International Color Consortium®

White Paper # 44

Level: Advanced

Date: April 2016

Revised: October 2023

Visualization of medical content on color display systems

1. Introduction

Since 1993, the International Color Consortium (ICC) has worked on standardization and evolution of color management architecture. The architecture relies on profiles which describe the color characteristics of a device in a reference color space.

The ICC describes its profiles as *"… tools to translate color data created on one device into another device's native color space […] permitting tremendous flexibility to both users and vendors. For example, it allows users to be sure that their image will retain its color fidelity when moved between systems and applications"*.

This document focuses on results and recommendations for the correct use of ICC profiles for visualization of grayscale (GSDF [1]) and color medical images on color displays. These recommendations allow for visualization of medical content on color display systems, whereby DICOM GSDF images and pseudo color images can all be presented effectively on the same display. The results and recommendations in this document were first discussed in the ICC Medical Imaging Working Group (MIWG).

Hardware display calibration is usually found to give better results. Software calibration, including the use of ICC profiles, can lead to lower display accuracy, due for example to tonal distortion and lack of a 10-bit pipeline. Hardware calibration is also more computationally efficient.

2. Calibration requirements, standards and guidelines for medical displays

Medical displays that are used for primary diagnosis need to achieve minimum performance levels to ensure that subtle details in radiological images are visible to the radiologist.

The National Electrical Manufacturers Association (NEMA) and The American College of Radiology (ACR) developed a Standard for Digital Imaging and Communications in Medicine (DICOM) [1]. Part 14 describes the Grayscale Standard Display Function (GSDF). This standard defines a way to take the existing Characteristic Curve of a display system (i.e. the relationship between the Luminance Output for each Digital Driving Level or pixel value) and modify it to the Grayscale Standard Display Function.

GSDF has become the generally accepted worldwide standard for calibration of display systems for use in radiology. In many countries throughout the world calibration to GSDF is a requirement in local regulations, laws and guidelines for display systems that are used to diagnose and review radiological images. In some countries a distinction is made between displays used for diagnosis versus review of radiological images, where specifications for primary diagnosis displays are set higher than for clinical review displays.

For medical displays systems used for applications other than radiology, the situation is less clear. There is no generally accepted standard today that describes how display systems for non-radiology modalities, such as ophthalmology, endoscopy, pathology, etc., need to be calibrated. Sometimes these displays are also calibrated to GSDF, but in other cases another standard such as sRGB is used. Typically, the display system in such applications is closely coupled to the acquisition device. It is typically the manufacturer of the acquisition device that designs, integrates and tests the display, ensures that it has appropriate characteristics and obtains the required regulatory approvals.

There is currently no agreed upon standard for the specific color characteristics that are needed for radiology displays. There are works in progress on this matter and one of these is described in more detail in section 3 of this document.

Displays that are not used by medical professionals to make diagnosis or review clinical images typically are not calibrated to GSDF. This is the case, for example, for displays used by referring physicians looking at patient records.

The recommendations in this whitepaper therefore fit better in situations where displays are calibrated to DICOM GSDF such as displays used for primary diagnosis or review of radiological images. It is important to note that while display calibration is a necessary condition for diagnostic displays, other requirements must also apply in certain countries (e.g. [2]).

3. Different uses of color in medical imaging

This section describes three typical uses of color in medical applications.

Accurate color (also known as 'true color') is explained first, followed by monochrome images, and finally pseudo color. For pseudo color, three types are discussed: sRGB, sRGB+GSDF, and GSDF+perceptually linear.

3.1. Accurate reproduction of color medical images

In recent years, medical imaging data has made increasing use of color images. As of this writing, color medical imaging has not been standardized, although there are several works available which discuss this matter [3] [4].

Depending on the specific field of medicine, requirements for the representation of colors may vary. For instance, for the interpretation of wound photographs, color is an indicator of the healing state of the wound and correct representation of colors is important, and dermatologist are demanding standardization [5].

The use of ICC profiles for achieving a good color reproduction across devices [6] is already a well-established practice in different fields including pathology [7], [8]. By connecting the acquisition device profile to the display profile, it is possible to achieve good color reproducibility thanks to the colorimetric rendering intents.

ICC MIWG has defined best practices for digital color photography in medicine in *White Paper 46: Improving Color Image Quality in Medical Photography* [9]. Recommendations given in this document fully applies to medical images requiring high color fidelity.

It is recommended to use ICC profiles with colorimetric rendering intents, and to follow the guidelines of WP46 [9] for all cases where accurate color reproduction is needed, including digital pathology, endoscopy, laparoscopy, dermatology, and medical photography.

3.2. Perceptually uniform visualization of monochrome medical images

3.2.1. DICOM Grayscale Standard Display Function (GSDF)

A known issue with the distribution of images on hardcopy or softcopy media is that images are usually inconsistent and can have different perceptions [10]. This means that depending on the hardware, images will have different contrast values or luminance differences. DICOM [1] has proposed a standard, called GSDF, for the purpose of ensuring that grayscale radiology images are presented consistently on different devices.

GSDF is a relative calibration method which aims to linearize the perception of luminance of a display without reducing its luminance dynamic range. The perceived variation is expressed as Just Noticeable Difference (JND) and is based on Barten's Contrast Sensitivity Function [11][12]. A good introduction to GSDF is available by Fetterly et al. [13].

This standard display calibration is applicable for any grayscale medical imaging modality, such as X-ray based medical image modalities, fluorescence, magnetic resonance, computed tomography, mammography, angiography, and echography, even if combined with pseudo colors (annotations, fusion of modalities like PET-CT...), and it has positive impact on diagnostic performances [14]. However, GSDF is defined for monochrome images, and therefore is not applicable to true color medical images like endoscopy or dermatology.

3.3. Pseudo color for medical images

As described in paragraph 3.1, certain medical disciplines require accurate color representation. However, some modalities use colors to display numerical/quantitative information on top of grayscale images as illustrated by Figure 1. The exact color used to visualize quantitative information is less of importance as long as differences are easily perceivable, and it is easy for the observer to visually determine what quantitative value is being represented by a specific color. This is the case for example (but not limited to) PET-CT fusion, dose distribution of IMRT, thermal camera images, Doppler in echography and image analysis or AI software outcomes overlaid on top of grayscale images.

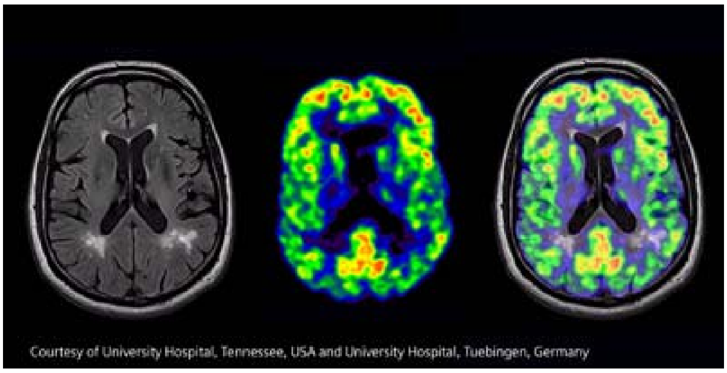

Figure 1: Example of PET-CT hybrid image. Positron Emission Tomography gives a quite accurate localization of metabolic activities represented in color and super projected on an X-ray Computed Tomography.

This quantitative imaging approach typically relies on color scales or 'colormaps' to represent calculated values. Figure 2 is an example of a "rainbow" color scale.

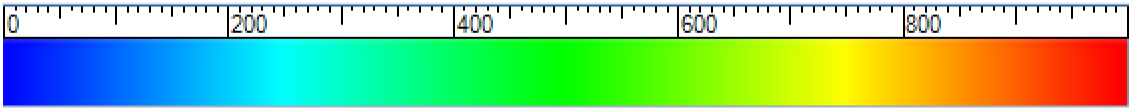

Figure 2: Rainbow color scale

Pseudo color is used to highlight features in the image that may be difficult to see using only luminance differences. Different color scales are used, according to the type of modality and preferences of diagnosing physician.

Because the display is the component that in the end generates the colors, the choice of color scale and display hardware affects the comparative visual analysis of pseudo-color images [15].

An example of such various color scales is a "rainbow" color scale [16]. This example associates colors with values from 0 to 1000. A large part of the color scale is covered by green, making it difficult to differentiate values from 350 to 650. On the other hand, only a thin band close to 750 is yellow, making this value clearly distinguishable. Also, depending on the quantitative value (and the color it corresponds to) it may be easy or difficult to perceive small differences in that quantitative value.

Various color scales have been conceived so that features can be properly detected even in conventional sRGB. However, in this case of rainbow color scale, a perceptually linear distribution of the colors could help to optimize color visualization for quantitative values and might reveal hidden details in the image. This can only be accomplished by matching the color scale specifications with the colors and luminance response of the display used for visualization, or by calibrating the display itself in a perceptually linear way. (A colour space is perceptually uniform when the Euclidean distance between any two colours in the space is proportional to the visual difference between them. [17])

Possible solutions are discussed below.

3.3.1. sRGB

sRGB representation is one widely-used approach, especially where interoperability across display devices and/or ensuring compatibility with existing color scales is a priority. The sRGB standard is defined by IEC 61966-2-1:1999 [18]. The primaries (gamut) and the white point are shown in Table 1. The maximum luminance is specified at 80 cd/m^2, however, in practice higher luminances are normally used.

	Red	Green	Blue	D65
x	0.6400	0.3000	0.1500	0.3127
y	0.3300	0.6000	0.0600	0.3290
z	0.0300	0.1000	0.7900	0.3583

Table 1: The primaries (gamut) and white point of sRGB

Full details of the sRGB encoding can be seen in the ICC 3-component encoding registry. sRGB has an electro-optical transfer function (EOTF) that defines a piecewise non-linear relationship between RGB drive signals and the displayed luminance. This function has some similarity to the function used to derive the perceptually uniform CIELAB L* lightness from luminance, and as a result sRGB drive signals correspond approximately to perceptually uniform steps over a limited dynamic range. In darker colors, and for higher luminances, this relationship becomes less uniform. The sRGB EOTF is sometimes simplified to 'gamma 2.2', where the exponent 2.2 is applied to the RGB signals to obtain the relative luminance.

If the perceptual uniformity of the color scales is more important, these can also be adjusted. For example, once a display has been calibrated to sRGB, a color scale can be created using the color differences between drive levels in sRGB, such that they remain constant. Such a color scale has been proposed, for example, by Nuñez et al [19].

Not only because sRGB standard color space devices have been the majority across the color imaging workflow in medical institutions, but also based on the fact that existing color scales have been established and utilized using sRGB by modality manufacturers; use of sRGB is recommended where interoperability across displays is required.

While sRGB is still in widespread use, modern display devices are commonly offering gamuts wider than sRGB, and now refer (for example) to the DCI-P3 standard (SMPTE-EG-0432-1:201). The static sRGB EOTF is also less applicable to High Dynamic Range (HDR) display technologies. HDR typically follows the PQ (SMPTE ST 2084) or HLG EOTF, both of which are defined in ITU-R BT.2100-2.

HDR and wide gamut characteristics are often combined in modern devices and applying a strict sRGB calibration loses the advantage of these additional colors, potentially reducing the distinguishability and readability of medical images in comparison to other approaches. Future discussion is needed on the use of these HDR and wide gamut colors for medical usages.

3.3.2. sRGB + GSDF

Calibrating the grayscale to GSDF makes it possible to maintain compliance with DICOM GSDF while achieving a degree of interoperability.

If the software supports color management, it may be possible to display monochrome and color images simultaneously by applying appropriate color management to images that should be displayed in either GSDF or the sRGB EOTF, based on profile information that characterizes the display.

3.3.3. Color Scale Display Function (CSDF)

The goal of a perceptually linear display calibration is that equal differences in RGB drive levels at the display input produce equal perceptual differences in the display output. For instance, DICOM Grayscale Standard Display Function (GSDF) is a perceptually linear calibration of the grayscale. Unfortunately, the definition of a Just Noticeable Difference (JND) given by DICOM [1] only considers luminance, and does not take chromaticity into account. For this reason, an extension of the GSDF to color cannot be achieved by using the same metric. A number of approaches have been proposed, including the use of CIEDE2000 [20], a commonly used color difference metric, for extending towards perceptually linear color behavior. As a drawback, a completely perceptual linear display purely based on the CIEDE2000 metric would not be DICOM GSDF compliant on gray. Moreover, calibrating a display in such a way that it is perceived as being linear throughout its complete color gamut in terms of CIEDE2000 is not an easy task and obtaining a completely uniform color space requires reducing the display gamut and to decrease display luminance and display contrast. This means that the full hardware capabilities of the display cannot be used. Such a calibration is described in [17].

A calibration using the CIEDE2000 color difference metric [20] to make a display as perceptually linear as possible has been proposed without shrinking its gamut and preserving the DICOM GSDF calibration of the grayscale [21]. This calibration is called CSDF (Color Scale Display Function) and is proposed as an extension of GSDF towards color.

