Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Diagnostic and clinical grade medical imaging display performance evaluation and optimization using quality assurance protocols

Master's Final Degree Project

Mindaugas Džiugelis
Project author

Prof. Dr. Diana Adlienė
Supervisor

Leonid Krynke
Adviser

Kaunas, 2021

Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Diagnostic and clinical grade medical imaging display performance evaluation and optimization using quality assurance protocols

Master's Final Degree Project

Medical Physics (6213GX001)

Mindaugas Džiugelis

Project author

Prof. Dr. Diana Adlienė

Supervisor

Prof. Dr. Liutauras Marcinauskas

Reviewer

Kaunas, 2021

Kaunas University of Technology

Faculty of Mathematics and Natural Sciences

Mindaugas Džiugelis

Diagnostic and clinical grade medical imaging display performance evaluation and optimization using quality assurance protocols

Declaration of Academic Integrity

I confirm the following:

1. I have prepared the final degree project independently and honestly without any violations of the copyrights or other rights of others, following the provisions of the Law on Copyrights and Related Rights of the Republic of Lithuania, the Regulations on the Management and Transfer of Intellectual Property of Kaunas University of Technology (hereinafter – University) and the ethical requirements stipulated by the Code of Academic Ethics of the University;

2. All the data and research results provided in the final degree project are correct and obtained legally; none of the parts of this project are plagiarised from any printed or electronic sources; all the quotations and references provided in the text of the final degree project are indicated in the list of references;

3. I have not paid anyone any monetary funds for the final degree project or the parts thereof unless required by the law;

4. I understand that in the case of any discovery of the fact of dishonesty or violation of any rights of others, the academic penalties will be imposed on me under the procedure applied at the University; I will be expelled from the University and my final degree project can be submitted to the Office of the Ombudsperson for Academic Ethics and Procedures in the examination of a possible violation of academic ethics.

Mindaugas Džiugelis

Džiugelis Mindaugas. Diagnostic and clinical grade medical imaging display performance evaluation and optimization using quality assurance protocols. Master's Final Degree Project; supervisor Prof. dr. Diana Adlienė; Faculty of Mathematics and Natural sciences, Kaunas University of Technology.

Study field and area (study field group): Health Sciences, Medical technologies (G09).

Keywords: medical display, quality control, monitor calibration, GSDF.

Kaunas, 2021. 57 pages.

Summary

Even though the performance of diagnostic displays used for primary interpretation of medical images has been well - characterized, lack of testing and knowledge are the main leading factors to their poor performance. There are two main categories of medical-grade displays – primary and secondary. Diagnostic displays belong to the primary display category and therefore are the most important, having the strictest requirements. Displays belonging to the secondary category still require quality assurance, but the requirements for their performance are not as strict (these are modality, clinical review, and electronic health report type displays). For such monitor quality assurance, performance monitoring is provided by the American Association of Physicists in Medicine (AAPM) protocol Task Group 270 together with previously released Task Group 18. Such protocols provide series of instructions, including image patterns with instructions on how to perform such tests.

Džiugelis Mindaugas. Diagnostinių ir klinikinių medicininių monitorių veikimo įvertinimas ir optimizavimas naudojant monitorių kokybės užtikrinimo protokolus. Magistro baigiamasis projektas. Vadovė prof. dr. Diana Adlienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Sveikatos mokslai, medicinos technologijos (G09).

Reikšminiai žodžiai: medicininiai monitoriai, kokybės kontrolė, monitorių kalibravimas, GSDF.

Kaunas, 2021. 57 p.

Santrauka

Nors diagnostinio tipo ekranų, naudojamų pirminiam medicininių vaizdų aiškinimui, našumo parametrai ir yra gerai aprašyti, tyrimų ir žinių stoka yra pagrindiniai veiksniai, lemiantys jų prastą veikimą. Yra dvi pagrindinės medicininių monitorių kategorijos - pirminiai ir antriniai. Diagnostiniai ekranai priklauso pirminiams, todėl yra svarbiausi - jų veikimui turi būti užtikrinami griežčiausi standartai. Antrinei kategorijai priklausantiems ekranams taip pat reikalinga kokybės kontrolė (šiai kategorijai priklauso modaliniai, klinikinės apžvalgos ir elektroninių sveikatos ataskaitų tipo ekranai). Tokiai monitorių kokybės krontrolei, jų veikimo stebėsenos rekomendacijas stebėjimą teikia Amerikos medicinos fizikų asociacijos (AAPM) protokolas nr. 270, kartu su ankstesniu, nr. 18. Šiuose protokoluose pateikiamos instrukcijos, įskaitant vaizdinius modelius su monitorių kokybės kontrolės bandymų paaiškinimais.

Acknowledgments

I sincerely appreciate my supervisor, prof. Dr. Diana Adlienė, Assoc. Prof. Dr. Birutė Gricienė, and Kirill Skovorodko for their time for consultations, scientific ideas, advice, and insights in preparing this master's degree thesis. I would also like to thank Leonid Krynke, who has always found time for his advice, insights, and consultations during this work preparation.

Table of contents

List of abbreviations

DICOM – Digital Imaging and Communications in Medicine.

GSDF – Greyscale Standart Display Function.

LCD – Liquid Crystal Display.

LED – Light-Emitting Diode.

OLED – Organic Light-Emitting Diode.

IPS – In-Plane Switching.

TN – Twisted Nematic.

VA – Vertical Alignment.

FFS – Fringe Field-Switching.

LC – Liquid Crystal.

PACS – Picture Archiving and Communication System.

ICU – Intense Care Unit.

CT – Computed Tomography.

AAPM – American Association of Physics in Medicine.

EFOMP – European Federation of Organisations For Medical Physics.

EHR – Electronic History Record.

TG – Task Group.

COTS – Commercial Off-The-Shelf.

CRT – Cathode Ray Tube.

TADF – Thermally Activated Delayed Fluorescence.

JND – Just-Noticeable Difference.

DDL – Display Driving Level.

LUT – Lookup Table.

GPU – Graphics Processing Unit.

LUDM - Luminance Deviation from the Median.

Introduction

In the medical field, digital imaging technologies can be divided into different categories, such as scanning, transcoding, receiving, and displaying. X-ray machines, CT, MRI, PET/SPECT scanners could be excellent examples of scanning devices. Computers are helping to transcode scanned information from imaging devices by performing specific calculations and algorithms [1]. For image transferring and receiving, computers, networking systems, and imaging modalities are involved. [2] Finally, medical display systems (monitors, in short) are used for such image observations, interpretations, diagnosing, etc. Such monitors are generally classified into diagnostic, modality, clinical review, and electronic health record displays [3].

Diagnostic displays are crucial for patient diagnosis interpretation; therefore, they must have quality assurance programs to monitor their performance. As Lithuanian Hygiene Norm HN 78:2015, American Association of Physicists in Medicine (AAPM) [3] and European Federation of Organisations for Medical Physics (EFOMP) [4] state, medical physicists should perform daily quality assurance for medical-grade display systems. One of the better examples of such performance monitoring is provided by the AAPM protocol Task Group 18, which in 2019 got succeeded by Task Group 270 [5]. Additionally, the DIN Standard for Quality Assurance of Diagnostic Displays, called DIN 6868-157, by initial analysis, is very similar to TG-18, created by AAPM and DIN 6868 57, NY PDM, NYC PDM, ACR, etc. Such protocols provide series of instructions, including image patterns with instructions on how to perform such tests. These tests could be generally divided into quantitative and visual tests. Quantitative tests require a photometer and light meter, while visual tests are for visual evaluation only [3]. For qualitative measurements, it is required to have a loupe and optionally reflector[3, 5]. These tests monitor track medical display performance and suitability over time and compare different monitor performance.

The main problem is that medical grade display performance is not adequately evaluated in Lithuania, and this can lead radiologists to diagnostic evaluation problems.

The main aim of the work is to optimize diagnostic and clinical display performance parameters in the hospital according to display QA protocol recomendations.

The main tasks of this work are:

1. To calibrate monitors by changing display driving level values;
2. To evaluate the changes of each display performance by comparing its parameter results before and after the calibration;
3. To suggest recomendations for optimal monitor performance based on results.

1. Literature review

1.1. Medical grade displays and their classification

For reviewing radiology examinations, radiologists typically use medical-grade displays. Such displays are preferred due to their resolution, luminance, contrast ratio, noise parameters, and Grayscale Standard Display Function (GSDF), allowing to see more grey shades than typical monitor [3]. There are two main categories of medical-grade displays known – primary and secondary; however, according to newly released AAPM (American Association of Medical Physicists) report 270, there are currently four main display categories of displays: diagnostic (primary), modality (secondary), clinical review (secondary), and electronic health report (EHR) (secondary) displays, as there was a need to improve QA criteria due to display technological advancements [5].

AAPM report 270 provides guidelines and recommendations to periodical tests and optimizations of monitor performance, evaluating its state, and tracking changes in its condition over time. These tests consist of luminance tests (minimum and maximum), luminance response, luminance uniformity, veiling glare, environmental conditions testing, reflective monitor characteristic checks, visual quality checks, etc.

There are specific passing criteria for each type of monitor for such tests to be passed [5].

1.1.1. Diagnostic Display Category

The diagnostic display category is analogous to a previously existed "primary" display category defined in the TG18 AAPM report [6]. Because diagnostic monitors are mainly used for image interpretation, they are usually found in so-called "reading rooms" (such rooms usually are intended for image interpretation and have specific requirements for ambient light in the room, reflection from the monitor) featuring carefully chosen lighting and ergonomic considerations. Some authors suggest allowing not as strict requirements to ambient light due to eyestrain, which could impact more than the ideal performance of a display [7]. When determining how to define the desired output of any monitor used for primary analysis or routine medical diagnosis or decision making, users can also use the specialized diagnostic screen category (e.g., clinical or modality displays). Usually, diagnostic category displays have the strictest requirements as such devices are used for primary diagnostic image interpretation [3].

1.1.2. Modality Display Category

Modality displays are usually used while acquiring and generating medical images. Such monitors are not primary and are connected to some diagnostic imaging device (i.e., CT scanners, C-arms, etc.). Even though these monitors are not primary, they are still essential, as they are used for image reconstruction and image processing with some advancements, regions of interest drawing. Radiology technologists and doctors could use such displays in interventional radiology during fluoroscopy operations (especially cardiac operations) [8]. Because modality displays give direct feedback to the operator, their performance can affect operation procedures. Therefore, they are classified as very important (only less critical than diagnostic displays) and have relatively high requirements in testing [3].

1.1.3. Clinical Review Display Category

Clinical review displays are being used in such clinical environments as intense care units (ICUs) or emergency rooms. If there are no suitable conditions for radiologists to evaluate images correctly, this type of monitor will be used. Because of the possible direct impact on the clinical environment, such displays are essential to check from time to time. However, AAPM provides less strict requirements for it [3].

1.1.4. Electronic Health Record Display Category

This type of monitor is used for general patient radiological image viewing, acquired with the radiologist's description. Surgery planning or patient health history checks are excellent examples of such image reviewing. Such displays are equivalent to good quality commercial multipurpose monitors for everyday use. However, they still require to be tested, as sometimes the poor performance of such monitors may affect efficient work [3, 5].

1.2. COTS (Commercial-Of-The-Shelf) Monitors

In recent years, the technology for commercial displays changed, which lead them to better overall performance. As this is the case, more and more radiology departments are started using (or considered using) such displays for clinical purposes. There were numerous published studies in which diagnostic accuracy was measured of such monitors and compared them to diagnostic medical-grade monitors. It turned out that better performing COTS monitors might perform similarly to diagnostic grade monitors or even better when calibrated [9].

1.3. Display Technologies

As the AAPM presented the TG18 report, most of the displays for medical imaging were still using CRT (cathode ray tube) technology. As its name suggests, the CRT monitor was based on cathode ray tubes, a vacuum tube containing several electron guns, generating images by deflecting electron beams to the screen surface [10].

Most monitors use LCD (liquid crystal display) panel type, which uses liquid crystal cells [9]. LCD panels have multiple types, and it is crucial to understand their qualities, advantages, and disadvantages primarily due to the applicable methods and limitations of calibration.

1.3.1. LCD Monitor Technology

An LCD monitor is a flat panel display that mainly uses liquid crystals, backlight, and polarizers to produce images. In general, a liquid crystal is a rod-shaped material, flowing like a liquid while retaining some of the properties that are found in the crystals. Applying an electric current to liquid crystals makes it possible to change their structure, which allows to alter polarized light [9].

Every pixel in the LCD screen is divided by a layer of indium-tin-oxide, including parallel and perpendicular polarizers. If no liquid crystal is present, the light will be blocked by polarizers. However, due to the liquid crystal's unique chemical structure, which can change by applying electric current, controlled with light modulators (dependently on a monitor type), it is possible to control emitted light intensity.

There are four main types of LCDs - twisted-nematic (TN), vertical alignment (VA), fringe field switching (FFS), and IPS (in-plane-switching). Every one of them has its drawbacks and benefits.

Twisted-nematic (TN) panel is one of the first created LCD panel types [11]. Its technology is based on two plates which alignment changes between different electric field applications. There are two primary states of a display: voltage on when voltage is applied and off when there is no voltage. There is no electrical field in voltage off state applied to the liquid crystal; therefore, liquid crystal molecules are in a helical structure. When there is a helical structure present in liquid crystals, a backlight can pass through the polarizers because the light is twisted by 90 degrees due to the alignment layers, which arrange liquid crystals in such way. When there is a 'voltage on' state, the electric field between electrodes is applied; therefore, crystals change their structure, which impacts light twisting. As a result, the light will not pass through a polarizer. By decreasing the voltage, the light twisting gets closer to 90 degrees; therefore, more light will pass through a polarizer. Such technology is power-efficient, simple, and fast. As a result, such monitors have an excellent response time. However, such alignment affects viewing angles, as light passes only at a specific angle through the polarizer and limits color reproduction [11].

The visual alignment (VA) panel is invented in 1971 by Schiekel and Fahrenschon [11]. At first, the LC's are aligned vertically between polarizers. When there is no voltage, light travels through the first polarizer and refracts at a specific angle. Because there is no voltage applied, light does not get manipulated by LC's. Therefore, the frontal layer of the polarizer does not allow the light to pass through the frontal polarizing layer. As voltage is applied, LC's (as its material always has negative dielectric anisotropy) tilt the light closer to the original vertical orientation, allowing to pass more light through the polarizers. Due to specific polarizer angles and light tilt, there will be some birefringence (or material causing refraction, depending on polarization and light propagation). That is to say, the contrast ratio is reduced by the dispersion of color filter pigments or by the unintentional light output caused by the polarizers, or even possibly by electrode edges. This also means that display uniformity and color consistency might not be as great as other display types [12].

In-plane switching (IPS) is one of the most popular display types due to its production cost and image quality. It's a technology that solves the two critical problems with a regular twisted nematic (TN) type panel: color accuracy and viewing angles. In this panel type, both linear polarization filters (polarizer and analyzer) (see figure 1) have their transmission axes pointing in the same direction. The inner surfaces of glass substrates are handled to match the surrounding LC molecules at a correct angle to achieve a 90 degree TN structure of the LC layer in-between glass plates without applied voltage (OFF state). Such molecular structure is very close to TN panels TFT electrodes are arranged differently. They create an electric field parallel to the plate as they (electrodes) are in the same plane and on a single glass plate. The LC layer is thin compared to the gap between the electrodes. LC molecules have positive dielectric anisotropy, and therefore are parallel to an applied voltage. They align themselves with its long axis.

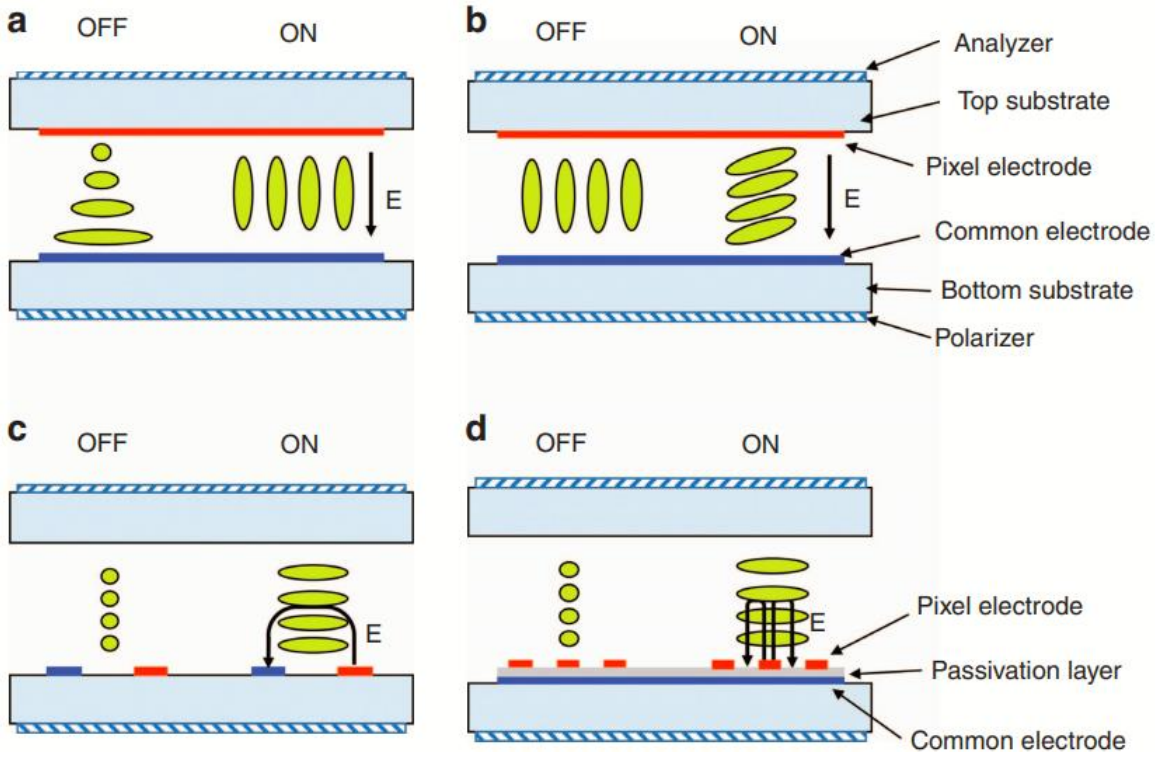

Figure 1. Schematic diagram of TN (a), VA (b), IPS (c) and FFS (d) LCD technologies [11]. In the "ON" state, liquid crystal molecules are adjusted by passing the electric field so that light can pass through the presented screen layers.

