart. J App Photog Eng. 1976;2:95–9.
38. Colorimeters [Internet]. [cited April 18, 2016]. Retrieved from: <http://www.xrite.com/colorimeter>
39. Spectrophotometry [Internet]. [cited April 18, 2016]. Retrieved from: <http://www.xrite.com/spectrophotometer>
40. Krupinski EA, Silverstein LD, Hashimi SF, Graham AR, Weinstein RS, Roehrig H. Observer Performance Using Virtual Pathology Slides: Impact of LCD Color Reproduction Accuracy. J Digit Imaging. 2012;25:738–43.
41. Kimpe T, Avanski A, Espig K, Rostang J, Van Hoey G, Xthona A. Influence of Display Characteristics on Clinical Performance in Digital Pathology. Diagn Pathol. 2016;8:7–8.
42. Hanna MG, Monaco SE, Ahmed I, Parwani A V, Pantanowitz L. Impact of Monitor Color Calibration on Digital Pathology Interpretation. Pathology Informatics. Pittsburgh, US; 2015.
43. Badano A, Revie WC, Casertano A, Cheng WC, Green P, Kimpe T, et al. Consistency and standardization of color in medical imaging: a consensus report. J Digit Imaging. 2015;28:41–52.
44. Bautista P, Hashimoto N, Yagi Y. Color standardization in whole slide imaging using a color calibration slide. J Pathol Inform [Internet]. 2014;5:4. Retrieved from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3952402&tool=pmc&rendertype=abstract>

- Accepted Article
45. Wei H, Brill MH, Park T. Evaluation of Targets for Color Calibrating Digital Images from an Optical Bright-Field Transmission Microscope. *Colour Res Appl.* 2014;1–8.
 46. Leica Microsystems (UK) Limited. Milton Keynes, UK;
 47. Clarke E, Revie C, Brett D, Wilson R, Mello-Thoms C, Treanor D. Colour Calibration in Digital Pathology: The Clinical Impact of a Novel Test Object (oral presentation). In: 13th European Congress on Digital Pathology, Berlin. Boston, US; 2016.
 48. Revie C, Clarke E, Shires M, Jackson P, Brett D, Cochrane R, et al. Development of a novel tissue-mimicking colour calibration assessment target for digital microscopy. (submitted).
 49. Clarke E, Review C, Brett D, Wilson R, Mello-Thoms C, Treanor D. Colour Calibration in Digital Pathology: The Clinical Impact of a Novel Test Object. (submitted).
 50. Clarke E, Revie C, Brett D, Wilson R, Mello-Thoms C, Treanor D. Colour Calibration in Digital Pathology: the Clinical Impact of a Novel Test Object (abstract). *J Clin Pathol Suppl.* 2016;238:S1–20.
 51. Hirschorn DS, Krupinski EA, Flynn MJ. IT Reference Guide for the Practicing Radiologist. 2013;1–12.
 52. Compton BK, Oosterwijk H. REQUIREMENTS FOR MEDICAL IMAGING MONITORS (part I). (part I):1–13.
 53. National Quality Assurance Coordinating Group for Radiography. Quality

assurance guidelines for mammography including radiographic quality control.

Natl Qual Assur Coord Gr Radiogr [Internet]. 2006;(63). Retrieved from:

<http://capitadiscovery.co.uk/bcu/items/1175568> \n <http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp63.pdf>

54. Rosslyn S. Digital Imaging and Communications in Medicine (DICOM) Part 14 : Grayscale Standard Display Function. *Medicine* (Baltimore). 2004;10:3–4.
55. Krupinski E, Roehrig H. The influence of a perceptually linearized display on observer performance and visual search. *Acad Radiol*. 2000;7:8–13.
56. Johnston R, Zimmerman J, Rogers D, Pizer S. Perceptual Standardization. In: 3rd Intl Conf on Picture Archiving and Communication Systems Proc SPIE 0536. 1985.
57. Butt A, Mahoney M, Savage N. The impact of computer display performance on the quality of digital radiographs: a review. *Austral Dent J*. 2012;57:16–23.
58. International Color Consortium. White Paper # 44 Visualization of medical content on color display systems. 2016;(April).
59. Kather JN, Weis C-A, Marx A, Schuster AK, Schad LR, Zöllner FG. New Colors for Histology: Optimized Bivariate Color Maps Increase Perceptual Contrast in Histological Images. *PLoS One*. 2015;10:e0145572.