CSDF relies on several color sweeps through the RGB cube as depicted in Figure 3:

- the GSDF calibration of the grayscale (from Black $(0{,}0{,}0)_{RGB}$ to White $(1{,}1{,}1)_{RGB}$)
- the CIEDE2000 calibration of:
  - The different edges of the RGB cube.
  - Sweeps from Primary colors (Red $(1{,}0{,}0)_{RGB}$, Green $(0{,}1{,}0)_{RGB}$ and Blue $(0{,}0{,}1)_{RGB}$) to White $(1{,}1{,}1)_{RGB}$
  - Sweeps from Secondary colors (Cyan $(0{,}1{,}1)_{RGB}$, Magenta $(1{,}0{,}1)_{RGB}$ and Yellow $(1{,}1{,}0)_{RGB}$) to Black $(0{,}0{,}0)_{RGB}$

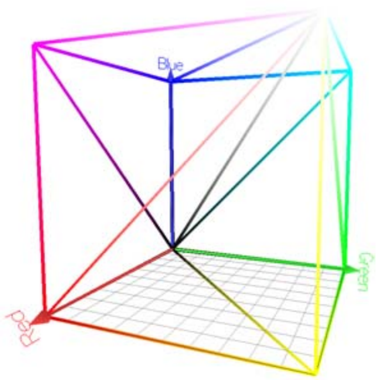

Figure 3: Color sweeps represented in the RGB cube that are made perceptually linear by CSDF. The Gray scale is linearized according to DICOM JND, and the other lines according to $\Delta\boldsymbol{E}_{\mathbf{2000}}$.

Both DICOM GSDF and CSDF are relative calibrations which aim at linearizing the perceptual differences between levels without shrinking the gamut of the display. It is therefore critical to accurately estimate the gamut of the device.

As shown in Figure 4, the average CIEDE2000 step varies between the primaries to maintain the gamut integrity. Other colors are then adjusted to ensure a smooth transition for the Gray GSDF to the CIEDE2000 calibrated colors. Therefore, perceptual linearities are maintained on color sweeps mentioned above, while other color combinations will not necessarily be perceptually linear.

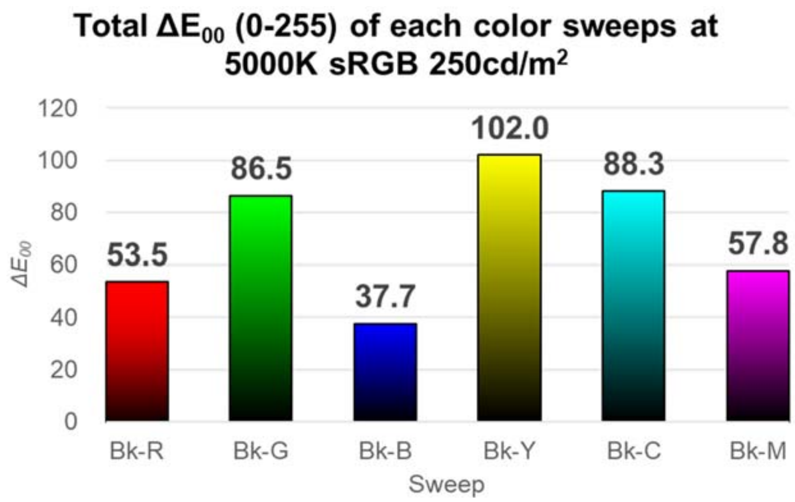

Figure 4: Sum of CIEDE2000 differences of primary and secondary color sweeps.

There is a wide range of colormaps for pseudo color scales, some of which may make it easier to see color differences and which are used for different purposes and preferences [22] [23].

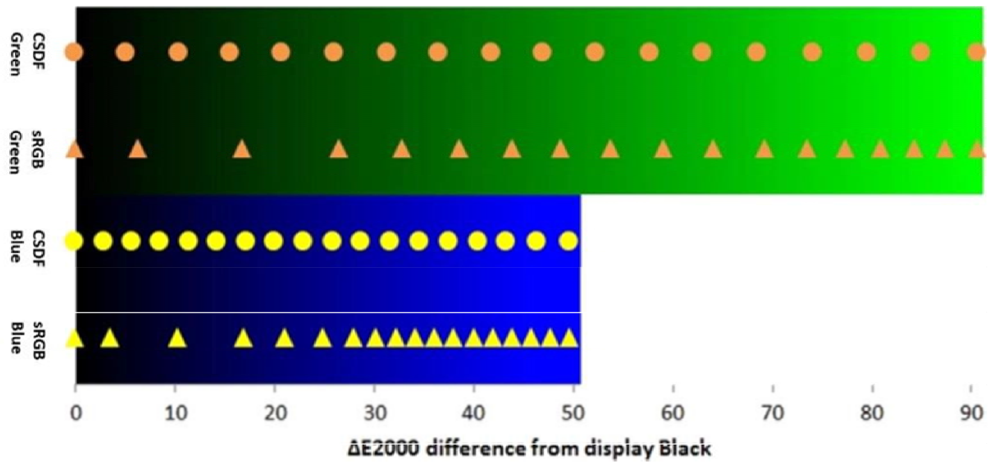

Figure 5: Representation of the color perceptual difference along Black-to-Green and Black-to-Blue color ramps for sRGB (triangles) and CSDF (disks). As the total perceptual difference changes from one ramp to the other, the linearized steps are also different to make use of the entire display gamut.

The rescaling will make perceptual differences between DDLs more uniform. To achieve this some will be made larger, others smaller. Overall, the minimum step along the ramp will be higher than with sRGB (Figure 5). One study has found no significant difference in the detectability of nuclear medicine images between CSDF and gamma 2.2 [24]. Another one has found lower error in color-to-value interpretation when looking at a rainbow color scale with CSDF calibration compared to sRGB [25]. Other studies have demonstrated that a display calibrated to CSDF results in a higher accuracy than one calibrated to sRGB or GSDF in PET/CT images [26].

3.3.3.1 ΔEITP

CIEDE2000 is defined for an illuminance of 1000 lux (equivalent to 318 cd m^{-2}). Medical displays conforming to DICOM requirements may have considerably higher luminances. An alternative color difference metric, ΔEITP, has been proposed for high dynamic range displays, and has a luminance function based on PQ or HLG in ITU-R BT2100. ΔEITP is defined in ITU-R BT.2124. HLG is similar to a gamma-shaped EOTF, while PQ, like GSDF, is built on Barten's Contrast Sensitivity Function [11], although the aim points may differ from those defined for GSDF.

3.4. Simultaneous display of monochrome and color images

The ICC framework is widely used for correct visualization of color. The ICC published guidelines for best practices for digital color photography in medicine, ICC White Paper 46 [9], which covers the use cases where high color fidelity is required.

One approach to realise this task is to apply DICOM GSDF for grayscale images, and apply the chosen color calibration for color images. This subsection focuses on how the ICC framework can be used together with DICOM GSDF and a color calibration to achieve the required color accuracy.

The proposed calibration method of this section is capable of simultaneously allowing visualization of medical content on color display systems, whereby DICOM GSDF images and color images can all be presented effectively on the same display.

Medical display systems are usually able to perform automatic GSDF calibration and internally stabilize their brightness. In such situations the display continuously or periodically measures its own characteristics and the ambient light conditions by means of sensors and consequently adapts its behavior when necessary. For example, a medical display could alter internal look up tables and other settings to make sure it remains compliant with the DICOM GSDF standard. It is also possible that such a display reacts to changes in ambient light by changing its display luminance or by changing the calibration curve. For these reasons, the display behavior and therefore the "display profile" can change whenever the display adapts its internal calibration settings.

Therefore, it is only possible to use color management with these self-calibrating displays if their ICC profiles are updated each time they change their behavior. In the case of non-self-calibrating displays it is even more important that ICC profiles for these displays are regenerated at a sufficient frequency (see sections 6 and C.4), since non-self-calibrating displays can have a behavior that fluctuates significantly over time.

The proposed method consists of creating two ICC profiles when characterizing a display. The first one describes its native color behavior (display profile) and a second one describes the ideal calibration based on its properties (calibration profile). The proposed color management workflow is schematized in Figure 6.

The ICC framework defines several rendering intents. The method focuses on the Colorimetric intents as it aims to perfectly match the colors from the source profile on the display.

If the software supports the option, it is possible to display monochrome and color images simultaneously by applying appropriate color management to images that should be displayed in monochrome or color, selecting the appropriate calibration profile for each image type.

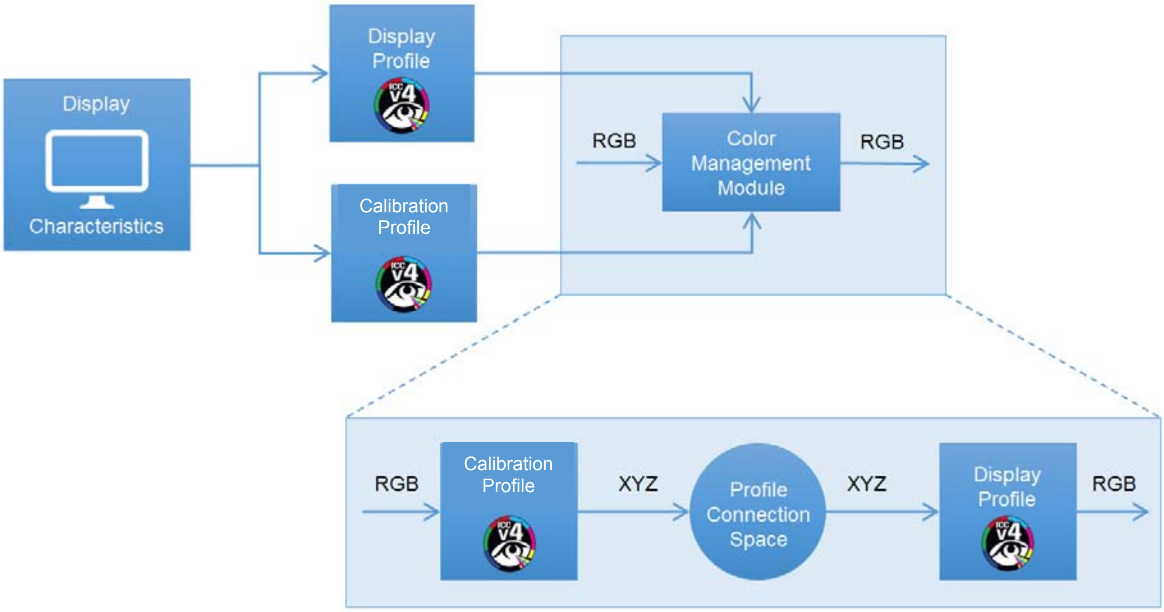

Figure 6: ICC workflow for software color management

There are systems that can automatically switch between different calibrations using hardware. In this case, the software does not need to support color management.

4. Creation of the profiles

ICC specifies different type of profiles balancing between performance and memory foot-print. Annex A details the structure of each of them for the present use-case.

To achieve a DICOM GSDF only calibration, the use of monochrome profiles is possible. Monochrome profiles are designed to be used with monochromatic devices. They can be used to calibrate grayscale displays to DICOM GSDF at a low processing cost. They can also be used to calibrate color displays to DICOM GSDF. By combining a monochrome source profile and a color display profile, CMM will return RGB triplets where $R = G = B$. Thus, the color monitor will not display colors anymore.

Matrix-TRC profiles and N-component LUT based profiles are designed for color devices. Matrix profiles have a simple structure. They perform very well in describing theoretical display standards or models as the ones we use in this study, but are unable to describe the internal constraints of a real LCD display (Annex A.2). In a lot of application this limitation is not really an issue, but it could be the case here.

LUT based profiles are much more powerful, but also more complex. The N-dimensional Color Look-up table (CLUT) is the core element of this structure, and is the one allowing describing a CSDF calibration in an ICC profile (see section A.3).

A final possibility is to use DeviceLink profiles. Unlike ordinary source or destination profiles, DeviceLink profiles do not describe a specific color space but the conversion from a source to a destination color space. In the present use case, it is possible to use a DeviceLink profile to describe the color transformation to be applied on a display to calibrate it. This does not influence the performance of the calibration process (see Annex A.4).

5. Profile quality assessment methods

When creating ICC profiles, it is important to control their correctness. Two complementary tests are proposed to validate profiles.

- The first test compares a given display model $L^*a^*b^*$ output with the corresponding ICC profile output to estimate the accuracy of the Device-To-PCS conversion of the profile.
- The second test consists in Device-To-PCS followed by PCS-To-Device conversions using the same profile in order to perform a roundtrip. This second test estimates the invertibility of the profile, and combined with the test above, indirectly measures the accuracy of the PCS-To-Device conversion.

Both tests are described in detail in ISO 23564.

Results from the fidelity test suggest that matrix-based profiles cannot be used to describe the target CSDF calibration and thus a LUT-based profile is needed.

The round trip test results point to the importance of the size of the CLUT in LUT-based profiles to correctly calibrate to CSDF.