In the OFF state (left), the polarizer linearly polarizes an entering light. The rotated TN LC layer twists the moving light's polarization axis by 90 degrees, ensuring no light gets through a polarizer. In the ON state, an appropriate voltage is generated between the electrodes, resulting in electrical field production, which adjusts LC molecules as presented on the right side of figure 1 (c part). Light will pass via polarizer and analyzer in this case. Other implementation schemes exist with a particular LC molecule structure – for example, no twisting in the off state. Both electrodes take up more area than TN matrix electrodes since they are on the same plate. The contrast and brightness are also reduced as a result of this [13].

Fringe field switching (FFS) panel is one of the latest LCD technologies found by Korean scientists in 1998. Such a panel has high contrast, large viewing angles, low discoloration, good response time, etc. [11]. The FFS panel seems to have the same operational concept as IPS, but a small passivation layer separates the common electrodes and pixels, meaning that the width and distance of the electrode can be much narrower than that of IPS and that the fringe fields are much more robust, covering the regions of electrodes and gaps. Therefore, the empty area is smaller, resulting in the possibility to make screens with higher pixel density.

1.3.2. Organic Light-Emitting Diode (OLED) Technology

As OLED displays are getting in the medical field more and more popular, it is essential to understand its technology as well. OLED screens are unique due to different display technology, which does not require a backlight. It is made of organic stacks which are surrounded by anode and cathode. Electrons with holes move from electrodes to organic material, which recombines and emits light. OLED displays consist of multiple layers. Emitting layer, made from dopant and host organic materials, emits light very efficiently and fast due to carrier mobility. The electron transport layer help to bring

electrons to the light-emitting layer, same as the hole transmitting layer, transferring holes. Both hole and electron layers can carry electrodes and holes by applying voltage to their ends; therefore, they are placed between electrodes. After activating hole and electron transporting layers, which activate conductors located in injection layers, electron and hole carriers move to the organic layer, causing it to glow.

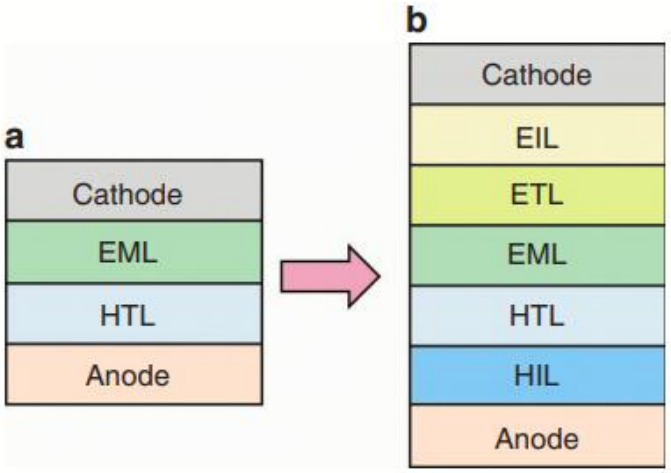

Figure 2. Schematic diagram of OLED (organic light-emitting diode) display layers. Scheme a represents the basic OLED screen structure proposed by Tang and VanSlyke. Scheme b represents current OLED screen technology. EIL – electron injection layer, ETL – electron transporting layer, EML – emitting layer, HTL – hole transporting layer, HIL – hole injection layer [14].

There are different materials, structures, and designs used to manipulate an image in an OLED display. For example, OLED displays can be divided into four main categories, based on inside EML layer existing emitting material chemical configuration.

Fluorescent OLED screens. When electrically excited, ¾ of the triplet organic stacks are formed, while the remaining ¼ part of the material is singlets. Those singlets are decaying radiatively through fluorescence in about one nanosecond. Such screens, therefore, are 25% efficient when it comes to organic material efficiency [14].

Triplet-triplet fluorescent OLED screens (TTL) are more efficient due to the fusion of two triplet excitons to form one singlet exciton. Other than that, this type of screen works very similarly to fluorescent OLED screens [14].

Phosphorescent OLED type screen working mechanism is based on phosphorescent emission process. With strong spin-orbital bonding, heavy metal atoms reduce triplet lifetime, leading to phosphorescent emission. A single exciton is passed to the threefold state to achieve a 100 percent IQE. Thanks to the extended life-cycle emission in phosphorescent OLED, the triplet can interact with other triplets and polarons, which leads to efficiency rollovers under high-current (triplet-triplet annihilation and triplet- polar annihilation, respectively). Such processes can cause hot excitons and warm polarons, especially for blue-emitting devices, to shorten operating life [14].

Thermally activated delayed fluorescent OLED has lower energy consumption due to lowered energies between singlets and triplets, minimizing the cost of energy exchange between particles. Recently, TADF materials quickly emerged and are expected to have more use in the future.

While OLED overall consumes less power (up to 60%), most of its savings come due to black levels, where no energy is used. When the OLED screen is white and fully illuminated, it can consume up to 3 times more energy than the LCD screen. Also, OLED screens are not as durable, as organic compounds degrade over time, resulting in burn-ins (areas where individual diodes due to constant use become not as bright or, in severe cases, fail) [14].

1.4. AAPM Display Quality Standards and Test Patterns

Despite the efforts to improve display quality by improving its technology, the standards to unify monitor quality are required to match its performance. Many different factors might affect display usability in the medical field. These standards can be divided into qualitative, when nothing is measured, just specific patterns are visually evaluated, and quantitative, when display parameters are measured and registered using a photometer or colorimeter.

The first criteria measured in medical-grade display evaluation is display luminance. It includes ambient illuminance, ambient luminance, maximum luminance, luminance ratio, and luminance response function (otherwise known as Greyscale Standart Display Function). Those parameters are dependant on each other. If one parameter fails, there is a high chance that others will fail as well (more detailed parameter dependence examples will be presented in the following sections).

Suggested limits of monitor luminance levels are published in AAPM newly released report TG 270, which set standards for display luminance criteria of primary (diagnostic mammography and non-mammographic diagnostic) and secondary (modality, clinical review, and electronic health record) displays:

Table 1. Recommended display luminance levels [3,5]

Display Type	Recommended values
Diagnostic (non-mammographic)	$L'_{min} \geq 1$ cd/m^2
	LR = 250-450
	$L'_{max} \geq 300$ cd/m^2
Diagnostic (mamographic)	$L'_{min} \geq 1.2$ cd/m^2
	LR = 250-450
	$L'_{max} \geq 350$ cd/m^2
Modality, Clinical review, and Electronic Health Record displays	$L'_{min} \geq 0.8$ cd/m^2
	LR = 250-450
	$L'_{max} \geq 250$ cd/m^2

Also, TG 270 suggests qualitative and quantitative QA test criteria and frequencies for primary (diagnostic) and secondary (Modality, Clinical review, and EHR displays) [3] (see tables 3 and 4):

Table 2. Display testing criteria (provided by AAPM TG 270 'Display quality assurance') [3].

Documented QA tests	Equipment	Patterns	Suggested passing criteria	
			Diagnostic	Modality, CR, EHR
Quantitative ambient luminance/illuminance	Photometer	Display off, N/A	AR (L_{amb}/L_{min}) < 1/4	AR (L_{amb}/L_{min}) < 1/4
			Illuminance 25 - 75 lux	
Qualitative ambient luminance/illuminance	None	TG270-sQC	Low-contrast features in darkest region visible in both no-light and normal light settings	
		TG270-pQC		
		TG18-QIQc		
Quantitative min/max luminance	Photometer	TG270-sQC	250<LR<450	250<LR<450
		TG270-ULN	$L'_{min} > 1.0$ cd/m^2	$L'_{min} > 0.8$ cd/m^2
		TG18-LN	$L'_{max} > 300$ cd/m^2	$L'_{max} > 250$ cd/m^2

Table 2 (continued). Display testing criteria (provided by AAPM TG 270 'Display quality assurance') [3].

Documented QA tests	Equipment	Patterns	Suggested passing criteria	
			Diagnostic	Modality, CR, EHR
Quantitative luminance response	Photometer	TG270-ULN TG18-LN TG270-sQC	Deviation from DICOM GSDF <10%	Deviation from DICOM GSDF <20%
Qualitative luminance response	None	TG270-sQC	All low contrast features visible under typical conditions	TG270-sQC: all ± 5 gray level patterns visible
		TG270-pQC		TG270-pQC: all ± 4 gray level patterns visible
		TG18-QIQc		
Quantitative uniformity	Photometer	TG270-ULN TG18-UL	LUDM <30% (if >15%, evaluate qualitatively and determine clinical impact)	
Qualitative uniformity	None	TG270-ULN TG18-UL	No non-uniformities that impact clinical use	
Qualitative noise	None	TG18-AFC TG270-ULN	No noise effects that affect clinical use	
Qualitative temporal resolution	Camera, photometer	TG270-TR	No temporal effects that affect clinical use	
Qualitative spatial resolution	Loupe	TG270-sQC	Pixel structure not visible at typical working distance, one-to-one pixel mapping from a graphics card	

Table 3. Suggested display quality assurance frequencies (provided by AAPM TG 270 'Display quality assurance'). Acceptance means that this test should be performed only for new monitors, which have not been used or evaluated [3].

Documented QA test	Display type			
	Diagnostic	Modality	Clinical review	Electronic Health Record
Qualitative Luminance Response	Quarterly	Quarterly	Annually	Annually
Qualitative Ambient Luminance/Illuminance	Quarterly	Annually	Annually	Annually
Qualitative Uniformity	Quarterly	Annually	Annually	Annually
Qualitative Spatial Resolution	Quarterly	Annually	Annually	Annually
Quantitative Min/Max Luminance	Annually	Annually	Annually	Annually
Quantitative Luminance Response	Annually	Annually	Annually	Acceptance
Quantitative Colour Assessment	Annually	Annually	Acceptance	Acceptance
Quantitative Ambient Luminance/Illuminance	Annually	Acceptance	Acceptance	Acceptance
Quantitative Uniformity	Acceptance	Acceptance	Acceptance	Evaluation
Qualitative Noise	Evaluation	Evaluation	Evaluation	Evaluation
Qualitative Temporal Resolution	Evaluation	Evaluation	Evaluation	Evaluation
Diffuse Reflection Coefficient (R_d)	Evaluation	Evaluation	Evaluation	Evaluation

1.4.1. Ambient Illuminance, Ambient Luminance and Reflection

In this work, ambient illuminance (L_{amb}) means the room's ambient lighting intensity. This parameter is important because ambient luminance cause reflections, which can negatively affect contrast,

luminance response, luminance ratio, and other related parameters [14][15]. It is measured in lux units by a luxometer.

Specular reflection produces a distinct reflection from the environment, which is mirror-like. This reflection type is common on reflective displays as they have a smooth surface, and light reflection does not get scattered [16].

Diffuse reflection scatters light uniformly in all directions [17]. The luminance of such scattering is proportional to illumination angle and viewing angle (usually, its intensity depends on the source). It increases reflection intensity (which will not be uniform), leading to a poor viewing experience. This is caused due to display polarizer degradation over time, leading to decreased contrast properties and image uniformity [16].

Haze is also a vital part worth mentioning as it might be present in both specular and diffuse reflections. As with specular reflection, its luminance spikes in the specular direction. However, unlike diffuse reflection, this reflection of the light is proportional to the incident light. Haze reflection occurs as a result of the antiglare matte top layer [16].

To avoid reflections, the positioning environment where medical displays exist is vital. The following step in the display luminance check-up is the minimum luminance check.

Ambient luminance (measured in cd/m^2 units) is the amount of light reflected from the surface of a display. The diffuse reflection coefficient of a display demonstrates the correlation between ambient light intensity and ambient luminance [8]. It shows how dark a monitor can get while showing images. If this requirement is ignored, results might be inaccurate after performing luminance response, especially in the first 50 gray levels (totally there are 1024 in diagnostic, 256 in the commercial type monitors) [3]. Displays which are used in a relatively bright environment are essential to meet the minimum luminance requirement. It allows noticing differences in small grey level changes. However, in bright environments maintaining low ambient light might be difficult obviously, therefore lights could be dimmed with proper angle adjustments of the displays themselves [3]. AAPM Report 270 suggests that ambient illuminance should be 25 - 75 lux in this case. Ambient luminance should stay underneath 0.2 cd/m^2 at such levels (based on the reflective features of the display) [3, 5]. Minimum display luminance Lmin after measuring is combined with ambient light and parameter L_{amb}, describing minimum observer's luminance. It represents the minimum luminance that the user from the display surface sees. The result is marked as L'_{min} and measured in cd/m^2.

Suppose ambient light requirements are not adequately met. In that case, display luminance response, luminance ratio, and maximum brightness will be measured inaccurately due to reflections (even if photometer is put directly on screen) that might go through a polarizer and ambient light gets scattered, making monitor, especially in dark levels, brighter than it is (figure 3 shows an example how display luminance might be affected by different levels of ambient light).

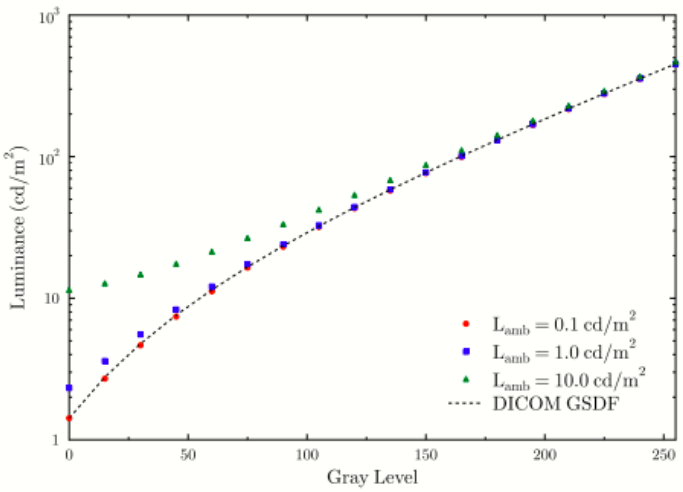

Figure 3. Effect on display luminance of various ambient light conditions. At darker greyscale points, where display luminance is lower when the ambient light level is much higher, environmental lighting effects are most noticeable in display luminance. The display was calibrated to DICOM GSDF when L_{amb} = 0.1 cd/m^2 in this case. This shows the importance of environmental conditions when diagnostic images are evaluated – it is impossible to properly calibrate display and see proper images when ambient light is high [3].

1.4.2. Luminance Ratio And Contrast Ratio

After measuring the minimum luminance of display, the same principle goes with maximum luminance L'_{max}, which equals the sum of L_{amb} and L_{max}, showing how bright a display can get.

Luminance ratio is a measure showing the difference between minimum L'_{min} and maximum luminance L'_{max}. It describes how different absolute brightest and absolute darkest grey levels are.

The minimum luminance ratio of 350 is recommended by AAPM TG 270 report. Such level is minimally supported by diagnostic monitors, which ensures to prevent washed-out effect on the screen (luminance response test, described in section 1.3.3, might be passed, but the difference of each grey level might become unnoticeable).

The contrast ratio is calculated by dividing monitor's minimum illuminance from maximum illuminance L_{max}/L_{min} measured by photometer in cd/m^2. To measure this, there are several test patterns: TG270-sQC and TG18-LN (first and last grey level).

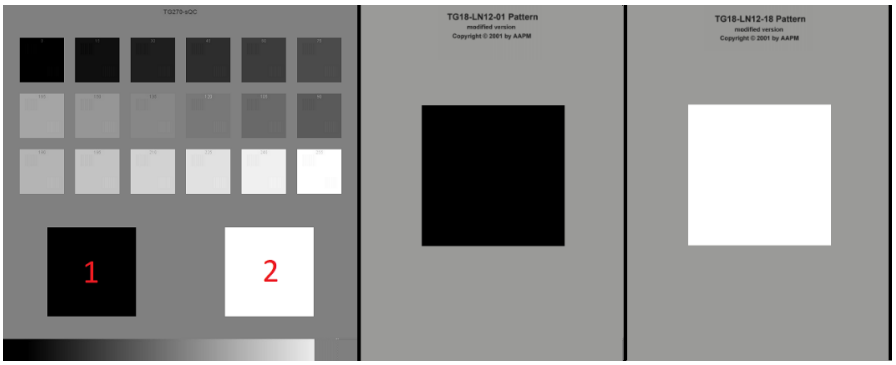

Figure 4. TG270-sQC, TG18-LN-01 and TG18-LN-18 patterns. In TG270-sQC patterns, marked 1 and 2 are equivalent to TG18-LN-01 and TG-18-LN-18 patterns. Black square shows minimum brightness of the display, whereas white square shows maximum [3,6].

1.4.3. Display Luminance Response and DICOM GSDF

Luminance response is one of the medical monitor tests, where a series of consistently increasing grey levels are measured by a photometer. It allows measuring different consistently increasing gray levels over the full 10-bit gray level scale. Luminance response could be measured by checking equal increments of gradually increasing greyscale brightness of 18 or 52 points, 256 points, and a whole pallet of 1024 points if needed.

Generally, the luminance response curve shows how consistent contrast change is between all set of measured grey levels. It is measured by measuring maximum error over all points, in percent. AAPM 270 report recommends using the DICOM (Digital Imaging and Communications in Medicine) GSDF (Grayscale Standard Display Function), which has been adopted as a standard in evaluating luminance response.