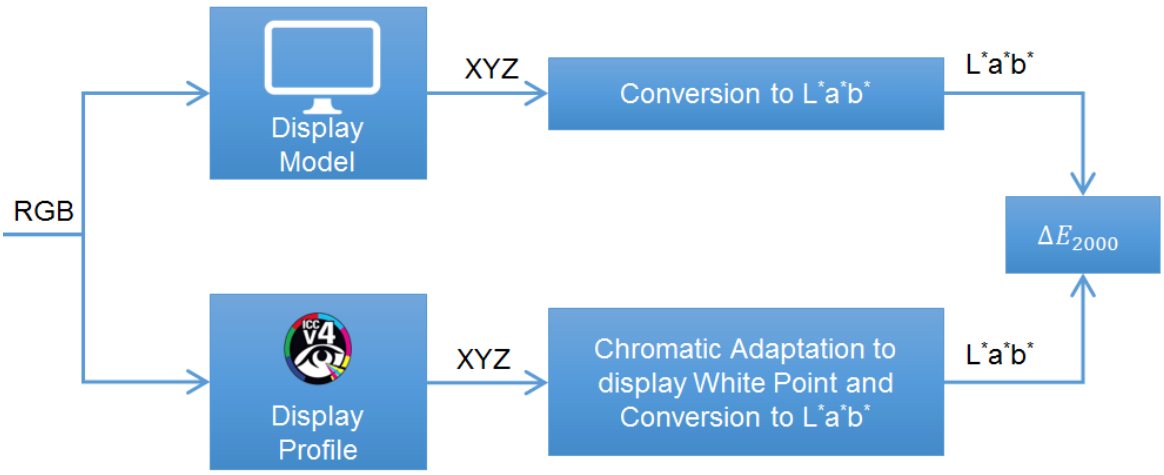

Figure 7: ICC Profile Model Fidelity test

5.1. ICC.2 profiles

ICC.1 profiles continue to be widely used across all applications, and at the time of writing the current version is ICC.1:2022 (version v4.4). In some situations ICC.1 is too restrictive to achieve the desired transform, and the advanced color management architecture ICC.2 (also known as iccMAX) may be appropriate. An ICC.2 profile is not limited to a fixed PCS or sequence of transform elements, and supports run-time execution through a custom color transform language. An example of its application to the workflow described in previous sections could be to encode both display and perceptually linear transform in a single profile, with the CMM making a run-time selection. ICC.2 PCS and transform elements are better able to handle the variable luminance requirements in GSDF, and to accept run-time input of viewing condition parameters through environment variables.

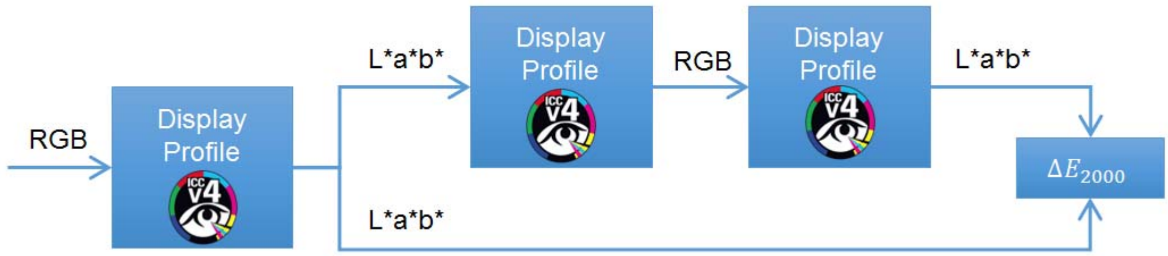

Figure 8: ICC profile roundtrip test

6. Recommendations

The following recommendations allow for visualization of medical content on color display systems, whereby DICOM GSDF images, pseudo color images can all be presented effectively on the same display.

When performing color management by software, it is recommended to accurately generate and apply source and destination profiles.

It is recommended that the source and destination profiles be the same (same display characteristics) in order to display pseudo color images accurately. The GSDF/sRGB/CSDF calibration may also be realized by hardware, and the chosen method should be capable of meeting the aims and tolerances of the calibration.

When realizing color transformation by using ICC Profiles and a CMM, the following recommendations below are provided. For non-calibrated displays, it is suggested that the recommendations below are followed with the goal to stay within 10% tolerance of the grayscale target and within 15% tolerance of the color target

- System configuration:
  - Only use ICC profiles that have been created for the specific display. Generic profiles do not offer sufficient accuracy, even if the display can be set to a reference state.
  - Every time a display setting is changed (e.g. display luminance or contrast settings), new source and destination profiles need to be created and used. Use at least $10\ bit$ connections from application to software when a most accurate calibration is needed, since $8\ bit$ ones are clearly not sufficient for these use cases. It is to be noted that $10\ bit$ connections can be achieved by means of $8\ bit$ display controllers that use appropriate dithering algorithms.
  - Display luminance and contrast should be stabilized to the value that was used when creating the profile, since luminance and contrast deviations result into reduced calibration accuracy (See Figure 14 and Figure 17).
  - If the luminance cannot be stabilized, a “warming-up” period should be respected before the display can be used. A period of $2\ hours$ (see Figure 16) is recommended, but this time may be reduced if the stability and warm-up of the display is known and reproducible.

- ICC Profile and CMM:
  - Both source and destination profiles might take the ambient light into account.
  - For sRGB, the tone curve and matrix elements that convert from RGB to XYZ PCS can be encoded in an AToBTag, and their inverse in a BToATag. A matrix/curve type profile can also be used.

- For CSDF profiles, both source and destination profiles should be LUT based profiles using XYZ color space as PCS as described in section A.3. As explained in section 4, it is also possible to use DeviceLink profiles.
- For DICOM GSDF calibration of grayscale display, the use of monochrome profile is possible, and recommended.
- For CSDF calibration, the CLUT of the source profile (describing the calibration) should ideally have a size of at least $11 * 11 * 11$ points, but using at least $33 * 33 * 33$ points is recommended for a more accurate calibration. The display profile can be matrix-based, but we recommend using a more accurate LUT-based profile as depicted in section A.3.
- Special attention should be given to PCS-To-Device conversion of the Black point. This is critical to achieve an acceptable calibration. See Annex C.

- Calibration process:
  - The calibration compliance must be verified regularly since typical display behavior changes over time as Figure 21 shows. If the compliance test fails, the whole calibration process has to be repeated. This means renewing display measurements and regenerating the display profile based on these measurements. More frequent measurements are possible and could guide determining when recalibration is needed.
  - Ambient light should be stable. Otherwise, the calibration process must be repeated every time the ambient light conditions change (see Figure 26).

7. Patent statement

The International Color Consortium (ICC) draws attention to the fact that it is claimed that implementation of section 3.3.3 (on CSDF) of this White Paper can involve the use of a patent. ICC takes no position concerning the evidence, validity and scope of this patent right. The holder of this patent right has assured the ICC that he is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ICC. Information may be obtained from:

Barco N.V.
Healthcare Department
Beneluxpark 21, 8500 Kortrijk
Belgium

Whether specific implementations of a particular recommendation listed in the white paper are covered by the subject matter disclosed in these patents/patent applications is to be evaluated by any interested party.
The Patent Holder does not claim that these patent/patent applications are required for the implementation of a particular recommendation listed in the white paper. In case of doubt, any entity is welcome to contact Barco patent department for further clarifications on the matter.

Annex A. Detailed structure of the ICC profiles

Version 4.4 of the ICC specification [27] makes it possible to use different architectures to build ICC profiles.

For information about the structure of ICC profiles, the reader should consult the ICC specifications and other ICC White Papers.

A.1. Monochrome profiles

ICC defined monochrome profiles to describe grayscale devices. The Gray TRC tag contains the Gray Tone Reproduction Curve, representing the conversion from the device Digital Driving Level to the achromatic channel of the PCS. This curve can be composed of up to 4096 points, or being a predefined parametric curve.

The display profile then contains an accurate description of the “native” display function, while the input profile describes the exact DICOM target for the given display luminance and contrast.

In addition to required tags, the luminance tag was included because of the necessity to ensure the input profile contains a DICOM calibration corresponding to the luminance of the display, as the calibration depends on both the maximum and minimum luminance of the device. The luminance tag is informative and the CMM is not expected to perform any processing with it.

A.2. Three-component Matrix-TRC-based profiles

Matrix-based profiles perform very well in describing theoretical display standards or models as the ones we use in this study, but are unable to describe the internal constraints of a real LCD display. For instance, this architecture cannot deal with cross-talk in between the sub-pixels of a Liquid Crystal Panel. On an actual display, the Red TRC depends on the levels of Green and Blue as shown on Figure 9, and the simplicity of this model does not allow this phenomenon to be taken into account.

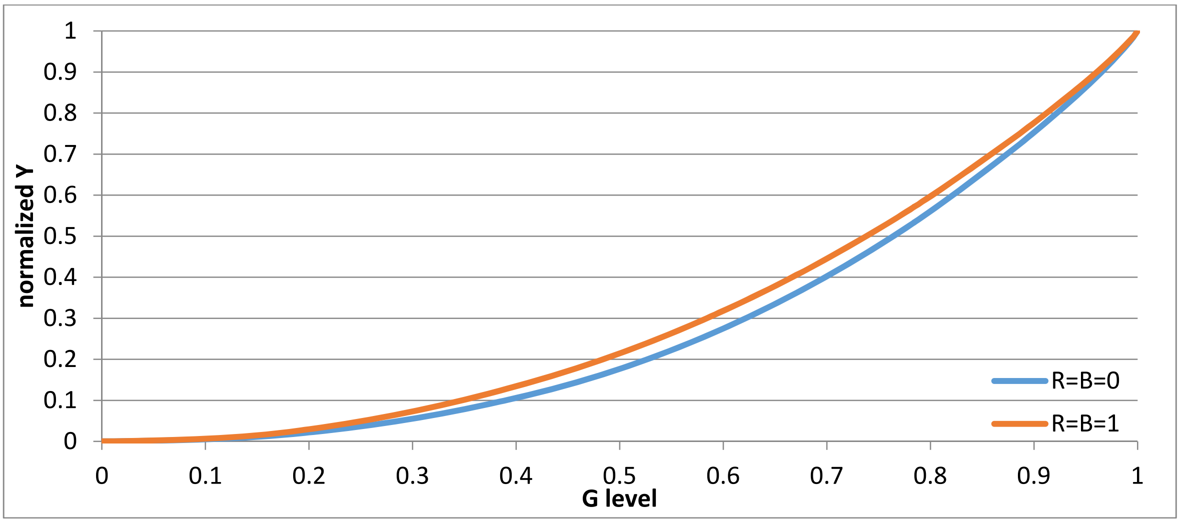

Figure 9: Comparison of the Green Luminance curve of a diagnostic color display, when Red and Blue Chanel are both set to 0% and 100%. The curves have been normalized for better readability. The horizontal axis represents the Green ddl, and the vertical axis is the normalized luminance (Y-Ymin)/(Ymax-Ymin).

This approximation may be acceptable depending on the use-case.

For the same reason, this profile architecture cannot be used to represent the CSDF calibration as the light output of the display for a given RGB triplet is no longer directly proportional to the light of the three primaries in this case.

For instance, CIEDE2000 calibration of the Black-to-Green sweep can be represented by the Green TRC tag, but this TRC is also applied on the Magenta-to-White sweep whereas its calibration is completely different. Figure 10 illustrates this by presenting an example of how the Green channel is impacted by the CSDF on different parallel color sweeps in the RGB cube.

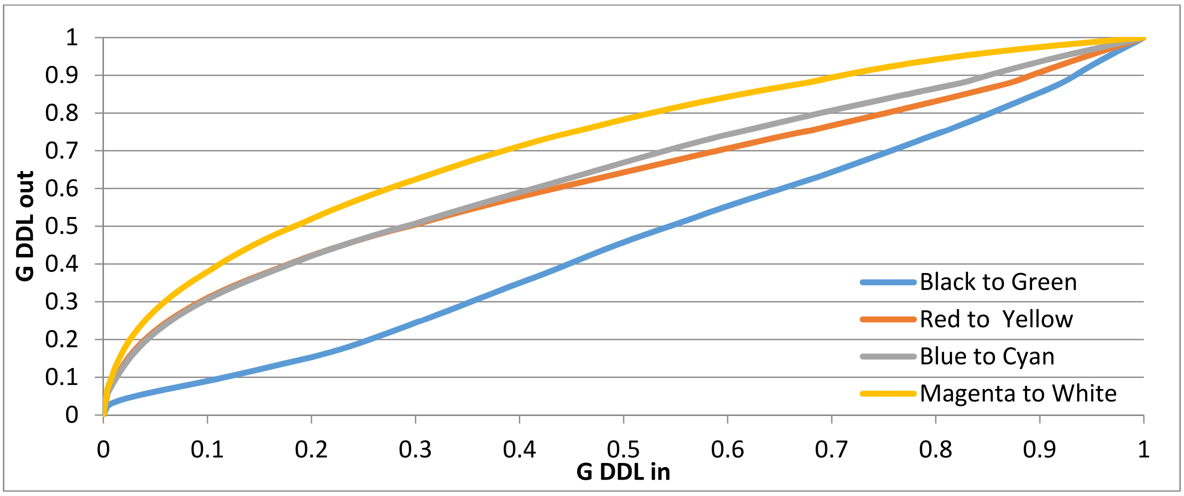

Figure 10: Green-to-Green 1D LUT for CSDF calibration of different color sweeps. All these sweeps are defined in RGB by a changing G value from 0 to 1 and with R and B constant.

A.3. N-component LUT-based profiles

In a LUT-based profile, complex conversion pipelines may be implemented, as depicted by Figure 11. The general principles of these pipelines and requirements of the different elements are available in the ICC specifications [27]. Below is described a possible implementation.

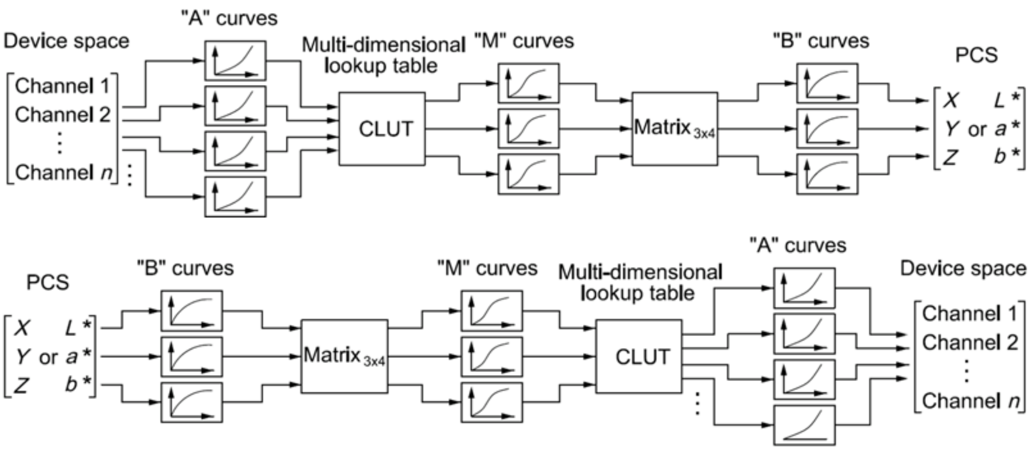

Figure 11: Device-to-PCS and PCS-to-Device conversion workflows for LUT based profiles. The different elements arround the Color LUT (CLUT) can be used to create a nonlinear repartition of the input values of the LUT, or set to identity.

During our experiments, the number of input and output channels were equal to 3. We used only cubic LUTs (same size on every dimension) and only a few sizes have been tested.

Profiles were constructed using the XYZ PCS.

- A-curves are unused in both Device-to-PCS and PCS-to-Device conversions and were thus set to be identity curves.
- CLUT stages are used for RGB-to-RGB conversions. The tables may contain individual point corrections to make the profile more faithful to the display it represents.
- M-curves are used to apply inverse RGB companding in Device-to-PCS conversion and RGB companding in PCS-to-Device conversion. This handles the RGB-XYZ non linearity and makes RGB linear.
- The matrix is used to finish the linear RGB-to-XYZ conversion and thus contains RGB reference primaries as XYZ values, written in column order. For encoding reasons, these values are all divided by 2.
- B-curves, as A-curves, were unused in both tags and were set to identity curves.

In the case of the aforementioned display models, identity 3D CLUTs were used in the CLUT stage since there is no physical display color point correction to apply on them. For CSDF profiles, a 3D color correction LUT is calculated and encompassed in the CLUT stage of the pipeline. CSDF profiles are always built from a DICOM model to ensure a good DICOM calibration by storing the GSDF in the profile M-curves.

A.4. DeviceLink profiles

DeviceLink profiles contain a single table similar to the one presented at the top of Figure 11. The DeviceLink profile can be built similarly to the N-component LUT based profile; except that every element related to PCS can be removed. In the end, only the RGB-To-RGB conversion is preserved in the CLUT element, and if necessary, the 1-dimensional RGB-To-RGB LUT ensuring GSDF calibration can be stored in the B-Curves.

DeviceLink profiles present as main drawback a lack of flexibility. Indeed, each profile corresponds to a single very precise situation. While the classical workflow allows for example to use the same source profile when a display's internal calibration state is changed from sRGB to DICOM, and only update the display profile, Here it is necessary to update the DeviceLink profile, and so to recalculate the RGB-To-RGB calibration LUT.

A.5. ICC.2 profiles

The ICC.2 profile format supports all the transform elements described in this Annex, with the exception of the matrix/TRC type of profile. Instead, matrices and curves are stored as elements in a LUT-based profile, with an identity transform in the CLUT, as described for display profiles in A.3

Annex B. Calibration quality assessment methods

This annex outlines methods for assessing color calibration quality for medical displays. sRGB and GSDF assessment are covered in detail elsewhere, so this Annex focuses mainly on CSDF.

B.1. How to evaluate the quality of the sRGB calibration

For sRGB, the methods described in established standards such as ISO 14861:2015 [28] (4.4.2.2 Colorimetric accuracy) are recommended in order to validate the accuracy of the calibration. The color differences between the measurement results and the predicted ideal values should be within the following tolerances: colors in gamut should conform to a mean CIEDE2000≤2.5 and 99th percentile CIEDE2000≤4.5.

B.2. How to evaluate the quality of the grayscale calibration

As grayscale must comply with the DICOM GSDF standard, the quality assessment procedure for grayscale also matches the DICOM definition.

Quantitative assessment of luminance response is accomplished by using defined test patterns and luminance meters to measure the display device's luminance response for a limited number of values. The measurement protocol is similar to the one described in Annex C of the DICOM standard [1].

The grayscale compliance evaluations presented hereafter are based on luminance measurements of 18 evenly spread driving levels. These correspond to RGB triplets that can be represented as:

$$(R, G, B) = (0,0,0), (\frac{1}{17}, \frac{1}{17}, \frac{1}{17}), \ldots, (\frac{17}{17}, \frac{17}{17}, \frac{17}{17})$$

For each of them the relative difference from the theoretical grayscale target to the observed luminance value, but also perceptual difference (JND) from one patch to the next one is evaluated.

Grayscale compliance of a system is summarized as the maximal error encountered on those 18 points. The lower the value, the better the calibration compliance score is. This score must fall within 10% ([29] section 4.3) for devices used for the interpretation of medical images (diagnostic).

Recommendations for display quality other than luminance response can be found in [29] and [30].

B.3. How to evaluate the quality of the CSDF calibration

This section suggests a methodology for quantifying compliance/accuracy of the color component of the CSDF calibration. It is very important to stress that in this section only a metric is being described for specifically assessing the *color* aspects of CSDF calibration. As explained before, CSDF calibration also requires that the neutral grey diagonal of the display complies with DICOM GSDF (see section B.2). For clarity reasons however, and to be able to clearly separate greyscale from color calibration performance, the reported results and graphs for "color" only refer to the metric described in B.3.

CSDF defines a different behavior for grayscale and saturated colors. For this reason, they have to be evaluated separately, and a valid calibration must comply on both criteria. This means that for a display system to be CSDF compliant, both the metrics/graphs of sections B.2 and B.3 need to be generated and the results need to be within the described tolerances.

In this recommendation, we quantify perceptual linearity of colors of a display based on the output obtained by sweeping primary and secondary colors. We define a series of 18 RGB triplets between Black and Red, with equal steps in the R channel value between them. Corresponding values are:

$$(\mathrm{R}, G, B) = (0,0,0), \left(\frac{1}{17}, 0,0\right), \ldots, \left(\frac{17}{17}, 0,0\right)$$

Likewise, evenly spread series of 18 RGB triplets are defined between Black and the other Primary colors (Green and Blue) as well as between Black and the Secondary colors (Cyan, Magenta, and Yellow). Corresponding sweep values from Black to Yellow, for example, are:

$$(R, G, B) = (0,0,0), \left(\frac{1}{17}, \frac{1}{17}, 0\right), \ldots, \left(\frac{17}{17}, \frac{17}{17}, 0\right)$$

Leaving out the Black duplicates this results in 120 unique RGB triplets for which the corresponding display output is obtained as XYZ. Because discrimination between colors is less relevant at low luminance levels, we discard measurements corresponding to a driving level which results in a luminance below $5\ cd/m^2$ in the sweep between Black and White.

For example, if the triplet $(R, G, B) = (4/17, 4/17, 4/17)$ is the first measurement in the Black-To-White sweep presenting a Y value of at least $5\ cd/m^2$, then the measurements $(R, G, B) = (0,0,0), \ldots, (3/17,0,0)$ of the Black-To-Red sweep are discarded. The same logic is applied on the other sweeps.

All non-discarded measurements are then converted to $L^*a^*b^*$ values by taking the XYZ of full White as the reference White point. Next, we calculate CIEDE2000 between consecutive points in each of the six sweeps for the Primary and Secondary colors, resulting in six series of CIEDE2000 values noted Δ_i with i representing the color sweep (Red, Green, Blue Cyan, Magenta or Yellow). Each value $\Delta_{i,j}$ within the set Δ_i is then normalized by dividing them by the series average.

$$\bar{\Delta_\iota} = \sum_{j=1}^{N} \frac{\Delta_{i,j}}{N} \qquad\qquad \forall j,\ \ \delta_{i,j} = \frac{\Delta_{i,j}}{\bar{\Delta_\iota}}$$

For an ideal perceptually linear display, the resulting normalized curves would all be constant with value 100% ($\forall i\ \forall j,\ \ \delta_{i,j} = 1$). For each of the six sweeps, we quantify the perceptual linearity D_i as the maximum deviation from 100%.

$$D_i = \begin{cases} \max_j(\delta_{i,j}), & \max_j(\delta_{i,j}) - 1\ \geq 1 - \min_j(\delta_{i,j}) \\ \min_j(\delta_{i,j}), & \max_j(\delta_{i,j}) - 1\ < 1 - \min_j(\delta_{i,j}) \end{cases}$$

The overall perceptual linearity is quantified as the maximum deviation encountered in any of the six curves.

$$D = \max_i(D_i)$$

If the perceptual linearity metric value is within a predefined tolerance range, e.g. $\pm 15\%$ (i.e. $0.85 < D < 1.15$), the display calibration is considered to be perceptually linear.

The color compliance evaluations below are presented as the relative deviation from the target (i.e. values below 15% are compliant, values above are not).

The tolerance threshold is defined as a relative value because absolute values can vary a lot depending on the gamut of the device (e.g. Adobe RGB gamut presents superior $\Delta_{i,j}$ than sRGB, and thus the same variation of R, G or B would induce a larger CIEDE2000 in Adobe RGB than in sRGB).

CIEDE2000 was derived for color differences < 5 in CIELAB units, and may not perfectly predict the perceptual uniformity of the color sweeps. Alternatives have been proposed, such as DEITP (see section 3.3.3) and CIECAM16-UCS.

The limit of 15% was selected based on what is achievable in practice and what is commonly used as a tolerance level for DICOM GSDF (see e.g. [29]). In case of DICOM GSDF most guidelines and local regulations require that for primary reading the maximum deviation shall be 10%, whereas for Review applications this can be up to 15%. It is possible that based on future studies specific thresholds could be defined for different modalities.

In parallel to the perceptual linearity of the colors, the DICOM GSDF compliance of the grayscale must be controlled too, as GSDF is part of CSDF calibration. The method to assess the GSDF quality is described in section B.2 and [1], [30].

B.4. Calibration smoothness

Green proposed in 2008 [31] a methodology for estimating the smoothness of a color transform from a transformed ramp. The color transform can be the result of the application of a 3D color LUT or of ICC profiles. The method is represented by Figure 12.

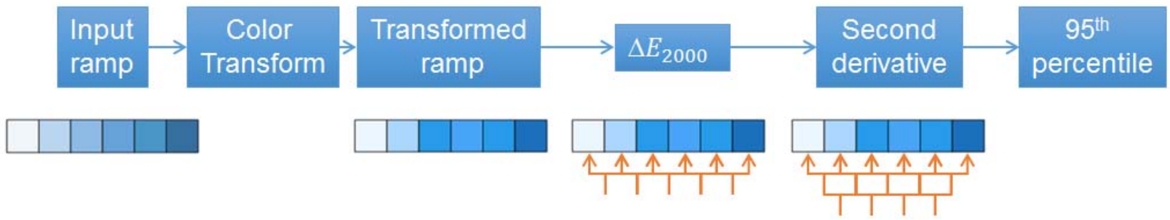

Figure 12: Green' smoothness workflow

For any input colored ramp with n pixels, the metric CIEDE2000 defined by the CIE is computed between CIELAB consecutive triplets of the ramp, resulting in $n-1$ CIEDE2000 values. From this resulting ramp, a second derivative is calculated by simply subtracting two consecutive elements of the CIEDE2000 ramp resulting in a set of $n-2$ values for which the 95^{th}percentile is calculated representing the smoothness of the color transform of the input ramp. (The optimum percentile level was determined to be 95th to best fit the subjective data from the measurement of the magnitude of tone jumps of 96 test gradations [31]).

In order to consider the entire calibration, Green's metric is applied on a large number of gradients through the RGB cube: The RGB cube is sampled to $50 * 50 * 50$ RGB triplets from which are built a total of 7500 ramps of 50 color shades.

Based on the 3D representation of the RGB cube, in the directions defined by the 3 main axes, 50-elements ramps are extracted as illustrated on Figure 13.