DICOM GSDF function is based on human eye contrast sensitivity [15] and is directly related to the ideal luminance GSDF standard. As human eye contrast sensitivity does not match computer contrast response (different shades of darker areas of an image for the human eye are harder to spot; therefore, typically dark shades of grey increments should be extended more than typical linear monitor display function), a new greyscale function has to be created. DICOM GSDF 3.14 standard (see figure 5) mathematically defines standards for greyscale display function. The primary purpose of such standard is to allow applications to measure the significance of measured monitor deviation from GSDF.

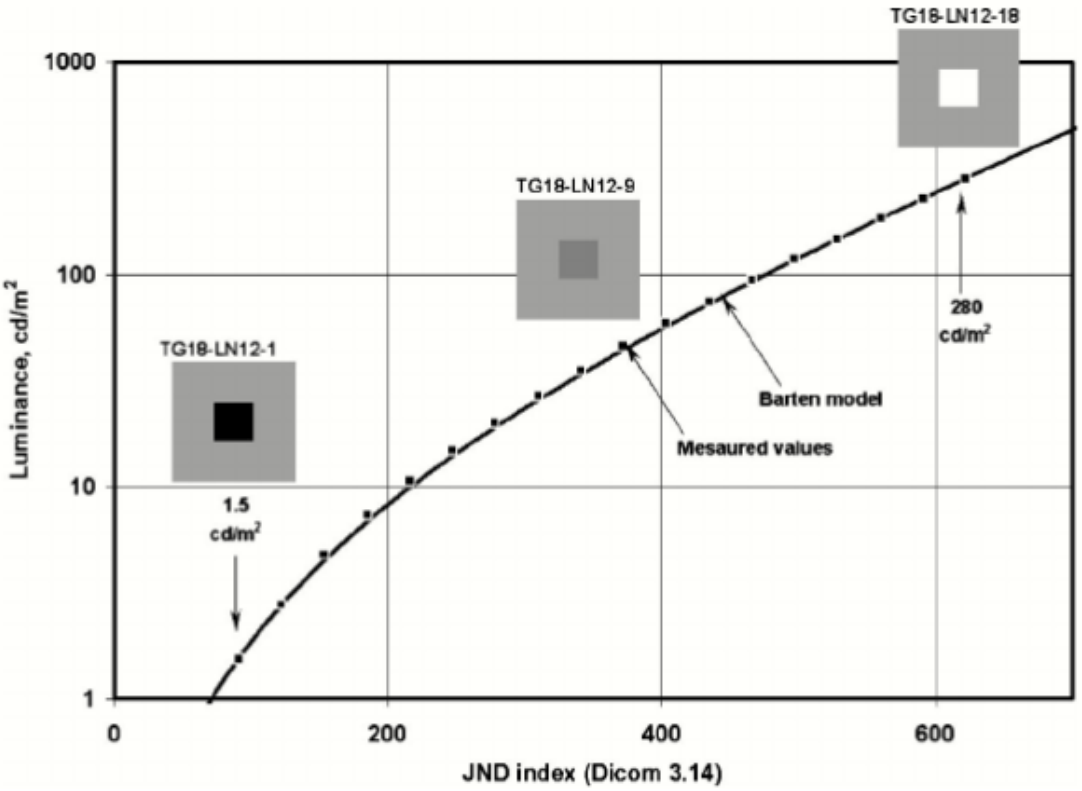

Figure 5. Grayscale Standard Display Function presented as logarithm-of-luminance versus JND-Index [6]. The X-axis shows the JND index, which shows grey levels. Y-axis shows luminance in cd/m^2. As the human eye hardly notices dark images, first grayscale levels should increase in wider increments until the grey level becomes bright enough. As a result, curvature in the curve disappears as increments of luminance in grey levels become even.

The data used for contrast response was taken from Barten's model, derived from the human visual system. As it is stated in DICOM GSDF 3.14 standard, "Grayscale Standard Display Function refers to Contrast Sensitivity for the Standard Target consisting of a 2-deg x 2-deg square filled with a horizontal or vertical grating with sinusoidal modulation of 4 cycles per degree." That square object is displayed in uniform (usually 20% of maximum luminance) background [15].

After obtaining all points measured in cd/m^2, all of the luminance values should be converted to just noticeable difference (JND) index j. The luminance gradient describes the desired JND (Just Noticeable Difference) at the Luminance L. It is a minimally distinguishable difference, required to notice by an average observer. This conversion thus provides an easy way to compare the calculated values with the intended luminance response. [3]. The coefficients a to m and A to I for equations are given in table 1.

These are the equations to convert between luminance, L, and JND index, j (equation 1 is used when JND indexes are known, equation 2 is used when luminance values in cd/m^2 for each greyscale value are known) [3]:

$$log_{10}\, L\,(j) \;= \frac{a + c \cdot ln\,(j) + e \cdot ln^2\,(j) + g \cdot ln^3\,(j) + m \cdot ln^4\,(j)}{1 + b \cdot ln\,(j) + d \cdot \ln^2(j) + f \cdot \ln^3(j) + h \cdot \ln^4(j) + k \cdot \ln^5(j)} \quad (1)$$

$$j\,(L) \;=\; A + B \cdot log_{10}(L) \;+\; C \cdot log_2 10(L) \;+\; D \cdot log_3 10(L) \;+\; E \cdot log_4 10(L) \;+\; F \cdot log_5 10(L) \;+\; G \cdot log_6 10(L) \;+\; H \cdot log_7 10(L) \;+\; I \cdot log_8 10(L). \quad (2)$$

Table 4. DICOM GSDF Conversion Coefficients (provided by AAPM TG 270 report) [3].

JND to Luminance		**Luminance to JND**	
a	−1.3011877	A	71.498068
b	−2.5840191E – 02	B	94.593053
c	8.0242636E − 02	C	41.912053
d	−1.0320229E − 01	D	9.8247004
e	1.3646699E − 01	E	0.28175407
f	2.8745620E − 02	F	−1.1878455
g	−2.5468404E − 02	G	−0.1801434
h	−3.1978977E − 03	H	0.14710899
k	1.2992634E − 04	I	−0.017046845
m	1.3635334E − 03		

After that, mean JND for each GSDF level is measured using the formula [3]:

$$\text{mean}\Delta JND/GL \;= \frac{j\,(L'_{max}) - j\,(L'_{min})}{GL_{max} - GL_{min}} \quad (3)$$

Where the denominator is equivalent to the difference of the minimal and maximum gray levels. Because the aim of the GSDF curve is to have an equivalent number of JND indices across each gray level, the average ΔJND/GL is the number of JND index that would be present for each gray level on a GSDF capable display. Next in the study, the ΔJND/GL across each observed gray level compared to the mean ΔJND/GL to determine the error at each measured phase.

There is an additional parameter to add, which takes into account the contrast threshold model. Equation 4 describes the slope of the luminance response function in the plot of gray level vs.

luminance[3]. As observers' eye cannot detect as sharply differences of contrast in high levels of luminance, the curve of GSDF is not linear. Relative contrast is indicated as dL/L.

$$dL/_{L}\, per\ JND = \frac{2(L'_i - L'_{i-1})}{(L'_i + L'_{i-1})(mean\frac{\Delta JND}{GL})(GL_i - GL_{i-1})}. (4)$$

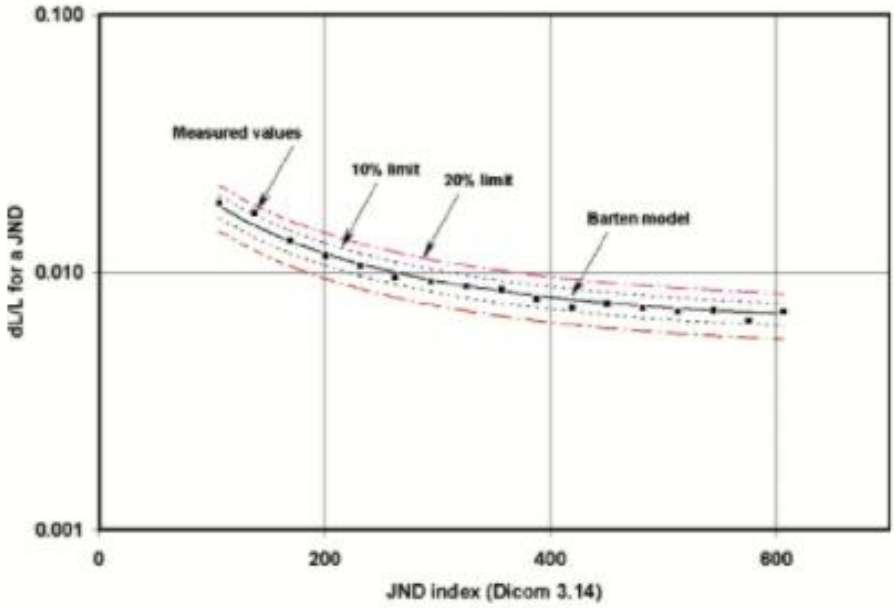

Figure 6. dL/L per JND plot, indicating 18 measured grey levels and comparing them with Barten's model. Measured values should not exceed 10 % limit if the monitor is diagnostic type and 20% if the monitor is classified as a clinical review, modality, or electronic health record type [6].

1.4.4. Display Luminance Uniformity

Display luminance uniformity shows how uniform luminance distribution is over the entire display panel. The main idea is that image should be displayed consistently despite the location of where the image is. Therefore, there are TG18-ULN10 and TG18-ULN80 test patterns, where luminance is measured, and after that, the maximum error is measured between values by using the formula [14]:

$$200 \times \frac{L_{\%maximum\ dev.} - L_{\%minimum\ dev.}}{L_{\%maximum\ dev.} + L_{\%minimum\ dev.}} (5)$$

When an image is displayed, it is expected that grey tone would still produce a constant luminance, regardless of the picture's position on the screen. When this expectation is not met, chances are that diagnostic interpretations with this monitor will become challenging, as with non-uniformities, it might be easy to confuse artifacts with examination image details. Non-uniformities appear in many different forms and a broad spectrum of sizes, shapes, and positions on the screen. Often is better to evaluate in qualitative assessment to get a general situation, as it might show non-uniformities in display in such places where doing qualitative evaluations might never be possible to spot. There are various types of reasons, i.e., OLED displays over a long time of use eventually develop burn-in, which result in darker places, LCDs might develop stuck or dead pixels due to damaged liquid crystals, backlight bleeding effect in bright areas might cause non-uniformities in bright screen, as polariser is physically damaged and light travels a different distance, causing white image to appear brighter in the physically impacted area. Because of this, TG18-UL10 and TG18-UL80 test patterns were premade, and they are just as, if not more, essential to evaluate for mentioned reasons [5].

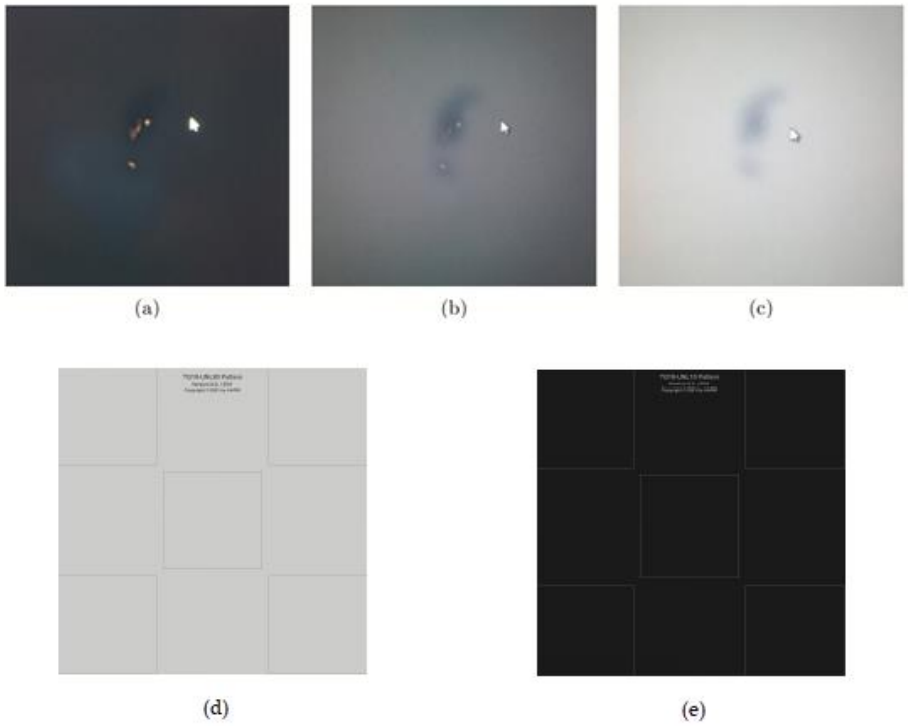

Figure 7. Display uniformity patterns. A, B, and C pictures show non-uniformities with dead (or stuck) pixels in TG18-UN10 (a), TG18-UN50 (b), TG18-UN80 (c) test patterns. For quantitative display uniformity evaluation, TG18-UNL80 (d) and TG18-UNL10 (e) patterns are used to evaluate each quadrant's luminance and use formula 5 to calculate the overall non-uniformity of the display. Dark and bright uniformity is evaluated as different values.

1.4.5. Display Noise

Noise is a critical requirement determining visibility. It may be various artifacts that are affected by noise. Observers' eye fatigue sometimes might also be a significant factor [18]. Apart from the image noise, a noise from the display itself may cause randomness in the data transmission of the displayed image. Noise may occur as a result of both temporal and fixed spatial variation (pixel by pixel), although there have been reports of stationary spatial variation being the most significant source of noise [19].

The spatial noise characteristics of the display can vary among various types of flat screens, whether because of fluid crystal inappropriation, cell thickness, the distance between them, voltage regulation issues, etc. [20]. In general, displays sold for medical imaging include displays chosen with high uniformity and can use pixel by pixel adjustments to make them more consistent.

1.4.6. Veiling Glare

Veiling glare causes display to lose contrast when surrounding light is too high due to diffused light-spreading or scattering effect within its various parts. It is different from different display technologies (especially many articles mention CRT and LCD differences) [9]. LCDs have a protective layer for TFTs (thin-film transistors), which might cause this effect. On the other hand, CRT displays are much more affected, as light rays produced by cathode ray tube are directed towards the viewer. This light ray guidance towards the screen is accomplished by inserting a reflective back into the phosphor layer's vacuum side. Light emitted by the phosphor particles must pass through a

thick faceplate, where several scatterings occur before reaching the observer's eye (as shown in figures 8 and 9). The phosphor layer's minimal light absorption facilitates several photon scatterings within the faceplate. In CRT emissive structures, lateral light diffusion triggers the spread of luminance functions with light rays that stretch across the entire viewing region (see figure 8) [21]. This diffuse aspect over broad areas reduces the device's full contrast capabilities [9].

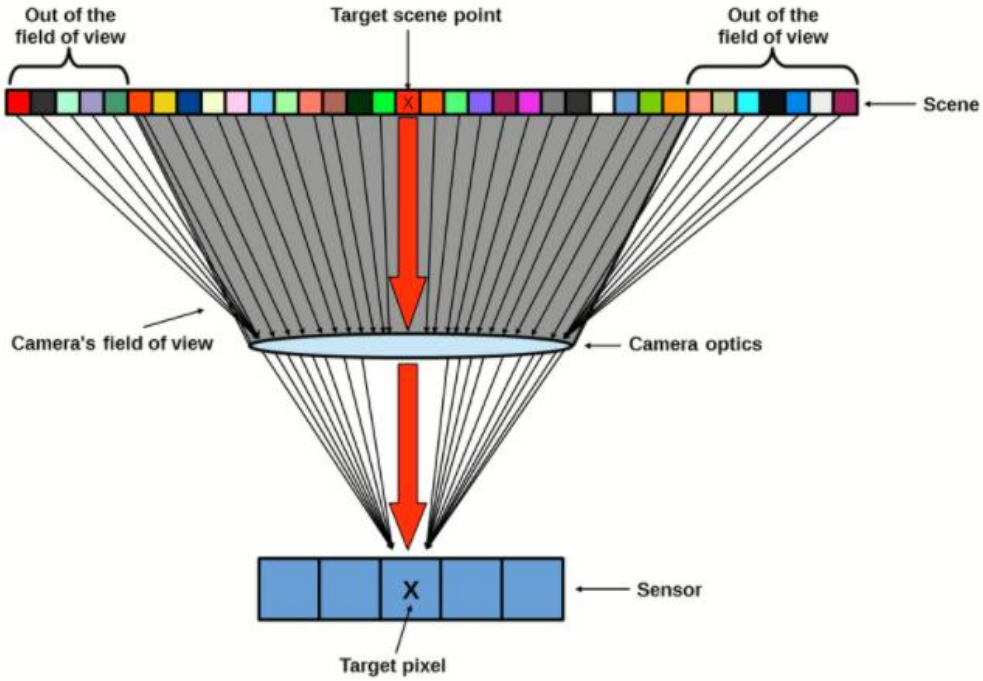

Figure 8. Veiling glare effect occurring scheme cross-section. Arrows show how veiling glare affects the response of pixels from each scene point to the target. [22]

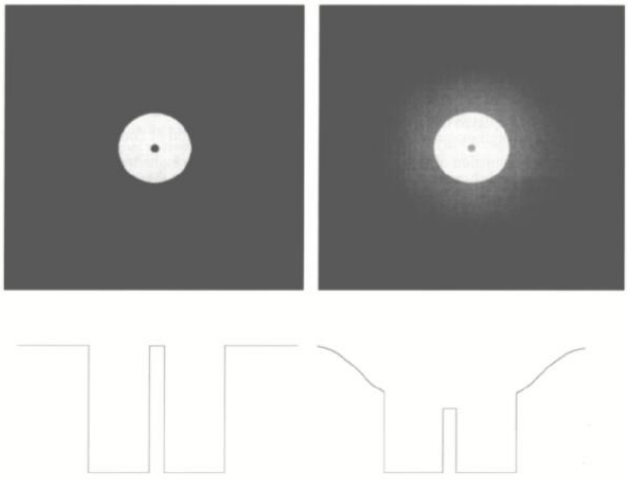

Figure 9.Veiling glare effect, seen on TG18-GV pattern. The curves shown below show angles, how light is scattered (the left image shows no light scatter, the right image shows veiling glare effect) [21].