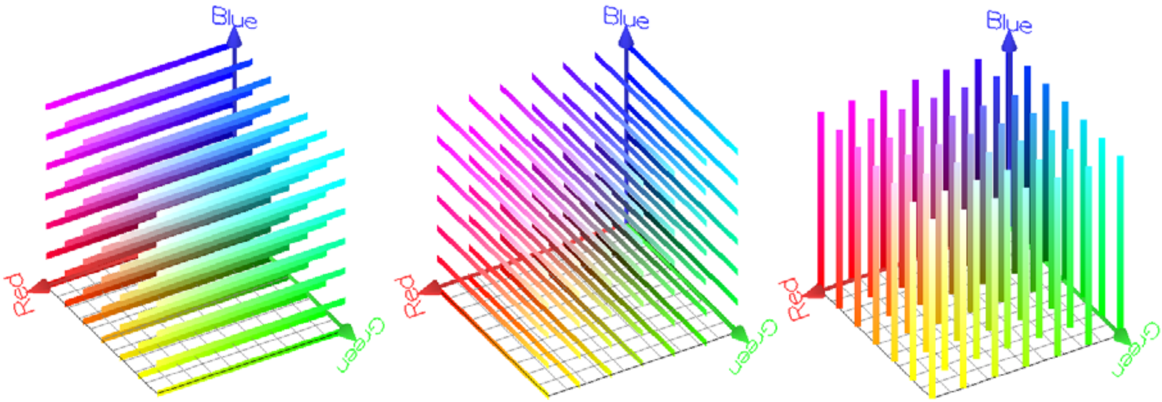

Figure 13: Examples of ramps used to evaluate the smoothness of a calibration. Here only $\mathbf{3}$ sets of $\mathbf{7} * \mathbf{7}$ ramps are represented while the tests involved a total of $\mathbf{3} * \mathbf{50} * \mathbf{50}$ ramps

The smoothness of the color transform for each input ramp is computed. Simple statistics can be calculated based on the 7500 smoothness values obtained: average, standard deviation, minimum and maximum. A perfect smoothness would have a value of 0.

Annex C. Impact of inaccurate profiling

Building ICC profiles relies on measurements which are sensitive to noise and ambient conditions. Since profiles must have a reasonable size and generation time, measuring every display color is not a viable approach. The content of a LUT-based profile may therefore rely on interpolation in the cases where all the color points of the 3D CLUTs were not necessarily measured.

Furthermore, many on-screen-display controls (OSDs) make it possible for the user to select a display function from a collection of reference presets such as Gamma 2.2, Gamma 1.8, sRGB… In this situation, one could be tempted to use generic ICC profiles instead of characterizing the display in its actual configuration. However, the same preset on different displays can results in very different color rendering, and may not even be close to the standard they supposedly match [32].

The following paragraphs present the results of different simulations evaluating the impact of a mis-evaluation of different characteristics of the monitor or the ambient conditions.

C.1. Display luminance

A potential mismatch between profile luminance and the actual display may affect grayscale compliance since it is based on luminance. The display profile was fixed to $600\ cd/m^2$ and the influence of the difference between its luminance and the actual display luminance was assessed. Results are observable on Figure 14.

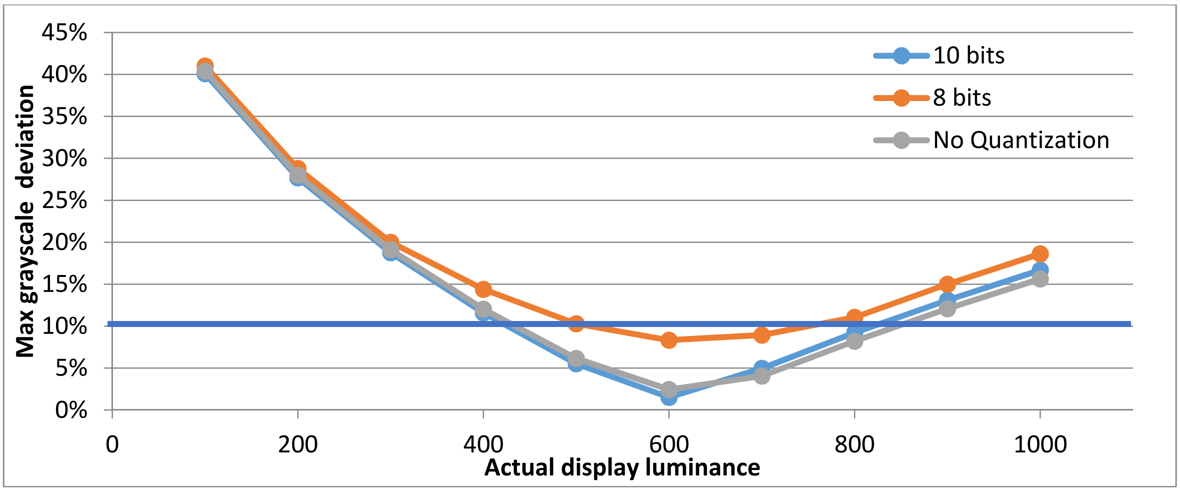

Figure 14: Influence on grayscale compliance of luminance mismatch between an sRGB profile describing a luminance of $\mathbf{600\ cd/m^2}$ relatively to the actual display.

Misestimating a display luminance in its ICC profile leads to large deviations from grayscale calibration: up to 14% error for a $200\ cd/m^2$ overestimation and 11% error for a $200\ cd/m^2$ underestimation.

For 8 bit systems, profile luminance misestimating invalidates the grayscale component of the calibration from a $100\ cd/m^2$ overestimation and approximately a $150\ cd/m^2$ underestimation. On the other hand, results remain compliant for both theoretical and 10 bit values if overestimation and underestimation does not respectively surpass $175\ cd/m^2$ and $215\ cd/m^2$.

With the presented method, a display's grayscale calibration will thus remain valid if the ICC profile describing the display does not encompass a luminance value that deviates largely from the actual one.

Tests have been repeated with different display profile architectures and different display native behaviors without observing noticeable differences.

It is also interesting to notice that misestimating the luminance of a display has no impact on the color component of the color compliance, as depicted by Figure 15. There is no visible difference between the theoretical values and 10 bit quantization.

It is possible to observe some variations of luminance on short term because of temperature variations within the display. Backlight efficiency depends on the lamps temperature. Liquid Crystals are also sensitive to temperature.

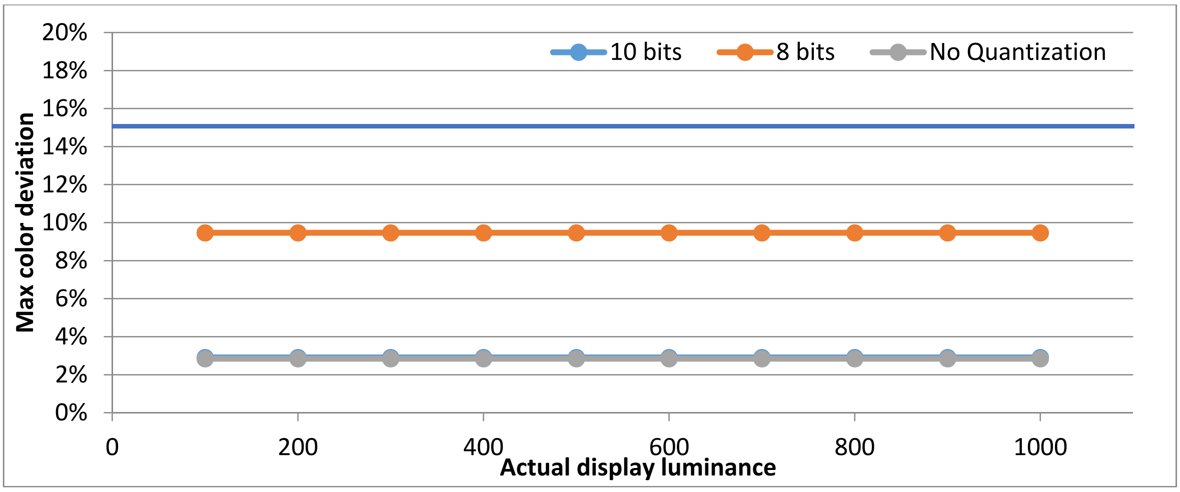

Figure 15: Influence on color compliance of luminance mismatch between an sRGB profile describing a luminance of $\mathbf{600}\ \boldsymbol{cd/m^2}$ relatively to the actual display.

Figure 16 shows how a display's luminance evolves from the moment it is turned on.

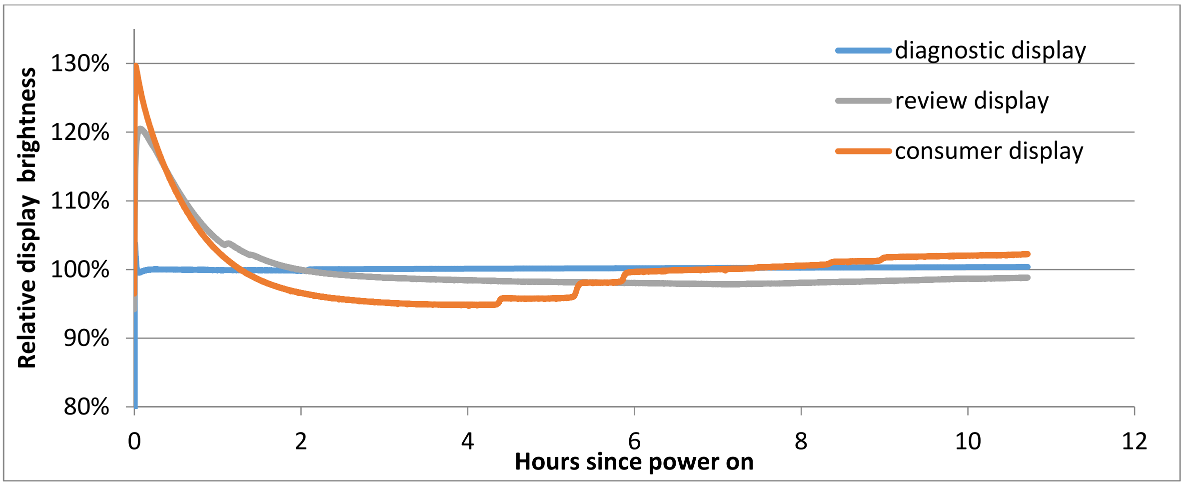

Figure 16: Short term evolution of a display luminance from start up. Diagnostic display is equipped with front and back sensors for real time stabilization. Clinical Review only has a back sensor. Consumer display is not stabilized at all. The different curves have been normalized to their average for an easier comparison. Some of the data presented here comes from [33].

Warming up creates an important overshoot during the display's first 2 hours of use, making it un-calibrated until the luminance has reached a normal level if the display cannot compensate it.

C.2. Display contrast

As with luminance inaccuracies, errors profiling contrast induce error in calibration accuracy.

Results for grayscale are assessed by evaluating grayscale compliance for several differences between display profile contrast and actual display model contrast. Results are presented in Figure 17.

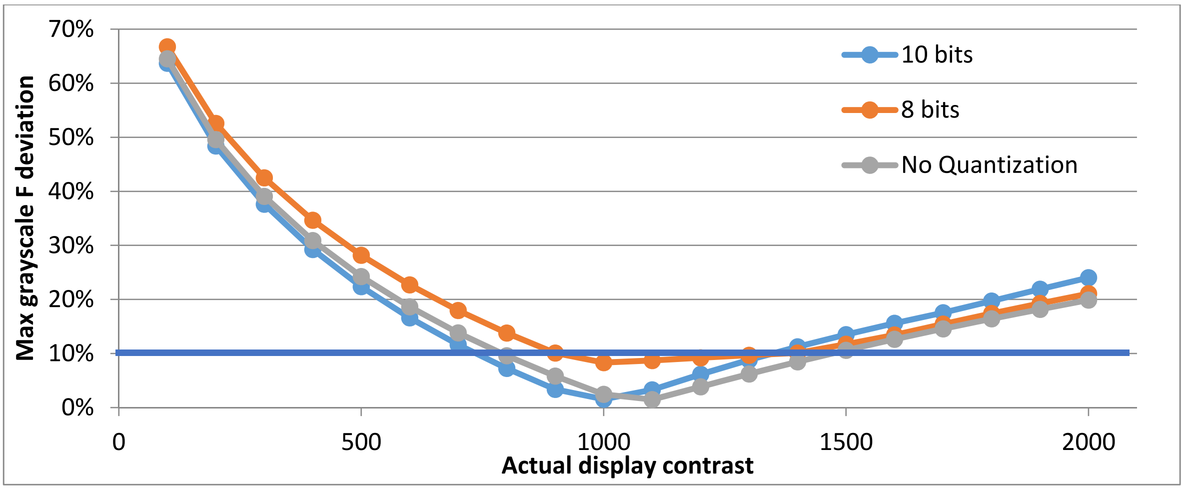

Figure 17: Influence on grayscale compliance of contrast mismatch between an sRGB profile describing a $\mathbf{1000{:}1}$ contrast ratio relatively and the actual display on GSDF deviation.