Veiling glare typically has only visual evaluation and can be performed using TG18-GV and TG18-GVN pattens.

In sequential observation of the TG18-GVN & TG18-GV images (TG18-GV pattern shown in figure 10), the user should distinguish the appearance of the low-contrast elements. Since seeing the bright field will cause the human eye to adjust, it is essential that the bright field is completely blocked from

sight and no reflected light from the bright field is visible. This can be done using a mask or cone that protects the eye from patterns surrounded luminance. There should be no distinguishable difference in contrast of the image content between the two pictures (TG18-GV and TG18-GVN). For diagnostic displays, at least three items should be easily visible in either pattern. At least one 5th goal corresponds to the corresponding objects for secondary class (modality, clinical review, and electronic health record) displays.

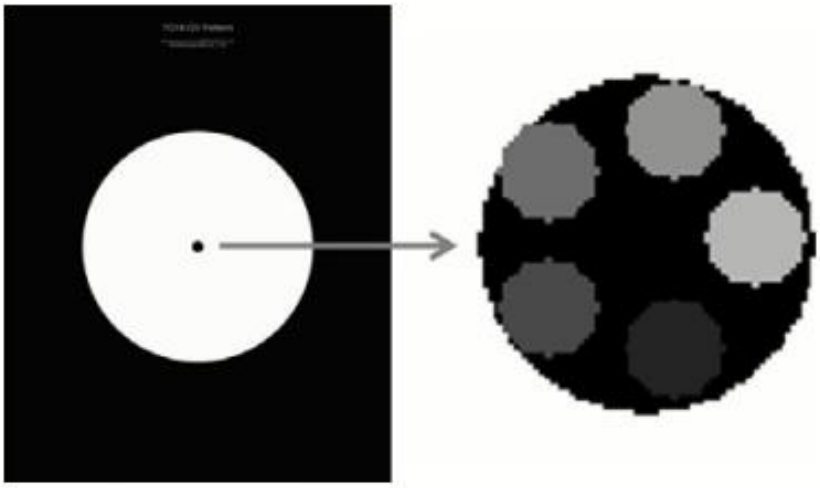

Figure 10. TG18-GV pattern, with the five contrast elements in the center (elements enlarged in the bottom right). The same elements are in the TG18-GVN pattern, where the background is entirely black. Through these patterns veiling glare is checked visually by counting how many fewer contrast elements are seen in the TG18-GV pattern [23].

Quantitative evaluation of veiling glare also exists and is performed using a collimated luminance meter with TG18-GQ, TG18-GQB, and TG18-GQN patterns. White region of the test pattern should not exceed 20 centimeters. White region, when measuring center element luminance (L), should be blocked. Luminance elements in the center, white region luminance (L_B), and the background luminance (L_N) are then measured using a high-collimation photometer. After obtaining the luminance results, a formula to calculate glare ratio (GR) is used [6].

$$GR = \frac{(L_B - L_N)}{(L - L_N)} (6).$$

Veiling glare effects for diagnostic displays should not exceed 20% for contrast ratio and 25% for luminance ratio. If the maximum to minimum luminance ratio is recommended to be at least around 250:1, this means glare ratio should be around 1000:1. However, TN and VA type LCD panels, especially older ones, cannot pass such criterion. Therefore, glare ratio, exceeding at least 350 for diagnostic and 150 for modality, clinical review, and electronic health record display devices, is considered acceptable by the AAPM. [3]

1.5. Calibration

Since the actual luminance (and thus perceived contrast) characteristics of various LCDs may fluctuate significantly, calibration of these monitors' luminance response is needed to ensure that users' perception of the image is consistent. As mentioned before, displays used in medicine should meet the requirements mentioned in previous work sections. Calibration can fix the luminance

response of the screen. The main goal for this is to ensure the best possible contrast and luminance appearance for radiographic images.

To properly calibrate the display, several requirements must be met before: ambient luminance should not exceed recommended 20 lx, as well as display should be positioned so that its reflectance and surroundings would not affect its performance due to in previous chapters mentioned reasons.

1.5.1. Selection of Luminance Range

To calibrate a display, minimum and maximum luminance should be defined first, as LCDs have a backlight that tends to decrease its luminance output over a long period of use. It usually is measured while determining the contrast ratio. Color displays typically have contrast ratios of 250 to 400, whereas grayscale displays can have contrast ratios of 600 or more when brand new [9]. To properly select the desired maximum luminance, it is essential to note that during calibration, peak luminance might variate according to mathematical calculations applicable to Barten's model [9]. L_{min} usually is selected to be higher than the minimum value as well, due to the same reason.

After L_{min} & L_{max} value selection, $JND_{min,}$ as well as JND_{max} values are calculated using conversion equation, previously shown with coefficients (see table 1 in section 1.3.3):

$$j\,(L) = A + B \cdot log_{10}(L) + C \cdot log_2 10(L) + D \cdot log_3 10(L) + E \cdot log_4 10(L) + F \cdot log_5 10(L) + G \cdot log_6 10(L) + H \cdot log_7 10(L) + I \cdot log_8 10(L).$$

The difference between JND_{min} and JND_{max} shows how each greyscale point should distribute. For example, a display with minimum luminance of 0.5 cd/m^2 and L_{max} with 200 cd/m^2 translates to 46.6 JND_{min} and 572.2 JND_{max}. This shows that the total JND difference is 525.6. As most screens have 8-bit graphics, a full palette of greyscale consists of 256 values. The main technical calibration objective is to set luminance levels for every digital driving level (every possible brightness value of, in this example, grey levels, shortened as DDL) input value in such way that a difference would be changing throughout the whole DDL range consistently in JND index (DDL increment changes in GSDF when the display is calibrated and uncalibrated to DICOM GSDF are shown in GSDF curve example in figure 11). To do so, the total count of JNDs have to be divided by the total number of DDLs, then subtracting one. This specifies the calibrated display's average amount of JNDs per single DDL (JND_{ave}). So, JND_{ave} equals to 525.6/255 = 2.06. Each DDL is then calculated using the formula

$$JND_{DDL} = JND_{min} + DDL * JND_{ave} \quad (7).$$

Then for each DDL to obtain luminance values, an equation is used:

$$log_{10} L\,(j) = \frac{a + c \cdot ln\,(j) + e \cdot ln^2\,(j) + g \cdot ln^3\,(j) + m \cdot ln^4\,(j)}{1 + b \cdot ln\,(j) + d \cdot \ln^2(j) + f \cdot \ln^3(j) + h \cdot \ln^4(j) + k \cdot \ln^5(j)}$$

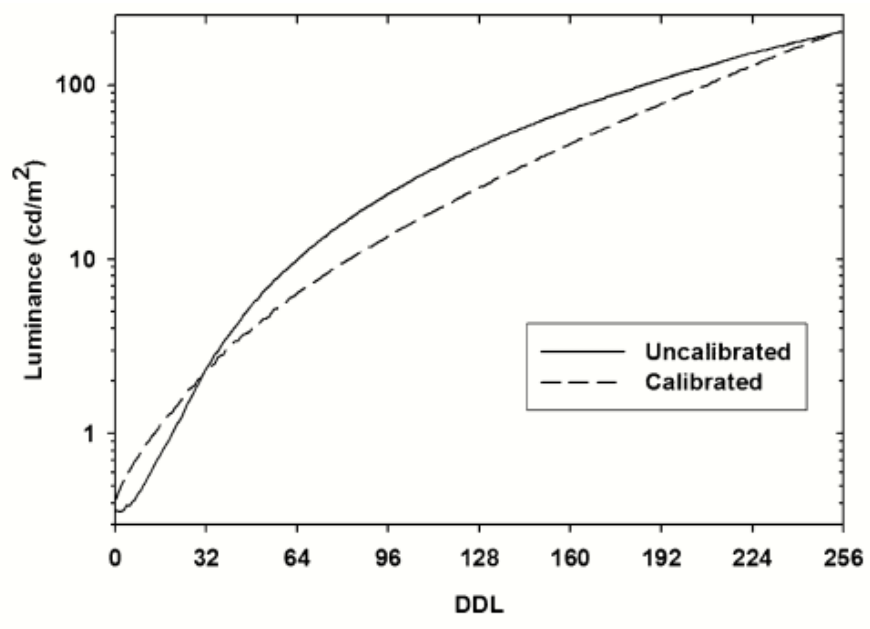

Figure 11. Luminance intensity examples for calibrated and uncalibrated displays for every digital driving level (DDL) of grayscale palette [9].

After obtaining JND and luminance values, a look-up table (LUT) has to be created. It is a file where new DDL values are created for each DDL matching DICOM GSDF. Using a specific program (i.e., pacsDisplay), this file can give instructions to change input values into specific output values for each DDL. The program then can transform the existing display's profile, which changes its properties that are more likely to match DICOM display standard. This conversion is simple, however, not perfect. 8-bit to 8-bit calibration satisfies overall perceptual contrast, but adjacent input pixel values (sub-pixel red, green, and blue values can be adjusted and manipulated for coloured monitors to gain better contrast and more accurate DDL values) are being unchanged; therefore, adjacent input levels change drastically, leaving monitor with poor contrast and missing grey shades (as transitions might be too drastic in some grey levels). Since 8-bit-to-8-bit calibration usually results in local contrast loss along with unnecessarily excessive local contrast, it cannot be an adequate calibration.

It can be stated that grayscale calibration based on the initial luminance measurements at every possible level of LUT value gives the best calibration quality as it is, related to how many points are adjusted for a better match to DICOM GSDF. If every point has been adjusted, the quality of calibration will be the best possible. So, for more precise calibration, each sub-pixel value has to be modified. Commonly, sub-pixel values are adjusted so that all three of them increase at least by 1 level and, therefore, depend on each other. Independent sub-pixel adjustment allows to obtain more than 256 gray values and therefore is called bit stealing, otherwise known as spatial dithering [24, 25]. Monochrome monitors can obtain a palette of 766 and coloured ones up to 1786 gray colour tones[26,27]. This allows more accurately match DICOM GSDF model. Compared to the 8-bit palette, there is more choice of possible grey levels to match each grey value for DICOM GSDF after bit stealing. This leads to a possibility to calibrate a monitor that more closely resembles an ideal curve.

With bit stealing, several vendors have used temporal modulation of subpixels to further improve the luminance variation between LUT palette levels [24]. Temporal modulation improves the time-

averaged luminance output by using fast sequential adjustments in the luminance output of individual subpixels. It works at an undetectable pace for humans to notice by the naked eye.

It is important to note that display calibration to DICOM GSDF only ensures that all displays would perform at their best. The overall appearance of medical images can vary in many variables, like image obtaining parameters (mAs, peak kilovoltage (kVp), etc.) of scanning machine (i.e., CT scanner, X-ray machine, or other modality). Image processing algorithms, which improve over time, tools that can change display contrast also might help to see medical images better along with greyscale transformation mechanisms integrated into PACS program user interface [9].

1.5.2. Storing the calibration settings

As there are many different ways to store the calibration data, one of the main is by integrating the calibration look-up table as International Color Consortium (.icc) file into operating system's GPU drivers. Then computer's graphics processing unit (GPU), along with its drivers, can manipulate DDL values by reading the file and generate different output. Also, calibration data can be stored in monitors provided software which includes drivers, directly linked to computers GPU.

2. Methods

Before calibration, it is essential to measure all monitors to collect information about their qualitative and quantitative parameters. Selected display categories for measurements were diagnostic (primary) and clinical (secondary). To properly overview the situation, it is essential to prepare a measurement procedure document for a fast and effective measurement process.

While measuring, the conditions of the environment and the monitor settings must be left the same as being used for medical examinations. After initial measurements, it is vital to determine which monitors have to be calibrated, which parameters are not acceptable, and what has to be changed.

2.1. Initial Measurements

As measurements were divided into qualitative and quantitative, different equipment was used for each category. There were three main stages – initial measurements, calibration, and measurements after the calibration. After measurements, Student Two Sample t-test was performed to evaluate the difference in monitor performance before and after the calibration. All monitors before any measurements were set to DICOM mode if it was possible.

For initial quantitative display measurements, a Piranha RTI L-100 photometer was used. TG18 test patterns were used for minimum and maximum luminance measurements, luminance response, veiling glare, noise, and display uniformity. Also, ambient light intensity was measured using the same Piranha RTI L-100 device, using an ambient light detector (figure 12).

Before performing any tests, the display itself has to be enabled for 30 minutes. It is essential due to possible inconsistencies occurrence to disappear, especially in older monitors. Also, it is essential to inspect whether the display is clean, as fingerprints and dust residues might affect emitted light optical properties – it can cause distortions or emitted light refraction.

Testing patterns were displayed at native monitor resolution for the best result. All parameters were registered using Ocean 2014 software, which recorded measured values from the Piranha RTI L-100 device.

Figure 12. Equipment used for quantitative monitor measurements. Number 1 in the picture presents a cap for ambient light measurements, number 2 indicates a cap for luminescence measurements, number 3 is a device itself that processes the data and sends it to the computer, and number 4 is an extra detail for luminance measuring cap to make measurements more accurate by making them more directional.

2.2. Quantitative measurements

At first, ambient light was evaluated. This test is simple – after connecting the photometer to the computer and opening Ocean APK 2014 program, the program automatically detects the attached sensor and selects to measure ambient light. Ambient light is measured in lux units.

Then, display reflection (while the display is turned off) measured using a photometer from a 50 centimeter distance. This measurement is essential because intense ambient light might negatively affect the visibility of low-contrast images. Dark images might be difficult to see, as reflections get in the way, etc.

The luminance ratio was evaluated using the TG270-sQC pattern (as shown in figure 4). Black square measurement was added to ambient light luminance and marked as minimum luminance measurement. White square measurement was done precisely the same, just with the white patch in TG270-sQC pattern and considered to be maximum ambient luminance that monitor can produce. Dividing minimum luminance from maximum luminance, a result of the luminance ratio was calculated.

For luminance response evaluation, TG18-LN (LN1 to LN18) patterns were used. When measuring luminance response, a photometer is placed on the center of the screen and measuring the luminance of each pattern. A monitor's luminance response is related to its luminance for each display driving level (DDL) of greyscale standard display function (GSDF). Measured luminance is typically combined with ambient light that is diffusely reflected from the monitors' surface. During this measurement, the luminance ratio is evaluated as well. It is the ratio of maximum luminance to minimum luminance, including ambient light. When ambient light levels are reduced, the contrast ratio can be computed by excluding the remaining reflected luminance from the display surface from the total display producing luminance.

The luminance uniformity evaluation includes measurements of high and low luminance values, covering the entire screen. Five quadrants are measured to obtain medium deviation of display backlight uniformity distribution – upper left, upper right, lower left, lower right, and center. For this assessment, TG18 test patterns UNL10 (minimum luminance) and UNL80 (maximum luminance) were used [6]. The mean percentage of LCD monitor's backlight deviation is calculated using the formula:

$$200 \times \frac{L_{\%maximum\ dev.} - L_{\%minimum\ dev.}}{L_{\%maximum\ dev.} + L_{\%minimum\ dev.}}$$

This test is marked as passed when the luminance deviation did not exceed 30%.

2.3. Qualitative tests

Luminance response was also evaluated visually using the TG270-sQC test pattern, which was recently succeeded the TG18-CT pattern. TG270-sQC pattern contains eighteen patches, covering an 8-bit greyscale value range. The square patches are supposed to increase in their luminance at equal increments of 15 greyscale driving levels. Each patch is marked with its digital value. Also, inside these patches, there are low-contrast bar patterns in the top-left and bottom-right corners. These low-contrast bars are offset ±5 digital values (-5 top-left; +5 bottom-right corner of the patch). The test is considered passed when low contrast bars from 15 to 240 digital levels are visible in the pattern.

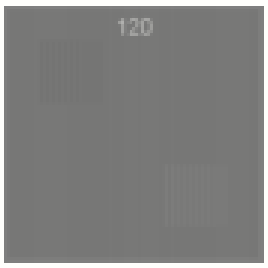

Figure 13. TG270-sQC patch containing low-contrast bars. Bars have slightly different digital display level compared to the background (the difference is -5 in the upper - left corner and +5 in the bottom - right corner)[27]

Another test from this pattern checks if ambient light is appropriately set for luminance response evaluation. If it is impossible to see low-contrast bars in the bottom-right corner in patch 0, then it means that ambient light needs to be dimmed until low-contrast bars are possible to see.

The final test of provided 18 patches checks whether the upper left bar is visible in the brightest patch. If it is possible to see this detail, the monitor likely meets qualitative requirements for luminance response.

Luminance uniformity was visually evaluated using TG18-UN10 (dark background) and TG18-UN80 (light background) patterns. These patterns should be evaluated from a 30-centimeter distance. No dead or stuck pixels and no luminance variations should be visible, which could be confused with something else to pass the test. If it is an OLED type of screen, no burn-ins or darker areas should be visible, as it is the most common fault of OLED screens.

Display noise in this work was evaluated only visually, using the TG18-AFC pattern. It is based on determining just noticeable luminance differences that vary in size. The pattern is divided into four main quadrants. Each of those quadrants is divided into squares in which there are square-shaped patches with different positions in every square. Every patch has a different location in the square. The size of patches is the same per quadrant by being the largest in the bottom-right quadrant and the smallest in the top-left quadrant part of the test pattern. Also, the test pattern in every corner and center has 16 size-varying square-shaped patches (five in total), which vary in size, being the largest at the bottom-right and smallest at the top-left. For diagnostic monitor type, test considered to be passed, when all 16 patches are visible in the corners and center, as well as all square-shaped patches in at least three quadrants are possible to distinguish. For the secondary type of displays, at least nine patches should be visible in the corners and center, as well as all square-shaped patches in at least two bottom quadrants are possible to distinguish.