Measurement devices used to quantify a display's luminance are usually much more accurate on bright levels than they are on dim ones. Even low end devices would not return an error of more than 10 to $20\ cd/m^2$ while measuring luminance close to $600 cd/m^2$. As it is observable on Figure 14, this kind of error, which would already be considered as huge, would not impact grayscale compliance significantly.

However, contrast ratio is both much more sensitive than luminance to small variations and has a bigger impact on the calibration results. For instance, the reference display model has a luminance of $600\ cd/m^2$ and a contrast ratio of $1000{:}1$, which means that the Black point luminance of this display is $0.6\ cd/m^2$. Measurement devices are much more likely to return an erroneous value for the dimmest luminance level of a display. In this case, even an $0.2\ cd/m^2$ error would lead to a drop from $1000{:}1$ to $750{:}1$ contrast ratio, which would invalidate the grayscale calibration.

Perceptual linearity of colors is evaluated in a similar fashion, and the influence of contrast differences on the color component of the calibration are shown in Figure 18.

Contrast overestimation by the profile has a much larger influence on perceptual linearity of colors than a corresponding underestimation. For instance, if the profile's black point luminance is twice as high as the actual black luminance, deviation from color calibration raises from 3% to 8% on 10 bits systems. Contrarily, if the profile's black point luminance is twice as low as the actual one, deviation only raises from 3% to 5%.

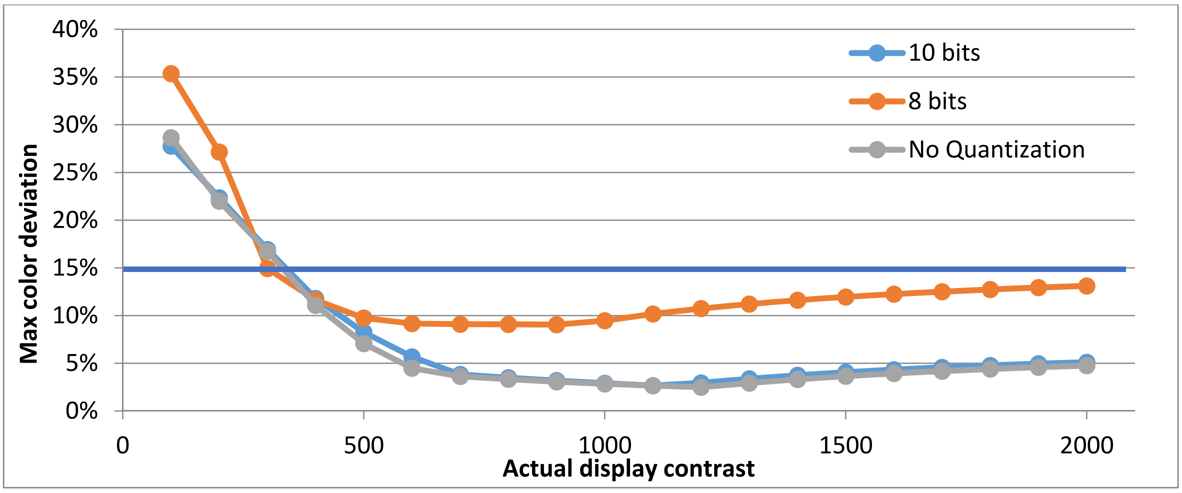

Figure 18: Influence on color compliance of contrast mismatch between an sRGB profile describing a $\mathbf{1000:1}$ contrast ratio relatively and the actual display contrast.

C.3. Display function (gamma)

The display function is typically the parameter that can be tuned from the display settings menu. Such menus usually propose to choose among a limited number of presets. Those presets are often common to different devices, and sRGB and Gamma 2.2 are available in almost every display. One could be tempted to use the display OSD in combination with a pre-created ICC profile.

Unfortunately, display presets are usually not accurate enough to allow such practices [33].

Figure 19 illustrates the fact that grayscale compliance is highly sensitive to imprecisions of the display functions contained in an ICC profile. An error of 0.05 in the estimation of the gamma is indeed enough to invalidate the grayscale calibration.

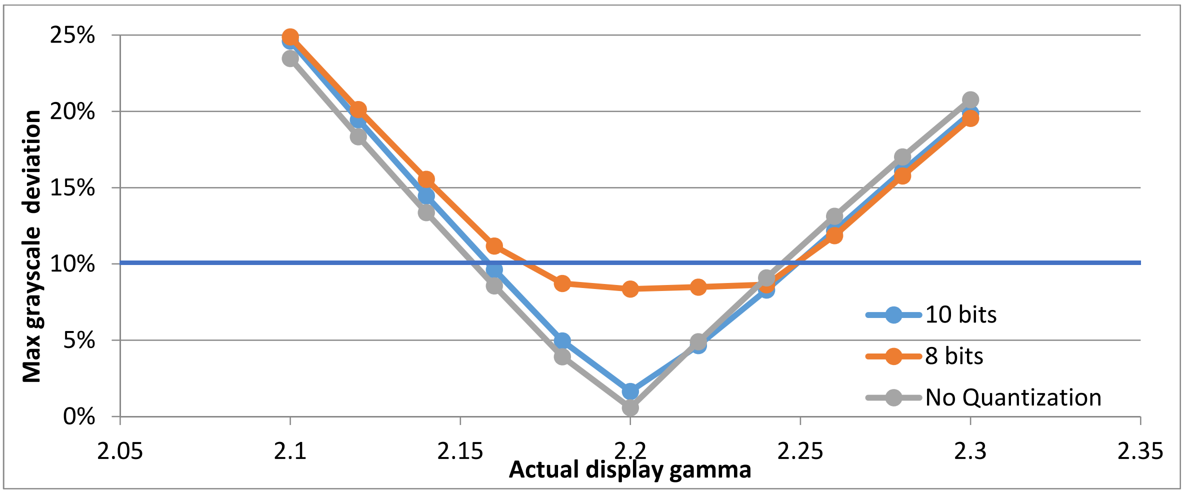

Figure 19: Influence of display function mismatch between a gamma2.2 profile and the actual display on grayscale compliance

On the other hand, the accuracy of the display function seems to be less critical for the calibration of the colors (see Figure 20) but remains a disturbing factor.

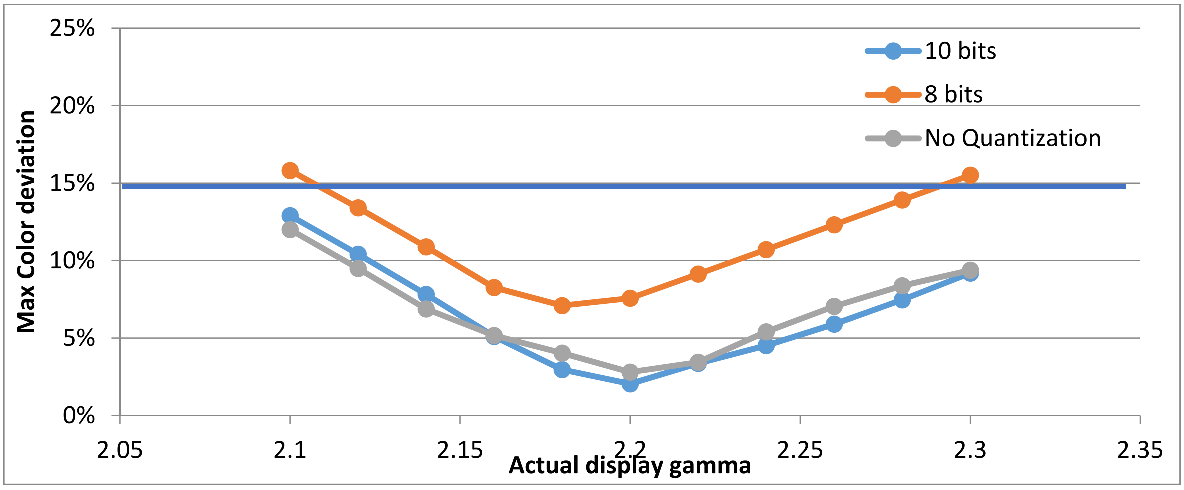

Figure 20: Influence of display function mismatch between a gamma2.2 profile and the actual display on color compliance

C.4. Display age

Because of the degradation of the materials composing the display, the colors it emits are susceptible to change in both chrominance and luminance over the lifetime of the device. Avanaki et al [34] have studied the effects of both of these variations on the interpretation of digital pathology images.

Figure 21 summarizes the variations observed while testing non-stabilized and stabilized displays of different types. By referring at section B.2 and Figure 21, it appears that aging is crucial in the grayscale compliance of the calibration.

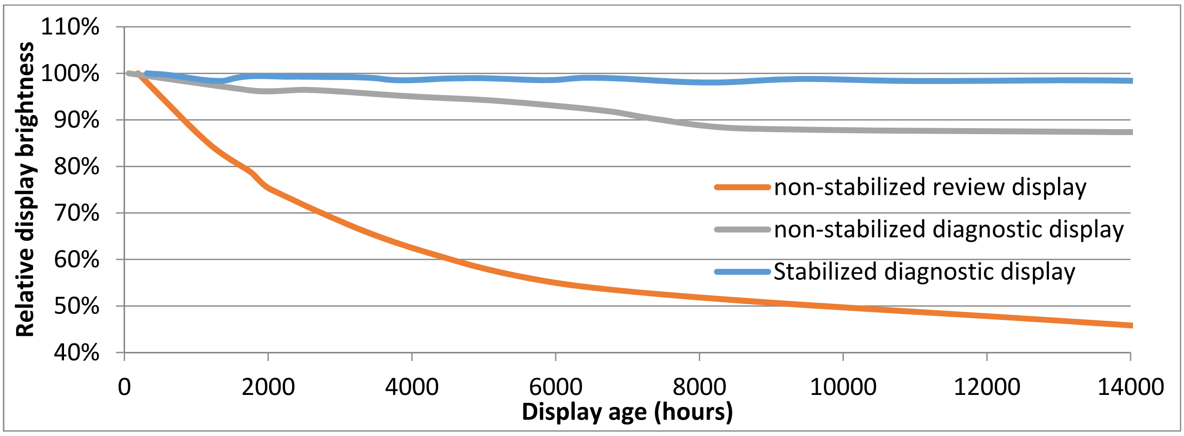

Figure 21: Long term evolution of the maximum luminance of different non-stabilized displays compared to stabilized displays

Please notice that different display systems can have a large difference in terms of performance and stability. While the luminance variation is almost only related to the decreasing efficiency of the backlight (CCFL or LED, where typically LED backlights are more stable over time [35]), and can be compensated by giving more power to the light sources, the color shift is more difficult to anticipate as it depends on the evolution of the optical properties of different layers of diffusers and filters. To evaluate the impact of this color shift on the calibration, a medical display has been characterized over its entire lifetime.

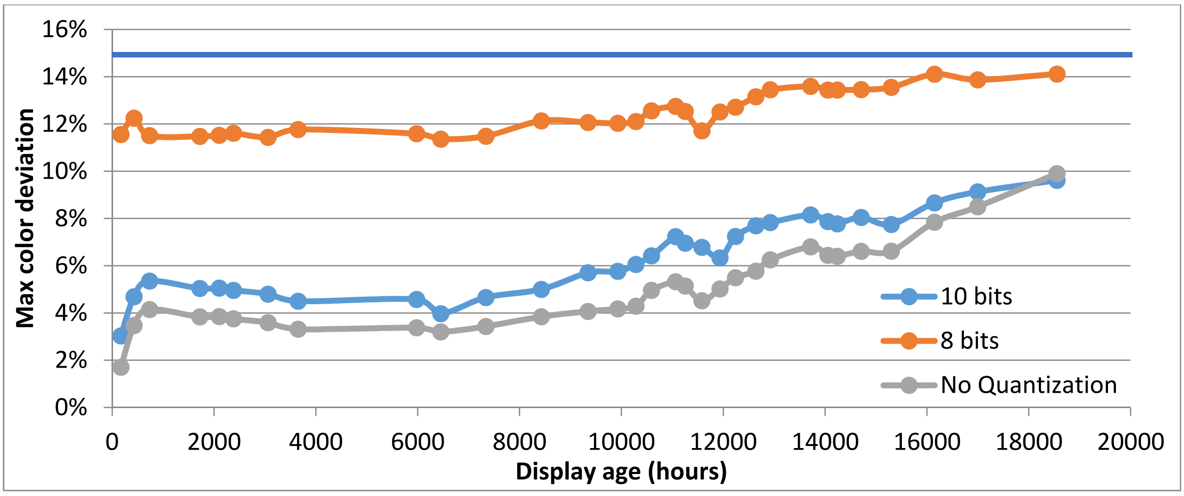

Figure 22: Effect of the aging of a diagnostic display on the color compliance of a calibration calculated at its production.

A calibration has been calculated based on the first measurements, and evaluated over the complete dataset. Figure 22 presents the results of these tests and shows that aging has a limited impact on the color compliance of the calibration.

C.5. Ambient light

The present study also takes into account ambient light chromaticity by considering a lighting color temperature of $5000K$ (D50).