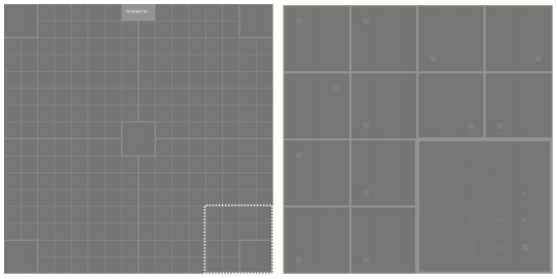

Figure 14. TG18-AFC pattern. On the right is visible small square patches, which should be visible [28].

Veiling glare was evaluated visually, using two test patterns - TG18-GV and TG18-GVN. TG18-GV pattern contains dark background, a white circle, and a black circle inside a white circle, in which also there are five low-contrast objects. The same five objects are in the center of the TG18-GVN pattern. TG18-GVN pattern consists of black background and five circular objects in the center of the test pattern. To pass the test, it is necessary under normal conditions to see at least three out of five objects in the middle of both patterns for primary displays. For secondary displays, it is essential to see at least one low-contrast object in both testing patterns to pass.

2.4. Calibration Hardware

For monitor calibration hardware, the X-Rite i1Display Pro (see fig. 13) photometer was used. It has a sensor for ambient light measurements (activated by placing a special matte lid on top) together with a photometer, allowing to measure monitors luminance, as well as a colorimeter, which measures the wavelength of emitted light from the monitor. It is capable to accurately measure luminance values in the range from 0.1 cd/m^2 up to 1000 cd/m^2 [29]. The ambient light measurement range can vary from 0.5 up to 5000 lx. Dependingly on measuring angles, its accuracy for luminance might vary; however, the device itself at maximum can make a 4 percent error.

Figure 15. X-Rite i1Display Pro calibration device. It has an integrated photometer, colorimeter and is capable of measuring ambient light intensity.

2.5. Calibration Software

Monitors were calibrated using PerfectLum™ 4.0 software (Qubyx Ltd, Horton, UK) (for application's user interface, see fig. 14). This software supports any display for its luminance response calibration to DICOM and Gamma 1.0 to 2.8 standards. It also has integrated quality assurance patterns, like AAPM TG18, DIN 6868-157, DIN 6868-57, NY PDM, NYC PDM, and ACR.

2.6. Preparations before calibration

Before the calibration, initial preparations were done to make sure that monitors will be calibrated accurately. At first, monitors that did not have a DICOM profile were set to standard mode (inside the built-in monitor menu). Monitors that had integrated DICOM setting were set to this setting. All monitors were set to maximum brightness if it was possible. After that, ambient light intensity and reflections from other monitors were minimized. Before calibrating, monitor surfaces were cleaned.

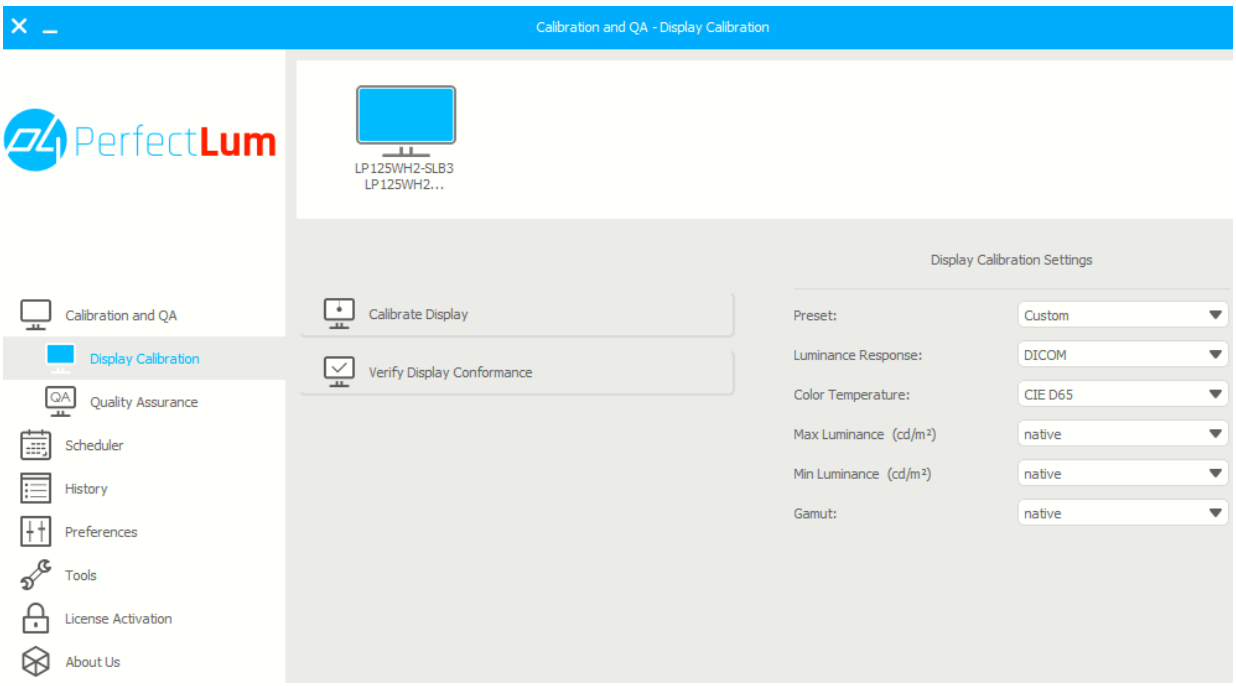

Figure 16. PerfectLum 4.0™ monitor calibration software.

2.7. Calibration procedure

Before starting the calibration, an X-Rite i1DisplayPro photometer was connected to the laptop. Once it is done, a PerfectLum calibration software was opened, and in the Calibration and QA tab selected Display Calibration following by the step Calibrate Display. It is essential to select the display that has to be calibrated.

After pressing Calibrate Display command, the display will turn black with the contours on the screen to place the photometer (as shown in figure 15). It is crucial that the photometer is placed accurately and is stable in its position, as movements might fail the calibration process.

After calibration is finished, the program notifies the user whether it was successful (see fig. 16). Monitors then were re-evaluated using a Piranha RTI photometer. It was worth to mention, that re-evaluation was performed using Piranha RTI photometer using the same methods as described in sections 2.1.1. and 2.1.2. because measuring conditions and measurements would not be affected by equipment, environment, or any other factor.

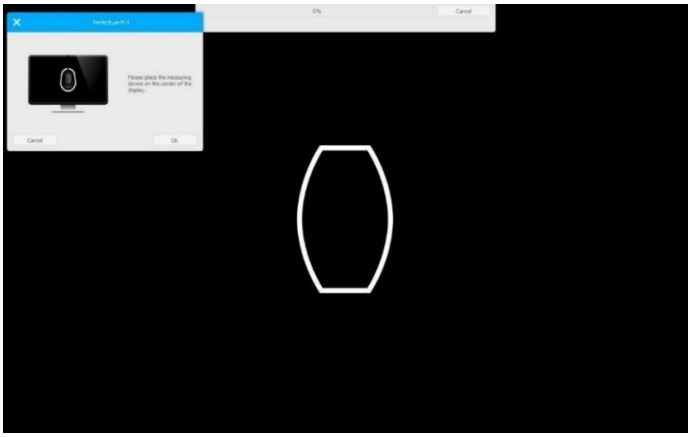

Figure 17. Calibration window, showing contours where the photometer has to be placed.

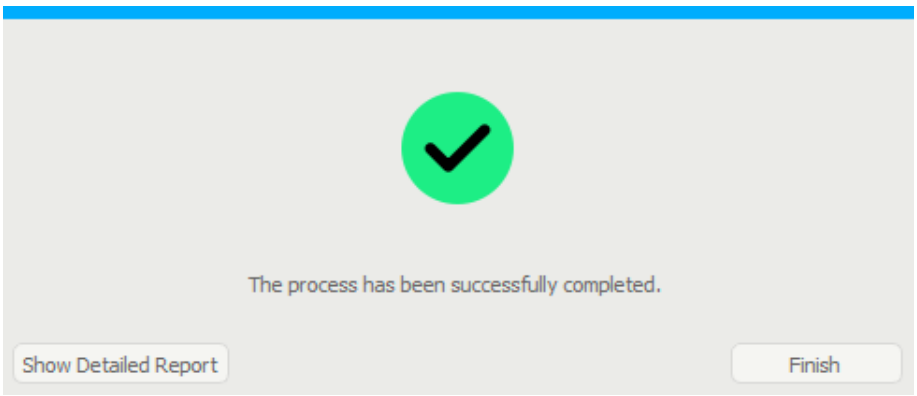

Figure 18. The notification from the calibration software about the successful monitor calibration.

3. Results and discussion

In total, 28 diagnostic and 11 clinical review displays were evaluated (n = 39). They were located in 3 different departments. 5 (NDS and Wide) monitors were greyscale type, all of which were diagnostic. Three displays (Wide) supported 14-bit input, four displays (JVC) were capable of supporting 12-bit input, 21 displays supported 10-bit input. All clinical review type monitors were capable of 8-bit LUT calibration and supported color (see table 5).

Before the calibration, initial measurements are presented. The criteria included qualitative and quantitative measurements. For the quantitative criteria, illuminance, monitor reflection, minimum luminance, DICOM GSDF absolute deviation percentage, maximum monitor brightness, luminance ratio, and monitor backlight uniformity deviation in the bright and dark background were measured. For qualitative measurements, noise, veiling glare, and defects for display nonuniformities (dead pixels, backlight bleeding, darker and brighter spots) under bright and dark background patterns (if any visible, the test is failed).

Table 5. Basic technical specifications of measured monitors.

No.	Type	Vendor	Model	Year of manufactory	Native resolution	Screen size	Bit support
1	Clinical	Dell	U2415	2018	1920 × 1200	24.1"	8
2	Clinical	Dell	P2417H	2019	1900 × 1200	24"	8
3	Clinical	HP	EliteDisplayE221C	2013	1920 × 1080	21"	8
4	Clinical	Dell	U2415	2018	1920 × 1200	24.1"	8
5	Diagnostic	NEC	LCD1990SXi	2008	1280 × 1024	19"	10
6	Diagnostic	NEC	LCD1990SXi	2008	1280 × 1024	19"	10
7	Diagnostic	NEC	LCD1990SXi	2008	1280 × 1024	19"	10
8	Diagnostic	NEC	LCD1990SXi	2008	1280 × 1024	19"	10
9	Diagnostic	NEC	PA272W	2014	2560 × 1440	27"	10
10	Diagnostic	Eizo	Radiforce MX193	2016	1280 × 1024	19"	10
11	Diagnostic	Eizo	Radiforce MX193	2016	1280 × 1024	19"	10
12	Diagnostic	Barco	MDRC-2324	2020	1920 × 1200	24"	10
13	Diagnostic	Barco	MDCC-6530	2020	3280 × 2048	30.4"	10
14	Diagnostic	JVC	CL-R211	2020	1200 × 1600	21.3"	12
15	Diagnostic	JVC	CL-R211	2020	1200 × 1600	21.3"	12
16	Clinical	HP	LA2306x	2012	1920 × 1080	23"	8
17	Clinical	Dell	P2412H	2018	1920 × 1080	24"	8
18	Clinical	Dell	P2412H	2018	1920 × 1080	24"	8
19	Diagnostic	Barco	MDRC-2324	2020	1920 × 1200	21"	10
20	Diagnostic	Barco	MDCC-6530	2020	3280 × 2048	30.4"	10
21	Diagnostic	Eizo	EV2456	2021	1920 × 1200	24.1"	10
22	Diagnostic	Eizo	EV2456	2021	1920 × 1200	24.1"	10
23	Clinical	Dell	P2417H	2019	1920 × 1080	24"	8
24	Clinical	Dell	U2412M	2018	1920 × 1200	24"	8
25	Clinical	Dell	P2719H	2019	1920 × 1080	27"	8
26	Clinical	Dell	U2412M	2016	1920 × 1080	24"	8
27	Diagnostic	Lenovo	T27Q	2021	2560 × 1440	27"	10
28	Diagnostic	Lenovo	T27Q	2021	2560 × 1440	27"	10
29	Diagnostic	JVC	CL-R211	2019	1200 ×1600	21.3"	12
30	Diagnostic	JVC	CL-R211	2019	1200 × 1600	21.3"	12
31	Diagnostic	Eizo	FlexScan S1923	2019	1280 × 1024	19"	10
32	Diagnostic	Eizo	FlexScan S1923	2019	1280 × 1024	19"	10
33	Diagnostic	NDS	Dome E2	2013	1200 × 1600	21.3"	10
34	Diagnostic	NDS	Dome E2	2013	1200 × 1600	21.3"	10
35	Diagnostic	Eizo	Radiforce GX540	2015	2048 × 2560	21.3"	10
36	Diagnostic	Eizo	Radiforce GX540	2015	2048 × 2560	21.3"	10
37	Diagnostic	Wide	2210E	2019	1536 × 2048	24.1"	14

Table 5 (continued). Basic technical specifications of measured monitors.

No.	Type	Vendor	Model	Year of manufactory	Native resolution	Screen size	Bit support
38	Diagnostic	Wide	2210E	2019	1536 × 2048	24.1"	14
39	Diagnostic	Wide	JMW1100KB2449F03	2019	1536 × 2048	24.1"	14

3.1. Initial quantitative measurements

3.1.1. Illuminance

Illuminance measurements were performed before each monitor measurement. Before calibration, ambient light varied in the range of 7 to 239 lux. The brightest room, where professional radiologists worked, evaluating mostly conventional radiology images, was measured during the bright day with blinds shut. The mean value of all illuminance measurements was 50.3 lux. In total, 15 monitors exceeded the maximum recommended illuminance value. For 12 displays, the illuminance values ranged between 0 and 20 lux; for 12 displays, the illuminance was between 21 and 50 lux, in the range of 55 to 70 lux fell 11 displays, and the remaining four displays were working in the range of 90-239 lx ambient light intensity.

Illuminance measurements varied due to the daylight intensity irregularities over the day, as well as some rooms had blinds, which were not as effective. According to AAPM guidelines, only nine displays in total were operating within recommended threshold values, five of them were located in nuclear medicine department (in the range of 25-50 lx). Nevertheless, it should be mentioned that in most instances, the ambient lighting was changing over the day and could always be adjusted. The illuminance value was measured when radiologists were interpreting the examinations. As a result, they were unlikely to develop disruptive eye strain.

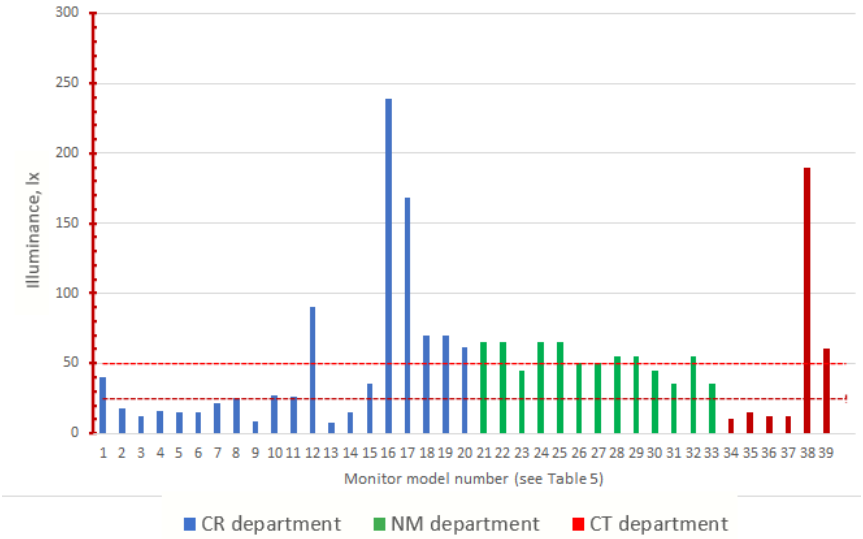

Figure 19. Measured illuminance values. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Red lines show the AAPM ambient illuminance recommended threshold values (25-50 lx). Numbers correspond to the monitor models in table 5.

3.1.2. Ambient luminance

Ambient luminance results were obtained by measuring display reflection 50 cm away from the monitor, which was turned off. In renewed AAPM technical report guidelines, the ambient luminance parameter should be less than 25% of the minimum monitor luminance value. Eight monitors (5 in

conventional radiology, 4 in nuclear medicine, 1 in computed tomography department) in total passed this criterion, 5 of which were diagnostic type (3 of them were located in conventional radiology department, 1 in nuclear medicine department and 1 in CT department, all clinical review type displays were located in nuclear medicine department). As the rooms were relatively small, it is very challenging to position monitors to pass this criterion, especially with those monitors that have a reflective surface.

Also, Pearson's product-moment correlation statistical data analysis showed that environmental illuminance and ambient luminance had moderately strong positive correlation (correlation coefficient = 0.5813173, p = 0.0001041), meaning that ambient luminance parameter moderately strongly depends on illuminance (the more illuminance increases, the more ambient luminance should increase as well).

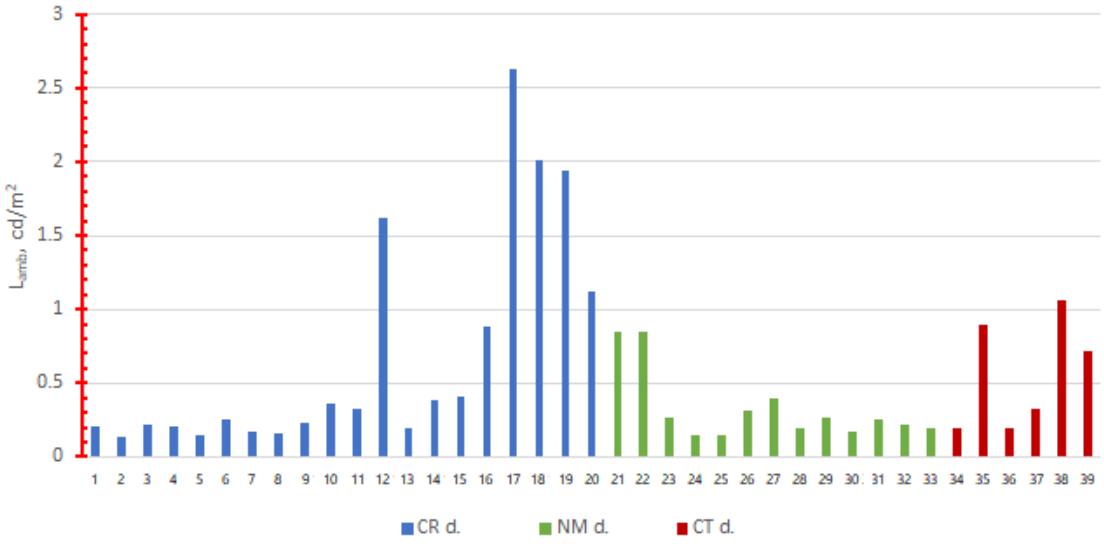

Figure 20. Measured ambient luminance values. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate monitor model number from table 5.

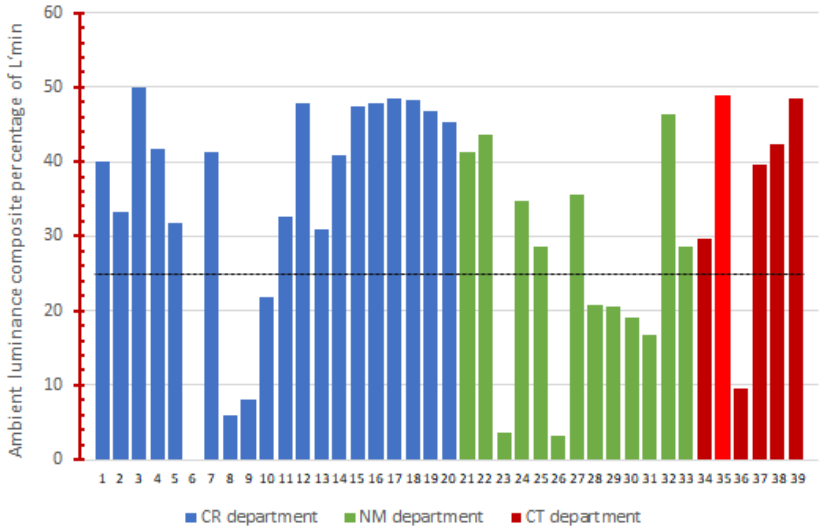

Figure 21. Measured ambient luminance composite percentage of L'min. Dashed line indicates the threshold value of maximum recommended ambient luminance composite percentage of L'min (25%). Numbers indicate monitor model number from table 5.

3.1.3. Minimum luminance (L'$_{min}$)

There are no recommended threshold values for this criterion in AAPM guidelines; however, it can be evaluated by adding ambient light intensity values to minimum luminance, only produced by the monitor. The final value L'min for primary displays should be 1 cd/m^2 or more, while for secondary displays, the smallest recommended value is 0.8 cd/m^2. Figure 23 shows measured values of L'min for diagnostic displays, and figure 22 shows L'min measurements of clinical review displays. After measuring diagnostic display L'min, 11 displays passed this criterion (6 of them were in conventional radiology department, 2 in nuclear medicine department and 3 in CT department) but only because of more intense ambient light in the rooms. All clinical review type displays failed this criterion. The mean value of minimum luminance was 1.212 cd/m^2 for diagnostic monitors, the minimum measured value was 0.17 cd/m^{2,} and the maximum reached 5.1 cd/m^2. For clinical review monitor type, the mean value of minimum luminance was 0.288 cd/m^2, minimum measured value was 0.21 cd/m^2, and maximum reached 0.44 cd/m^2.

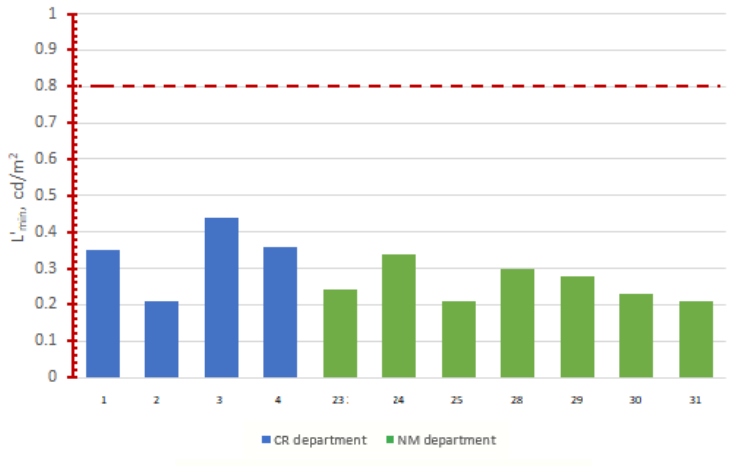

Figure 22. Measured L'$_{min}$ values for clinical review display type. CR – conventional radiology, NM – nuclear medicine department. The red dash line shows the minimum passable threshold value for L'$_{min}$ criteria. Numbers correspond to the monitor models in table 5.

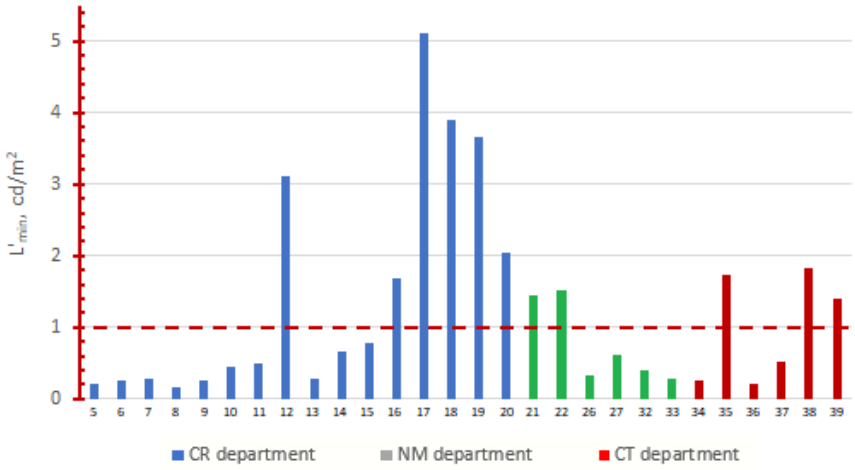

Figure 23. Measured L'$_{min}$ values for diagnostic display type. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. The blue dash line shows the minimum passable threshold value for L'$_{min}$ criteria. Numbers correspond to the monitor models in table 5.

3.1.4. Maximum luminance (L'_{max})

In figure 24 and 25, maximum luminance measurements are presented (diagnostic and clinical review type, respectively). According to the new guidelines, the threshold value of L'max for diagnostic display type should reach 300 cd/m^2 and for clinical review displays L'_{max} value should be no less than 200 cd/m^2. The minimum measured value for diagnostic displays was 113.1 cd/m^2, the mean value was 335.25 cd/m^{2,} and the maximum 794.9 cd/m^2. For clinical review type displays minimum measured L'max value was 78.22 cd/m^2, the mean value was 173.78 cd/m^2, and the maximum value reached 207.2 cd/m^2. In total, 13 (8 from conventional radiology, 1 from nuclear medicine, and 2 from CT departments) monitors passed this criterion, 11 of which were diagnostic type (those two clinical review type displays, which passed this criterion, were brand new in nuclear medicine department). It should be noted that some of the monitors in the hospital are working all the time. Some are very old (especially in the conventional radiology department), as shown in the presented graph below (figure 24)), which might cause backlight luminance degradation. Also, some radiologists prefer to use monitors not at its brightest settings, otherwise they experience eyestrain.

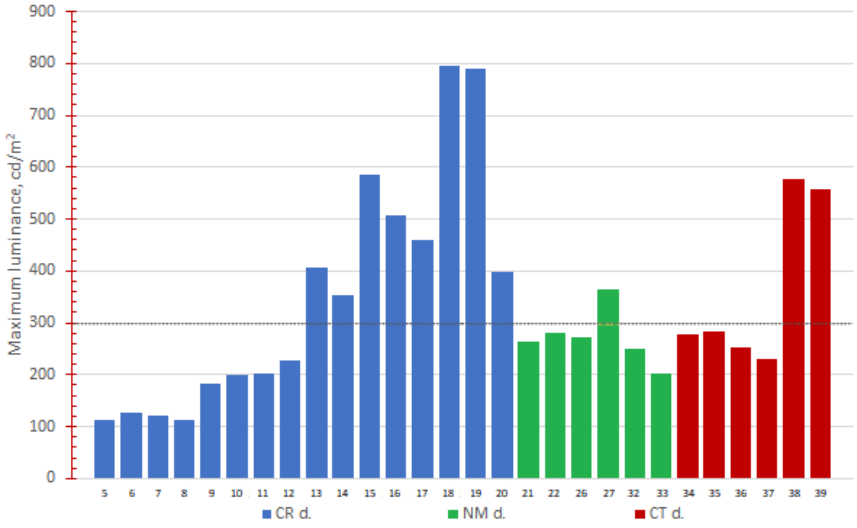

Figure 24. Diagnostic monitor type maximum luminance measurements. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. The grey dotted line shows the minimum recommended passable threshold value for L'_{max} criteria. Numbers correspond to the monitor models in table 5.

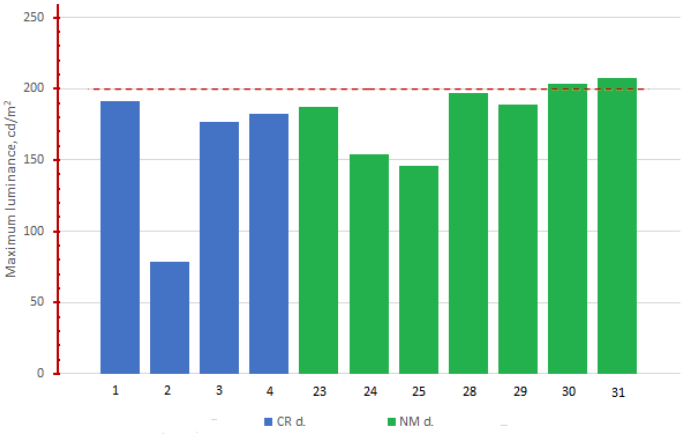

Figure 25. Clinical review monitor type maximum luminance measurements. CR – conventional radiology, NM – nuclear medicine. The red dash line shows the minimum passable recommended threshold value for the L'_{max} criteria. Numbers correspond to the monitor models in table 5.

3.1.5. Luminance response absolute error percentage from the DICOM GSDF

It should be noted that initial measurements, especially for clinical review type monitors, should not show that they are DICOM GSDF compliant, as these monitors do not have DICOM function, nor they were calibrated to it. The absolute error calculated in such steps: at first using formula 4 delta JND per grey level was calculated. After this, average JND value obtained by subtracting maximum and minimum luminance and dividing it from total possible grey values that monitor can produce. The obtained JND value was subtracted from the average JND value and divided from the same average JND value following with the final result multiplication to 100, converting final value to the percent.

In total, 26 monitors failed the luminance response test, including all 11 clinical review type displays to this count (because clinical review type monitors have less strict recommended DICOM GSDF deviation values (±20%), they were evaluated separately from there).

Thirteen diagnostic monitors passed, and 15 failed this criterion. 8 monitors passed this criterion were located in the conventional radiology department, 2 in nuclear medicine, and 3 in CT department. One sample t-test was performed for the diagnostic monitors to evaluate whether the true mean from the measured sample exceeds recommended 10% value. The test showed, that the true mean deviation of the tested sample values was 27.23321% ($t = 3.5242$, $df = 27$, $p = 0.0007679$). This means that the average DICOM GSDF deviation from all measured diagnostic type monitors in overall was significantly higher than 10%. After initial measurements, the minimum deviation of diagnostic monitors was 1.38%, maximum for this type reached 78.64%.

All clinical review type monitors failed this criterion; therefore, no statistical analysis was performed, as in such a case, it is evident that the mean DICOM GSDF deviation was significantly higher than 20%. After initial measurements, the minimum deviation of clinical review monitors was 26.62%, the maximum for this type reached 56.47%, and the mean deviation of all measured monitors of this type reached 39.67% deviation.

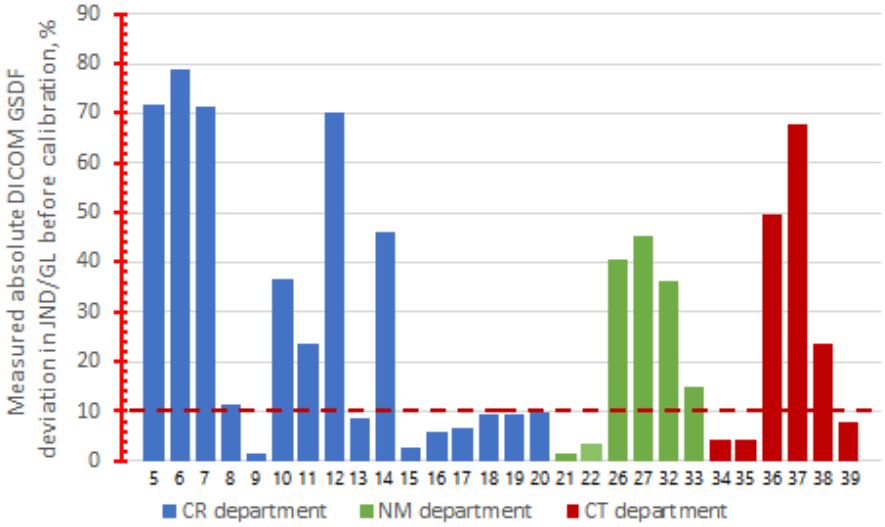

Figure 26. Diagnostic monitor type average JND/GL deviation percentage from DICOM GSDF. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. The red dash line shows the minimum passable threshold value for the L'_{max} criteria. Grey dashed line indicated the maximum recommended deviation for diagnostic displays. Numbers correspond to the monitor models in table 5.

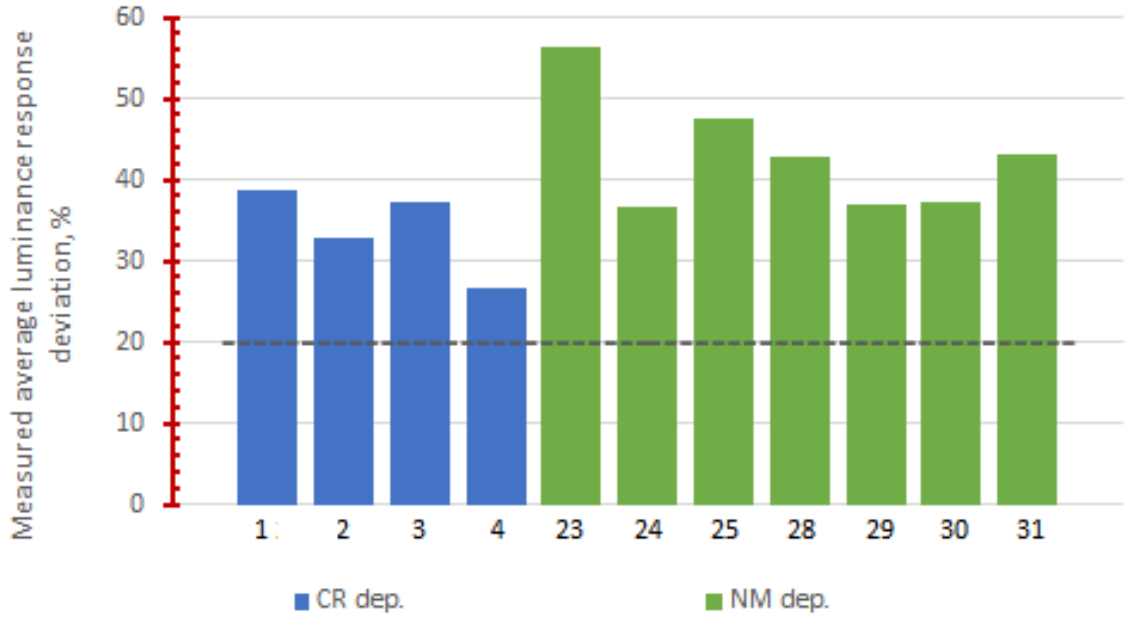

Figure 27. Clinical review display type average JND/GL deviation percentage from DICOM GSDF. CR – conventional radiology, NM – nuclear medicine. The red dash line shows the minimum passable threshold value for the L'_{max} criteria. Grey dashed line indicated the maximum recommended deviation for diagnostic displays. Numbers correspond to the monitor models in table 5.

3.1.6. Luminance and contrast ratio

The luminance ratio was calculated by dividing monitors minimum luminance and ambient lighting sum from maximum luminance and ambient light sum ($\frac{L'_{max}}{L'_{min}}$). The same equation was used to calculate contrast ratio, just without adding ambient light ($\frac{L_{max}}{L_{min}}$).

As the AAPM TG 270 report states, recommended values for the luminance ratio for diagnostic displays are in the range of 250 and 450. According to the report, an extremely high luminance ratio exceeds the human eye's abilities and thus is unnecessary; therefore, an upper limit is added.