Impact of the ambient light is modelled as an offset applied on the XYZ output of the display model. This offset is defined as follows with I being the illuminance level (in lux) and 0.01 the reflection coefficient of the display.

$$Y_{amb} = 0.01 * I$$

This offset differs from X and Z channel, according to the proportion of X Y and Z of White D50 $(0.96422,\ 1.0,\ 0.82521)$:

$$X_{amb} = 0.96422 * Y_{amb}$$

$$Z_{amb} = 0.82521 * Y_{amb}$$

C.5.1. Why considering the ambient light?

Ambient light partly reflects on any surface, including displays. The proportion of reflected light mainly depends on the material and reflecting surfaces geometries. This is usually characterized by a Reflection Coefficient associated with the display.

On certain medical displays, depending on the manufacturer and model, this coefficient may be higher than on consumer level displays because of the presence of a front glass adding two more interfaces (air-glass and glass-air) on top of the air-panel interface, creating even more reflections.

For this reason we decided to use a reflection coefficient of 0.01 in our simulations. In other words, we consider that the display reflects 1% of the ambient light. Figure 23 shows how the additional light from the reflection can adversely affect the grayscale part of a calibration if ambient light's effect is not taken into account.

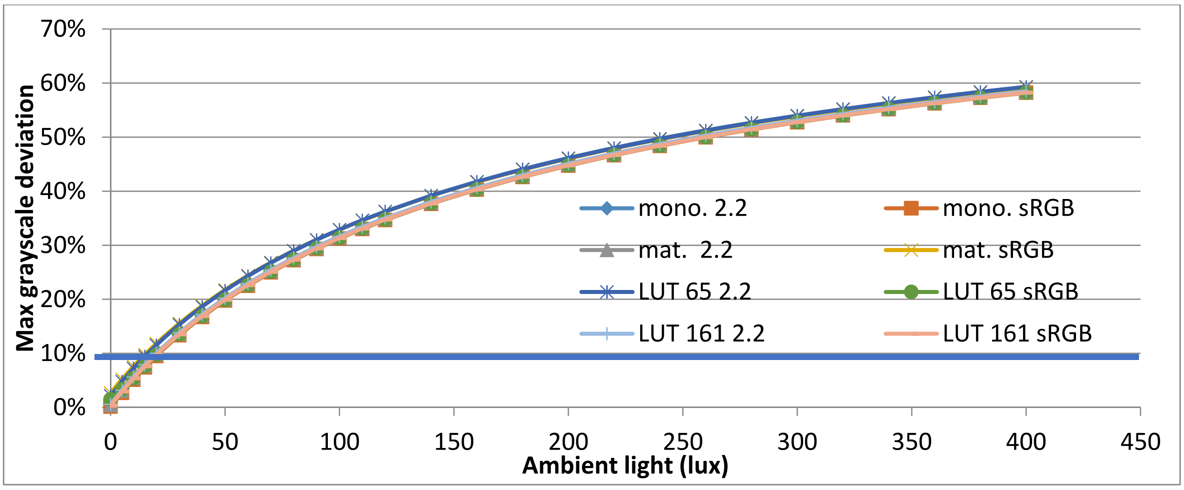

Figure 23: Effect of the ambient light on the grayscale compliance of a calibration evaluated with different type of display profiles.

It appears on Figure 23 that the effect of ambient light on the calibration is not linear. For this reason, it is important to detail its impact in different environments.

C.5.2. Diagnostic rooms

Diagnostic reading rooms are already used when establishing a diagnostic from quantitative imaging modalities, X-rays and other grayscale modalities where lighting conditions are controlled and illumination maintained low (2 to $10\ lux$ for x-rays, 15 to $60\ lux$ for CT and MR) [29]. In these conditions, knowing precisely the ambient light has its importance. Figure 24 shows how grayscale compliance varies with the ambient light while profiles were built considering an illumination of $5\ lux$.

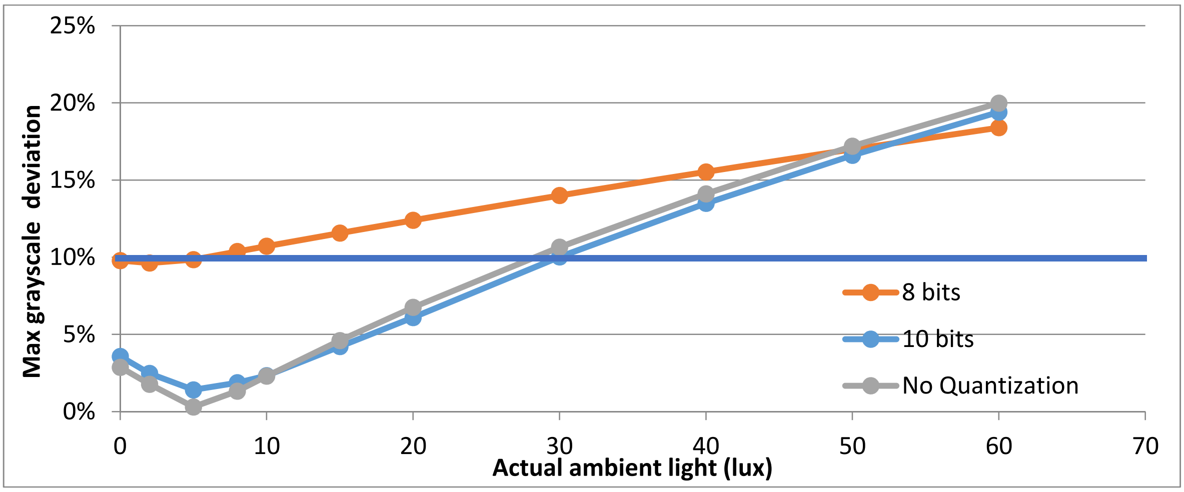

Figure 24: Effect of the ambient light on grayscale compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{5\mathit{lux}}$ with an $\mathbf{18 * 18 * 18}$ color LUT

If the lighting conditions are correctly controlled (no windows…) it is possible to assess a correct calibration by having a single estimation of the ambient light at the profile generation time and monitoring ambient light afterwards may not be required.

Figure 25 presents the impact of ambient light on the calibration of colors.

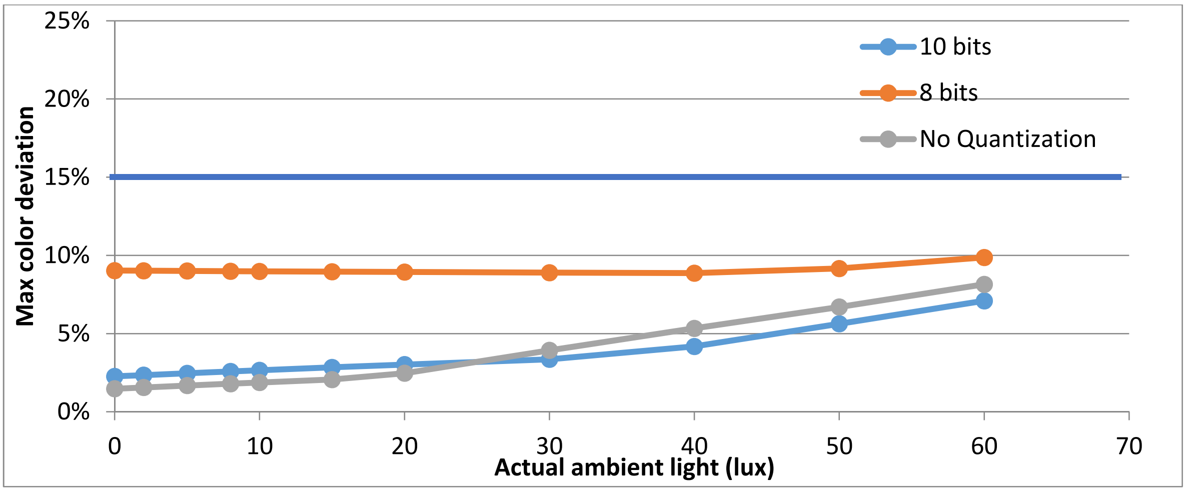

Figure 25: Effect of the ambient light on color compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{5}\boldsymbol{lux}$ with an $\mathbf{18} * \mathbf{18} * \mathbf{18}$ color LUT

It is clear here that color calibration does not suffer from an approximate estimation of the ambient light during the calibration process, at least for low illumination levels.

C.5.3. Staff offices

While quantitative imaging modalities are to be examined in dedicated reading rooms with reduced ambient light, pathology diagnostics are usually established in physician offices, where lighting conditions are not controlled and can vary from 50 to $180\ lux$ [29]. In such conditions, it is much more difficult to control the office's illumination as it highly depends on external parameters such as the weather which might abruptly and unpredictably change. It is therefore necessary to continuously measure ambient light and regenerate calibration profiles several times a day.

Figure 27 shows that higher relative variations of ambient light between the profile and the display it describes, have a larger influence on color compliance for staff offices than they do for diagnostic rooms. For instance, a 50% underestimation of the ambient light led to 9%, 3% and 2% maximal color deviations in diagnostic rooms, respectively for 8 bit, 10 bit, and floating point precisions. In the case of staff offices, a 45% underestimation already leads to 16%, 10% and 9% color deviations for 8 bit, 10 bit, and floating point precisions, respectively.

These variations of color calibration compliance remain rather limited in 10 bit systems compared to 8 bit architectures and could be considered as acceptable. However, this is not the case for grayscale calibration, as incorrectly estimating the ambient light by 30% is enough to make the calibration incompliant in an office (Figure 26), where the illumination is susceptible to drastically change throughout the day.

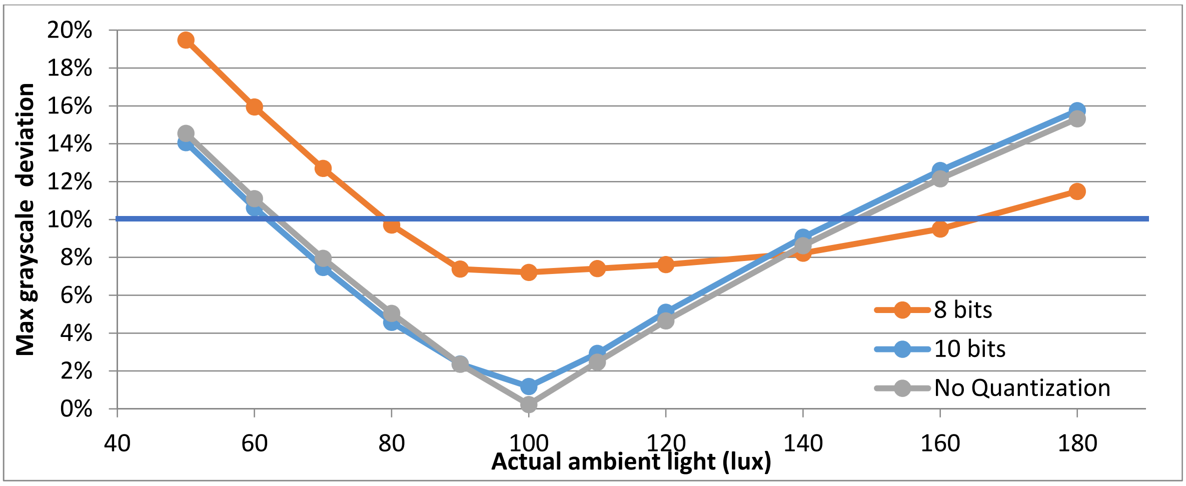

Figure 26: Effect of the ambient light on grayscale compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{100}\boldsymbol{lux}$ with an $\mathbf{18 * 18 * 18}$ color LUT

In order to preserve an accurate grayscale calibration using this method, ICC profiles would have to be regularly recreated according to the office's ambient light variations. If the presented method were used to obtain the most accurate grayscale calibration, staff offices would be improper for diagnostic purposes.

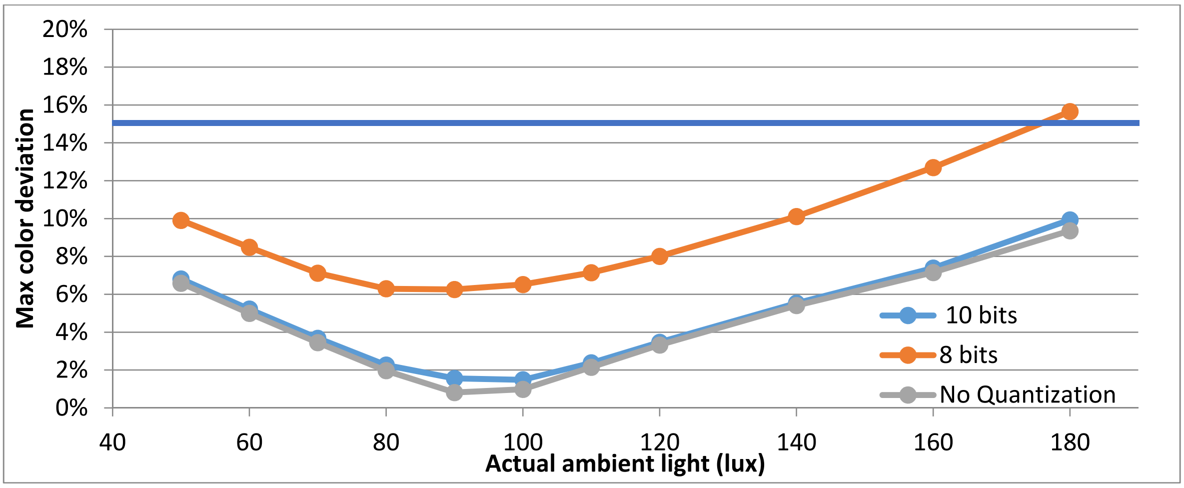

Figure 27: Effect of the ambient light on color compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{100}\boldsymbol{lux}$ with an $\mathbf{18 * 18 * 18}$ color LUT

C.5.4. Operating rooms

Color management and display calibration is also a concern for surgery. Reviewing scans and radios in an operating room happens and, in this case, DICOM GSDF calibration must also be respected.