The contrast ratio is a good starting point to determine the luminance ratio. If the contrast ratio does not achieve the minimum ratio of 250, it means that the luminance ratio on such a monitor will not pass this criterion. For the clinical review display type, the variation in contrast and luminance ratio is not that drastic, especially compared with the diagnostic displays (clinical review type display CR ranges between 565.53 and 1200.6, average value = 626.73, LR ranges between 372.48 and 967.14, average value = 869.88, diagnostic type display CR ranges between 174.8 and 1921.1, average value = 712.015, LR range is 73.53 - 1379.59, average value = 472.52). In total, 10 monitors pass this criterion, 2 of which are clinical review type. However, it is not appropriate to rely only on luminance ratio parameters to determine if this criterion is passed, as some monitors were measured in a relatively bright environment; therefore, the luminance ratio parameter might be inaccurate. A good example might be the 8th measurement in the diagnostic monitor results(figure 28), where luminance ratio and contrast ratio difference is drastic, meaning that ambient light should be decreased to achieve a better luminance ratio.

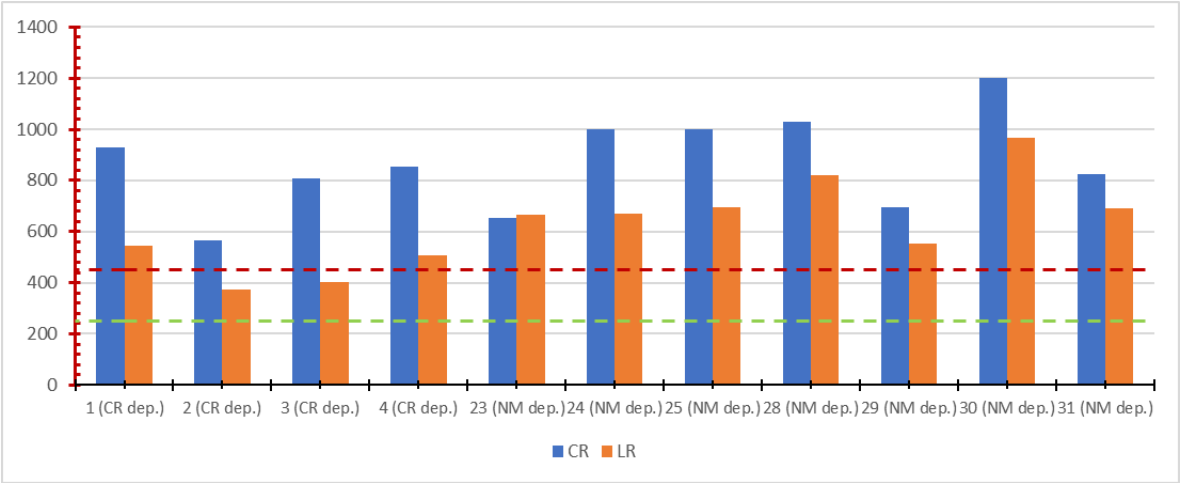

Figure 28. Contrast (CR) and luminance (LR) ratios of measured clinical review type displays. CR – conventional radiology, NM – nuclear medicine. Numbers indicate model number from table 5.

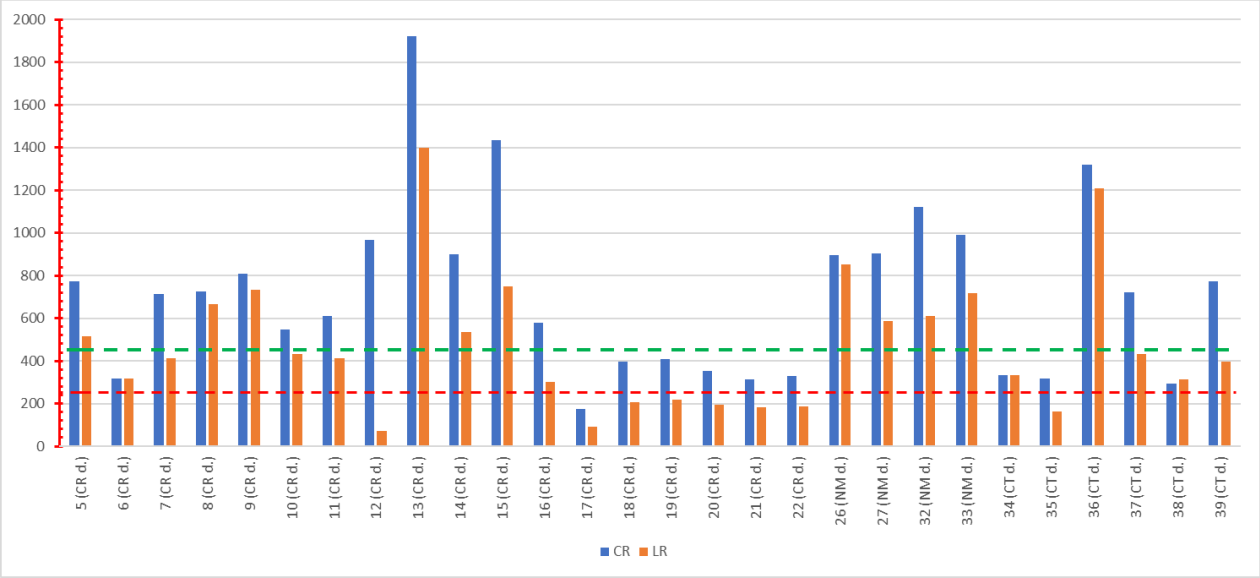

Figure 29. Contrast (CR) and luminance (LR) ratios of measured diagnostic type displays. CR d. – conventional radiology department, NM – nuclear medicine department, CT – computed tomography department.

3.1.7. Uniformity

Only quantitative results for monitor uniformity in this work will be presented as no non-uniformities were spotted in any of the displays by the naked eye. The monitor quantitative uniformity should not exceed 30% in both dark and bright test patterns (evaluated separately). The results calculated using formula 5, mentioned in section 1.3.4. of this work.

As seen in figure 29, only two monitors did not pass this criterion and only under bright background. On average, under light background monitor, backlight uniformity deviated 14.675%, the minimum value was 2.46% and the maximum 50.58%. Under dark background, monitor backlight uniformity deviated in the range of 1.8 to 23.4 %; the average value was 11.65%. Most of the measurements did not exceed 20% deviation (69 out of 78, 37 under dark background and 32 under light background), 23 (11 under light background, 12 under dark background) out of 78 measurements did not exceed

10% deviation. This time, most of the major nonuniformity deviations were present in a light background.

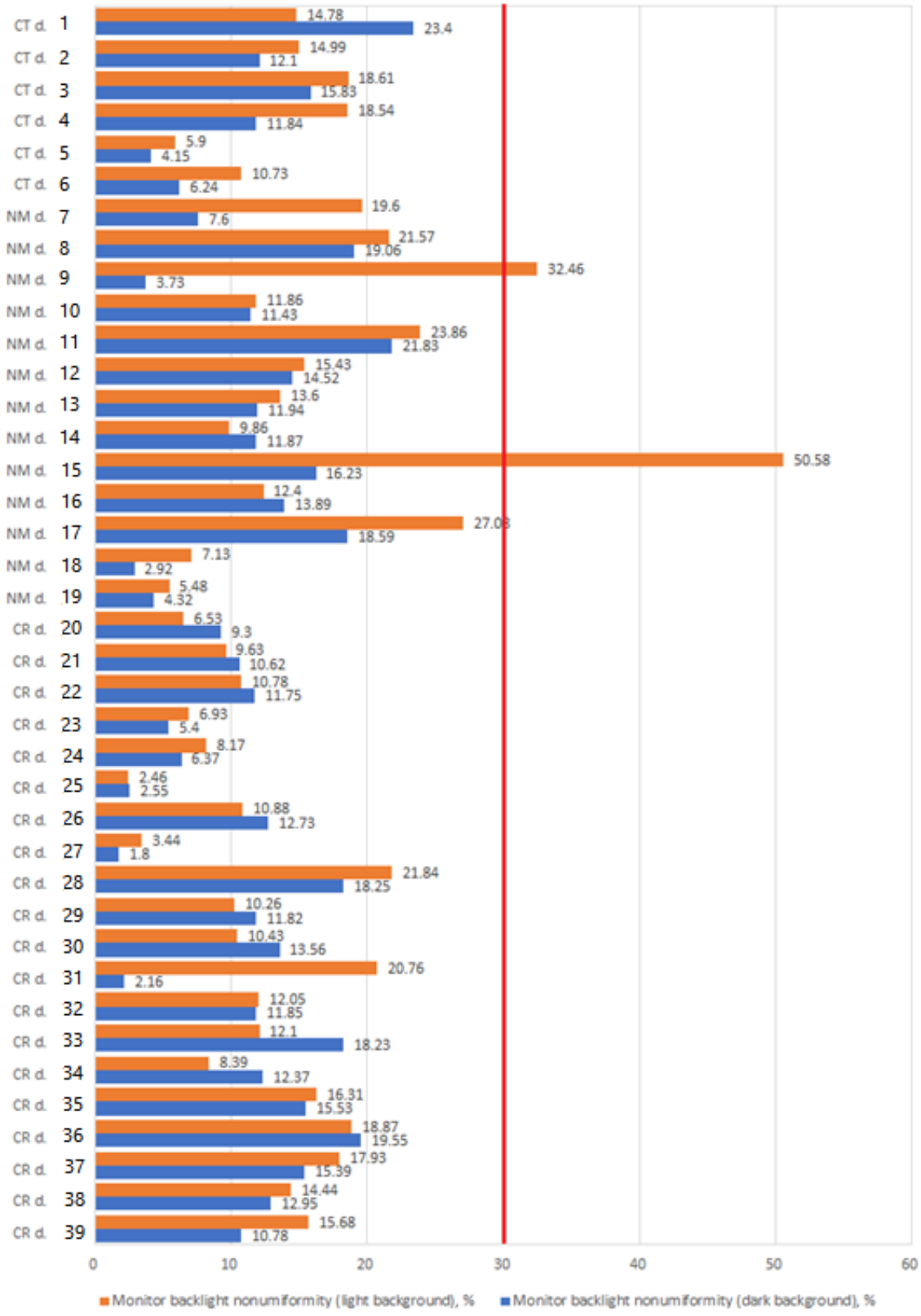

Figure 29. Monitor backlight nonuniformity measurements. As there are no classifications for this criterion, monitors were not classified in this work for this parameter. CR d. – conventional radiology department, NM – nuclear medicine department, CT – computed tomography department. Numbers indicate monitor model number from table 5.

3.2. Initial qualitative measurements

3.2.1. Qualitative veiling glare

In this work veiling glare parameter was evaluated only visually. The main reason for this was that there was no appropriate equipment to evaluate this parameter quantitatively. It was counted how many out of five circular figures in the center are visible in visual evaluation. AAPM report recommends that for diagnostic displays, at least 3 out of 5 figures and for clinical review displays, at least 2 out of 5 should be visible in the TG18-GV and TG18-GVN test patterns.

Table 6. Qualitative Veiling Glare results for clinical review display type.

Vendor	Model	Year of manifactory	TG18-GV	TG18-GVN
Dell	U2415	2018	4	5
Dell	P2417H	2019	3	5
HP	EliteDisplayE221C	2013	2	5
Dell	U2415	2018	3	4
HP	LA2306x	2012	3	3
Dell	P2412H	2018	3	5
Dell	P2412H	2018	3	5
Dell	P2417H	2019	4	5
Dell	U2412M	2018	3	5
Dell	P2719H	2019	3	5
Dell	U2412M	2016	3	5

As seen in table 6, no displays failed this criterion, although its effect is evident because almost all monitors performed worse in the TG18-GV pattern than TG18-GVN one.

Two diagnostic type displays failed this criterion out of 28, where none of the figures were visible in the TG18-GV pattern. In 10 diagnostic displays, three figures were visible on the same pattern, 7 displays of this type were good enough that it would be possible to distinguish four circular figures, and the remaining six monitors were able to show this test pattern so that all five figures were distinguished. In the TG18-GVN pattern, it was possible to see 5 out of 5 figures in 23 diagnostic monitors; in two monitors, it was impossible to distinguish any of the figures; in 1 monitor, it was possible to see 4 out of 5 figures, and another one showed a pattern with three distinguishable figures and in one monitor 2 out of 5 figures distinguished. For more details, see table 7.

Table 7. Qualitative Veiling Glare results for diagnostic display type.

Vendor	Model	TG18-GV	TG18-GVN
NEC	LCD1990SXi	4	5
NEC	LCD1990SXi	0	0
NEC	LCD1990SXi	0	0
NEC	LCD1990SXi	4	5
NEC	PA272W	3	5
Eizo	Radiforce MX193	4	5
Eizo	Radiforce MX193	4	5
Barco	MDRC-2324	3	5
Barco	MDCC-6530	4	5
Barco	MDRC-2324	4	5
Barco	MDCC-6530	3	5

Table 7 (continued). Qualitative Veiling Glare results for diagnostic display type.

Vendor	Model	TG18-GV	TG18-GVN
Eizo	Radiforce GX540	5	5
Eizo	Radiforce GX540	4	5
Wide	2210E	5	5
Wide	2210E	5	5
Wide	JMW1100KB2449F03	5	5
JVC	CL-R211	3	5
JVC	CL-R211	3	3
Eizo	EV2456	5	5
Eizo	EV2456	5	5
Lenovo	T27Q	3	5
Lenovo	T27Q	3	5
JVC	CL-R211	3	5
JVC	CL-R211	5	5
Eizo	FlexScan S1923	0	4
Eizo	FlexScan S1923	0	2
NDS	Dome E2	3	5
NDS	Dome E2	3	5

3.2.2. Noise

Display noise was visually evaluated using the TG18-AFC pattern. According to the AAPM standards, in 3 out of 4 quadrants should be possible to see all of the square figures for the diagnostic display type and for clinical review display type at least 2 out of 4.

Out of 28 diagnostic display evaluations, all passed this criterion. In 23 displays, three quadrants with all details were fully visible, and five displays were able to produce an image, where all four quadrants were visible, including all desired details.

Out of 11 clinical review type display evaluations, all passed this criterion as well; in 10 displays, three quadrants with all details were fully visible, and in 1 display, four quadrants were fully visible.

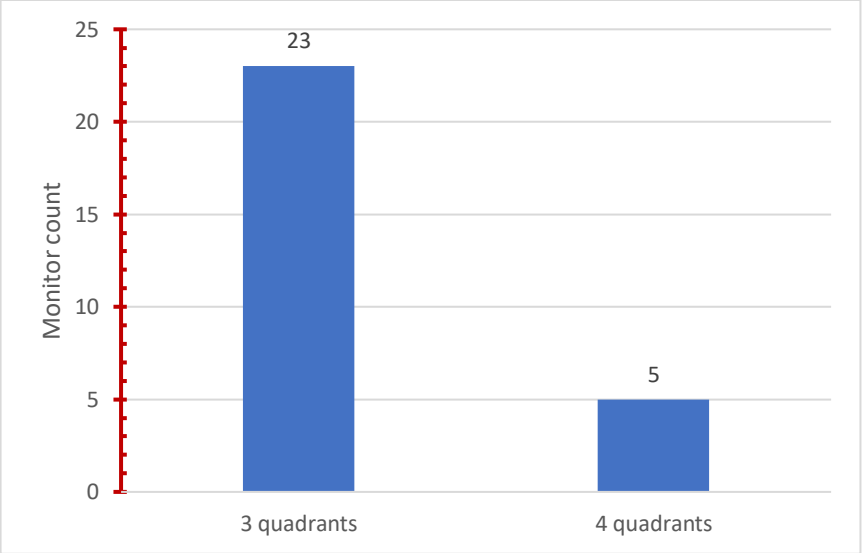

Figure 30. Diagnostic display category noise visual evaluation results. The column graph shows monitor count based on how many quadrants were fully visible with all required details.

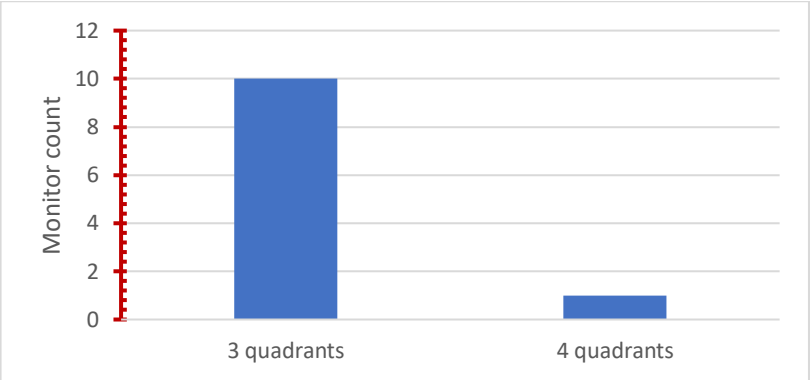

Figure 31. Clinical review display category noise visual evaluation results. The column graph shows monitor count based on how many quadrants were fully visible with all required details.

3.3. Post-calibration monitor parameter measurements and its change evaluation

3.3.1. Illuminance measurements after calibration

After the calibration, measured illuminance changed. 18 monitor working places (of which 2 had the clinical review monitor type and 16 had diagnostic type) still did not pass recommended AAPM TG 270 illuminance criterion. One diagnostic monitor place from conventional radiology department exceeded recommended value range, and 14 were under the recommended value range (of which were 2 clinical review type display places). None of display places from the nuclear medicine department failed this criterion. In the CT department, 3 diagnostic display places were under the recommended illuminance range. The minimum value for the illuminance was 6.25 lx, the maximum reached 92.56 lx, and the mean value was 27.48 lx. After comparing all measured values of illuminance before and after the calibration using Student Two Sample t-test, the true difference in means was significant with the p-value of 0.008489 (t = 2.7519, df = 45.354).