AAPM estimates that operating room illumination usually varies from $300\ lux$ to $400\ lux$. This is quite high and can produce important reflections, especially on displays equipped with a front glass.

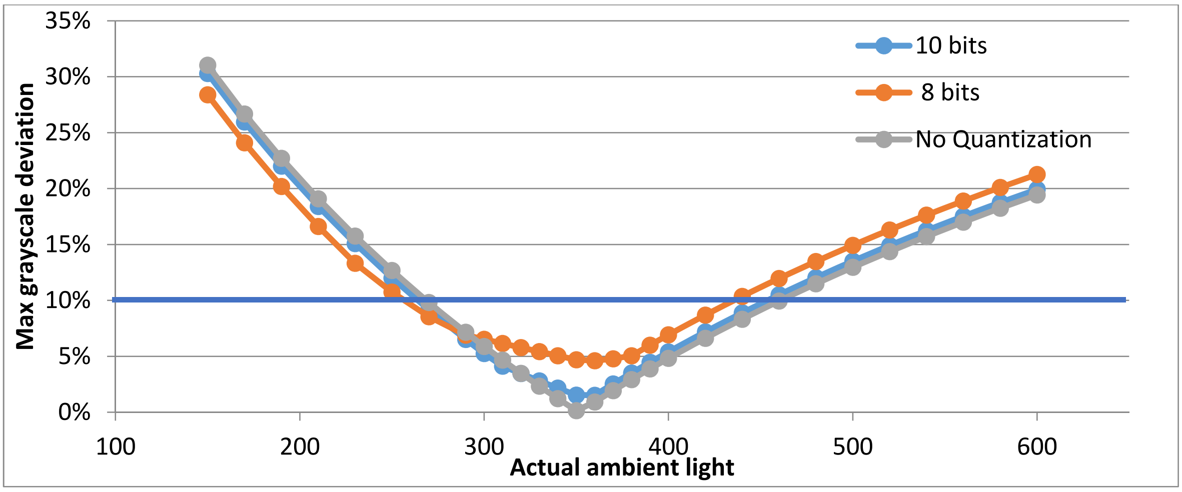

Figure 28: Effect of the ambient light on grayscale compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{350}\boldsymbol{lux}$ with an $\mathbf{18} * \mathbf{18} * \mathbf{18}$ color LUT

Figure 28 and Figure 29 show that ambient light variation in these conditions has a lower impact on the calibration compliance than it does in diagnostic rooms or staff offices. Relative variation appears to be similar: a 42% underestimation of ambient light by the profile leads to 15.5%, 9.3% and 8.6% color deviations for 8 bit, 10 bit and floating point precisions.

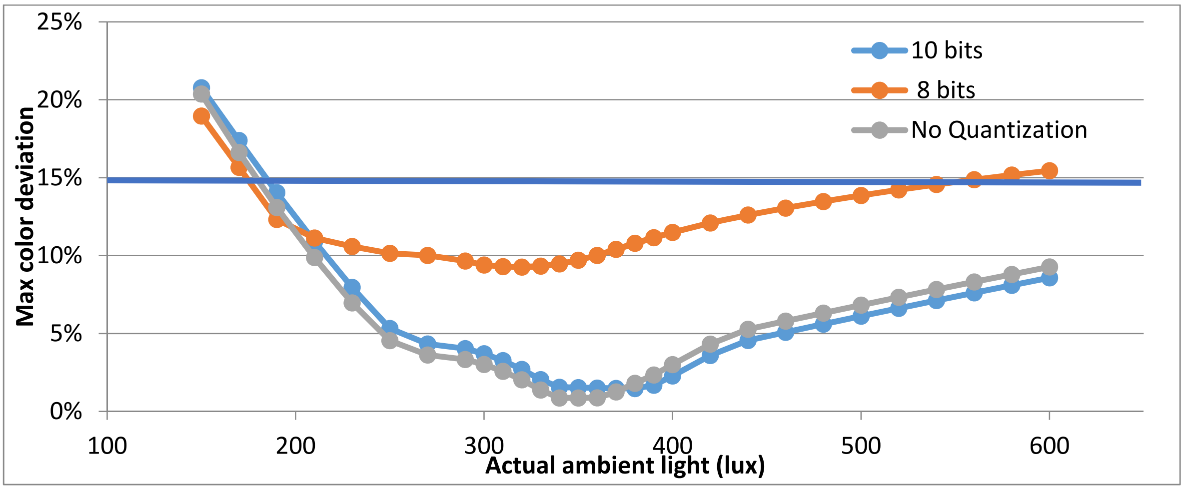

Figure 29: Effect of the ambient light on color compliance when ICC profiles used for calibration are built for an illumination of $\mathbf{350}\boldsymbol{lux}$ with an $\mathbf{18} * \mathbf{18} * \mathbf{18}$ color LUT

However, influence of absolute variation has far less impact than in other use cases, but the presence of very powerful directional light sources can be a concern for the quality of the color calibration.

Bibliography

1. NEMA, Digital imaging and communications in medicine (DICOM), part 14: Grayscale Standard Display Function, vol. PS 3.14, National Electrical Manufacturers Association, 2001.
2. Food and Drug Administration, "Display Devices for Diagnostic Radiology: Guidance for Industry and Food and Drug Administration Staff," 2022.
3. A. Badano, C. Revie, A. Casertano, W.-C. Cheng, P. Green, T. Kimpe, E. Krupinski, C. Sisson, S. Skrøvseth, D. Treanor, P. Boynton, D. Clunie, M. Flynn, T. Heki, S. Hewitt, H. Homma, A. Masia, T. Matsui, B. Nagy, M. Nishibori, J. Penczek, T. Schopf, Y. Yagi and H. Yokoi, "Consistency and Standardization of Color in Medical Imaging: a Consensus Report," Journal of Digital Imaging, vol. 28, no. 1, pp. 41-52, February 2015.
4. NEMA, DICOM supplement 100: Color Softcopy Presentation State Storage SOP Classes, 2005.
5. A. A. Marghoob for the International Skin Imaging Collaboration Melanoma Project Working Groups, "Standards in Dermatologic Imaging," JAMA Dermatology, vol. 151, no. 18, pp. 819-821, 2015.
6. L. Silverstein, S. Hashmi, K. Lang and E. Krupinski, "Paradigm for achieving color-reproduction accuracy in LCDs for medical imaging," Journal of the Society for Information Display, vol. 20, pp. 53-62, 2012.
7. W. C. Revie, M. Shires, P. Jackson, D. Brettle, R. A Cochrane and D. A Treanor, "Color Management in Digital Pathology," Analytical Cellular Pathology, 2014.
8. E. Krupinski, L. Silverstein, S. Hashmi, A. Graham, R. Weinstein and H. Roehrig, "Impact of Color Calibration on Breast Biopsy Whole Slide Image Interpretation Accuracy and Efficiency," in Breast Imaging, vol. 8539, pp. 744-748, Springer, 2014,.
9. ICC, "White Paper 46: Improving Color Image Quality in Medical Photography," International Color Consortium, 2017.
10. E. Krupinski, "The importance of perception research in medical imaging," Radiation Medicine, vol. 18, pp. 329-334, 2000.
11. P. Barten, "Physical model for the contrast sensitivity of the human eye" in Proc. SPIE 1666, p. 57-72 1992.
12. P. Barten, "Spatio-temporal model for the contrast sensitivity of the human eye and its temporal aspects," in Proc. SPIE 1913, Human Vision, Visual Processing, and Digital Display IV, 1993
13. K. A. Fetterly, H. R. Blume, M. J. Flynn and E. Samei, "Introduction to grayscale calibration and related aspects of medical imaging grade liquid crystal displays.," Journal of Digital Imaging, vol. 21, no. 2, pp. 193-207, 2008.
14. E. A. Krupinski and H. Roehrig, "The influence of a perceptually linearized display on observer performance and visual search" Academic Radiology, vol. 7, pp. 8-13, 2000.

15. S. Zabala-Travers, M. Choi, W.-C. Cheng and A. Badano, "Effect of color visualization and display hardware on the visual assessment of pseudocolor medical images," Medical Physics, vol. 42, no. 6, pp. 2942-2954, 2015.
16. C. Moler, "Origins of Colormaps," MathWorks, 2 2 2015. [Online]. Available: https://blogs.mathworks.com/cleve/2015/02/02/origins-of-colormaps/. [Accessed 3 2 2023].
17. Lissner and P. Urban, "Toward a Unified Color Space for Perception-Based Image Processing," IEEE Transactions on Image Processing, vol. 21, no. 3, pp. 1153-1168, March 2012.
18. IEC 61966-2-1:1999, "Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB" IEC, 1999.
19. J. R. Nuñez, C. R. Anderton and R. S. Renslow, "Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data," PLoS ONE 13(7): e0199239, 2018.
20. ISO/CIE 11664-6:2022 Colorimetry - Part 6: CIEDE2000 Colour-difference formula, ISO, Geneva
21. T. Kimpe, J. Rostang, G. Van Hoey and A. Xthona, "WE-D-204-04: Color Standard Display Function (CSDF): A Proposed Extension of DICOM GSDF," Medical Physics, vol. 42, no. 6, pp. 3670-3671, 2015.
22. NEMA, "DICOM PS3.6: Well-Known Color Palettes (Normative)," 2023. [Online]. Available: https://dicom.nema.org/medical/dicom/current/output/chtml/part06/chapter_B.html. [Accessed 3 2 2023].
23. A. Mikhailov, "Turbo, An Improved Rainbow Colormap for Visualization," 20 8 2019. [Online]. Available: https://ai.googleblog.com/2019/08/turbo-improved-rainbow-colormap-for.html. [Accessed 3 2 2023].
24. T. Shibutani, M. Onoguchi, Y. Hiroto, K. Nakajima, S. Matsuo, Y. Bamba and S. Saito, "Impact of calibration methods for color medical displays on interpreting brain SPECT images," Medical Physics, Vol 46, pp 2580-2588, 2019
25. F. Chesterman, H. Manssens, C. Morel, G. Serrell, B. Piepers and T. Kimpe, "Interpretation of the rainbow color scale for quantitative medical imaging: perceptually linear color calibration (CSDF) versus DICOM GSDF," Proc. SPIE 10136, Medical Imaging, 2017.
26. A. Avanaki, K. Espig, T. Kimpe, A. Xthona, B. Piepers and O. Vanovermeire, "Accuracy of Estimating PET Activity Depends on Medical Display Color Characteristics," in Society for Imaging Informatics in Medicine, 2018.
27. ICC.1:2010, Image technology colour management -Architecture, profile format, and data structure
28. ISO 14861:2015, "Graphic technology - Requirements for colour soft proofing systems," ISO, Geneva
29. E. Samei, A. Badano, D. Chakraborty, K. Compton, C. Cornelius, K. Corrigan, M. Flynn, B. Hemminger, N. Hangiandreou, J. Johnson, D. Moxley-Stevens, W. Pavlicek, H. Roehrig, L. Rutz, S. Shepard, R. Uzenoff, J. Wang and C.

Willis, "Assessment of display performance for medical imaging systems", Medical Physics, Vol. 32, pp 1202-1225, 2005

30. IEC 62563-1:2009, "Medical electrical equipment, Medical image display systems, Part 1: Evaluation methods," IEC, Geneva, 2009.
31. P. J. Green, "A smoothness metric for colour transforms," in Color Imaging XIII: Processing, Hardcopy, and Applications, 2008.
32. A. Avanaki, K. Espig, A. Xthona and T. Kimpe, "WE-D-9A-07: SRGB Displays: A Good Choice for Medical Color Images?," Medical Physics, vol. 41, no. 6, pp. 501-501, 2014.
33. N. Odlum, G. Spalla, N. Van Assche, B. Vandenberghe, R. Jacobs, M. Quirynen and C. Marchessoux, "Preliminary display comparison for dental diagnostic applications," in Proc. SPIE 8318 Medical imaging, 2012.
34. A. R. Avanaki, K. S. Espig, S. Sawhney, L. Pantanowitz, A. V. Parwani, A. Xthona and T. R. L. Kimpe, "Aging display's effect on interpretation of digital pathology slides," in Proc. SPIE 9420, Medical Imaging, 2015
35. D. S. Hirschorn, E. A. Krupinski and M. J. Flynn, "PACS displays: how to select the right display technology," Journal of the American College of Radiology, vol. 11, no. 12, pp. 1270-1276, 2014.