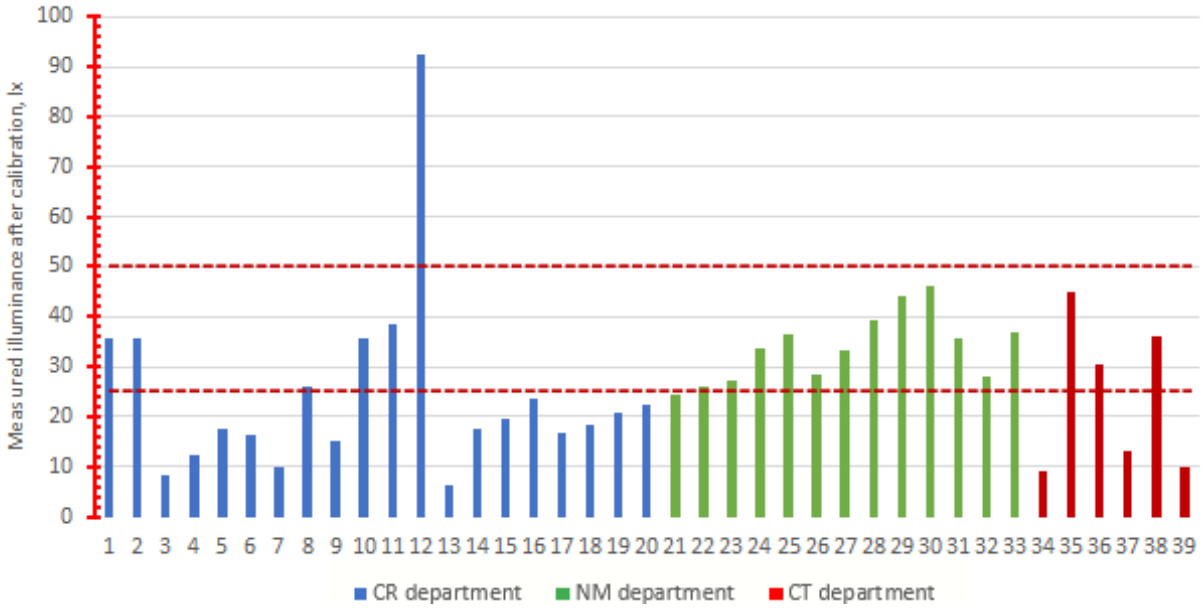

Figure 32. Measured illuminance values after the calibration. Red dashed lines indicate the range of AAPM recomended values of illuminance. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

3.3.2. Ambient luminance after calibration

After performing ambient luminance measurements, in total, 11 monitors passed this criterion (see figure 34). 7 of them were diagnostic type (3 of them were located in a conventional radiology department, 3 in the nuclear medicine department, and 1 in CT department) and 4 clinical review type (1 clinical review type displays was located in the conventional radiology department and 5 in the nuclear medicine department). The smallest measured illuminance value was 0.01 cd/m^2, the mean value was 0.2164 cd/m^2, and the maximum value was 1.53 cd/m^2. After comparing all measured values of ambient luminance before and after the calibration using Student Two Sample t-test, the true difference in means was insignificant with the p-value of 0.07507 (t = 1.8143, df = 55.215).

Just like before calibration, Pearson's product-moment correlation statistical data analysis showed that illuminance (after calibration) and ambient luminance (after calibration) had a moderately strong positive correlation, although a tiny bit higher (correlation coefficient = 0.6081267, p = 0.000402), meaning that this time ambient luminance parameter moderately strongly and slightly more, than before the calibration depended on illuminance.

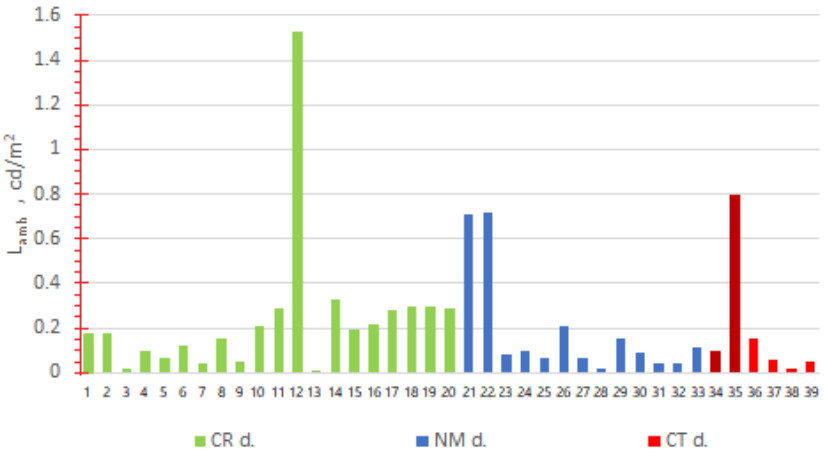

Figure 33. Measured ambient luminance values after the calibration. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

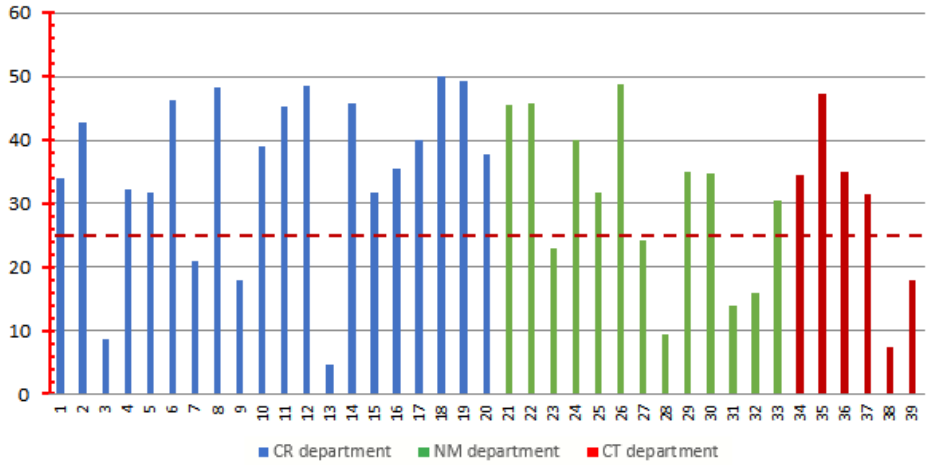

Figure 34. Measured ambient luminance composite percentage of L‘min. Dashed line indicates the threshold value of maximum recommended ambient luminance composite percentage of L‘min (25%). CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

3.3.3. Minimum luminance (L'$_{min}$)

After doing post-calibration measurements, none of the clinical review type displays have met the minimum (recommended by the AAPM TG270 report) threshold value of 0.8 cd/m^2. From the diagnostic type displays, only one met this criterion, exceeding 1 cd/m^2 threshold value, but only due to high illuminance (92.56 lux for this monitor was the highest illuminance value measured after the calibration). For the clinical review type displays, the minimum L'$_{min}$ value after calibration was 0.15 cd/m^2 (while for the diagnostic monitors was 0.13 cd/m^2), the mean value was 0.2227 cd/m^2 (for diagnostic type it was 0.3789 cd/m^2), and the maximum value was 0.33 cd/m^2 (for diagnostic monitor type it was 1.62 cd/m^2). After comparing clinical review type display measured values of minimum luminance before and after the calibration using Student Two Sample t-test, the true difference in means was insignificant with the p-value of 0.3831, (t = 0.89251, df = 19.315). Comparing the true difference in means for diagnostic monitor type, Student Two Sample t-test showed significant difference (t = -2.1853, df = 39.233, p-value = 0.0349), which is due to changed ambient light conditions.

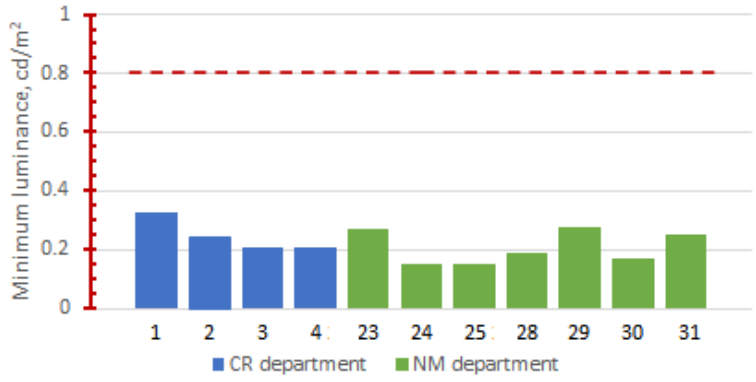

Figure 35. Measured minimum luminance values for clinical review type displays after the calibration. The red dashed line indicates the minimum passable criterion value. CR – conventional radiology, NM – nuclear medicine. Numbers indicate model number provided in table 5.

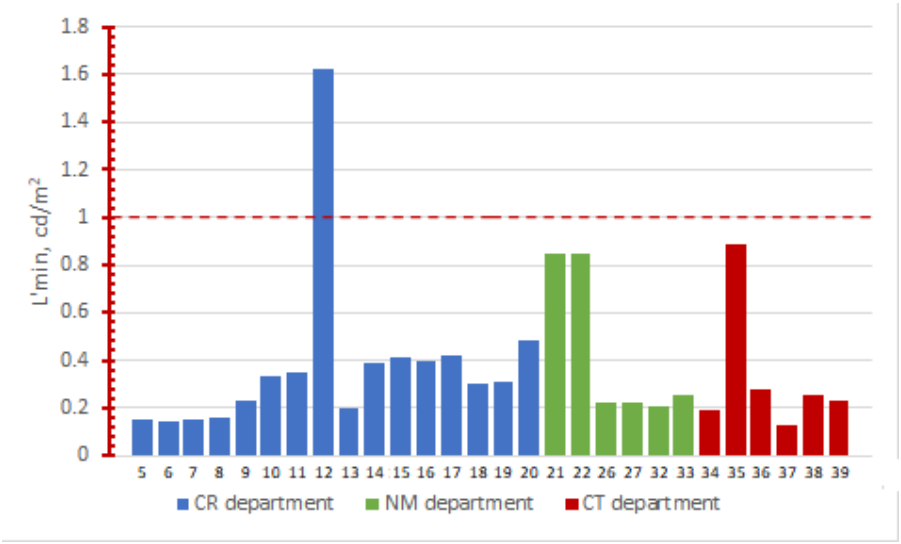

Figure 36. Measured minimum luminance values for diagnostic type displays after the calibration. The red dashed line indicates the minimum passable criterion value. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

3.3.4. Maximum luminance

The minimum acceptable value for diagnostic displays according to AAPM guidelines (300 cd/m^2) luminance passed 10 diagnostic displays (8 from conventional radiology, 2 from CT department), and for the clinical displays, minimum 250 cd/m^2 value was surpassed by 4 clinical review type displays, three of which belonged to the conventional radiology department. For the clinical review type displays, the minimum L'_{max} value after calibration was 150.5 cd/m^2 (while for the diagnostic monitors was 111.9 cd/m^2), the mean value was 201.8 cd/m^2 (for diagnostic type it was 322.5 cd/m^2), and the maximum value was 306.3 cd/m^2 (for diagnostic monitor type it was 794.9 cd/m^2). After comparing clinical review type display measured values of maximum luminance before and after the calibration using Student Two Sample t-test, the true difference in means was insignificant with the p-value of 0.1223, (t = -1.6158, df = 19.386). Comparing the true difference in means for diagnostic monitor type, Student Two Sample t-test showed insignificant difference (t = 0.2601, df = 53.855, p-value = 0.7958).

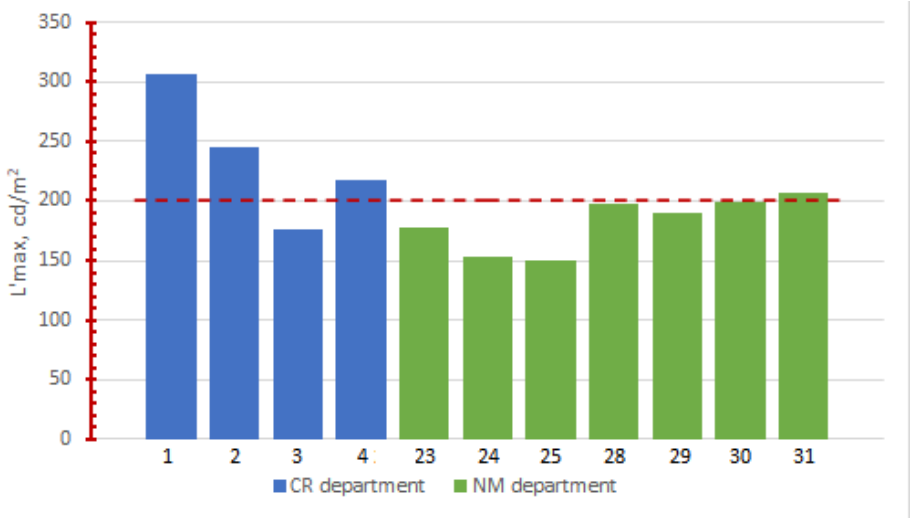

Figure 37. Measured maximum luminance values for clinical review type displays after the calibration. The red dashed line indicates the minimum passable criterion value. CR – conventional radiology, NM – nuclear medicine. Numbers indicate model number provided in table 5.

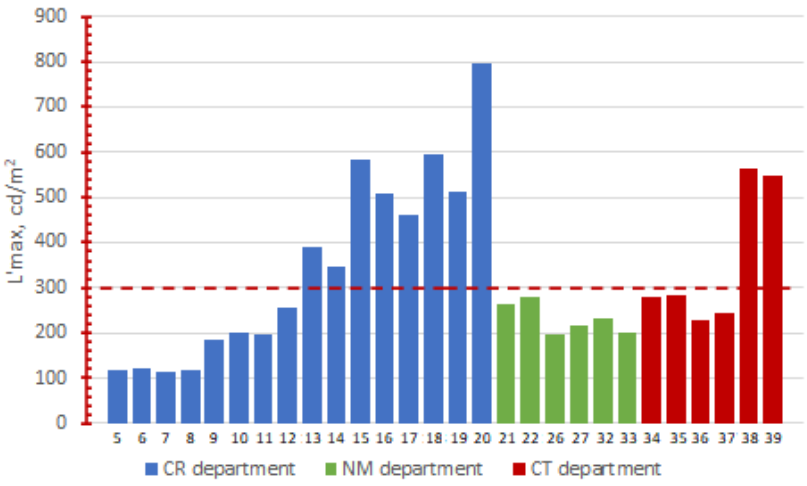

Figure 38. Measured maximum luminance values for diagnostic type displays after the calibration. The red dashed line indicates the minimum passable criterion value. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

3.3.5. Luminance response absolute deviation percentage in just noticeable difference per grey level (JND/GL) after calibration

From all measurements after calibration, four monitors in total (3 from conventional radiology, one from CT department) failed the luminance response test after calibration, all of them were diagnostic type. These monitors were impossible to calibrate. Part of the reason was that they were very old and worked for a very long time.

One sample t-test was performed for the diagnostic monitors to evaluate whether the true mean from the measured sample exceeds recommended 10% value. The test showed, that the true mean deviation of the tested sample values was 6.020714% ($t = -5.4123$, $df = 27$, $p = 5.04 \times 10^{-6}$). This means that the average DICOM GSDF deviation from all measured diagnostic type monitors in overall was not significantly higher than 10%. After initial measurements, the minimum deviation of diagnostic monitors was 1.38%, maximum for this type reached 15.07%.

All clinical review type monitors passed this criterion; therefore, no statistical analysis was performed, as in such a case, it is evident that the mean DICOM GSDF deviation was lower than 20%. After initial measurements, the minimum deviation of clinical review monitors was 2.13%, the maximum for this type reached 18.34%, and the mean deviation of all measured clinical review type monitors of this type reached 7.856% deviation.

After comparing clinical review type display measured values of DICOM GSDF absolute deviation percentage in just noticeable difference per grey level (JND/GL) before and after the calibration using Student Two Sample t-test, the true difference in means was significant with the p-value of 2.128×10^{-9}, ($t = 11.054$, $df = 17.791$). Comparing the true difference in means for diagnostic monitor type, Student Two Sample t-test showed significant difference aswell ($t = 4.2897$, $df = 28.22$, p-value = 0.0001898). This means that both for clinical and diagnostic monitors, calibration improved DICOM GSDF parameter, according to AAPM recomendations,by significantly reducing its absolute deviation percentage in just noticeable difference per grey level.

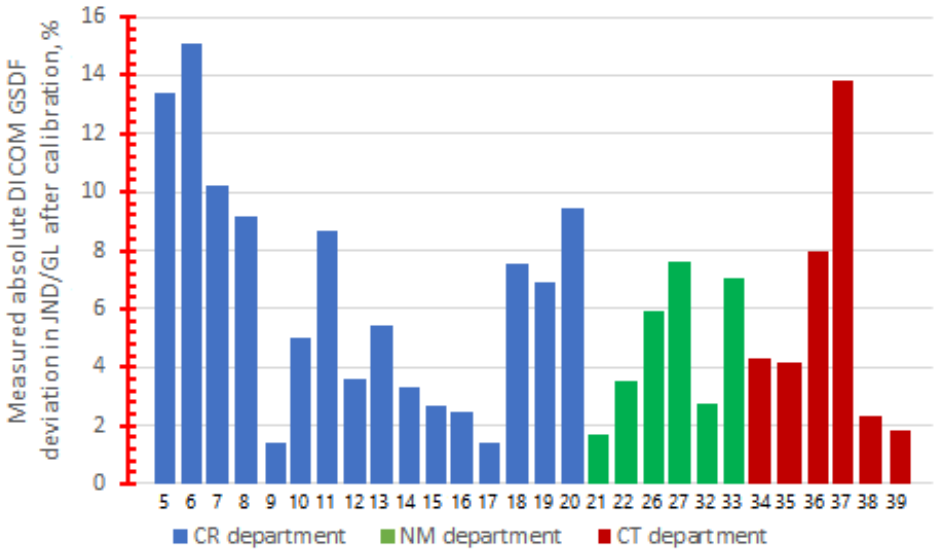

Figure 39. Measured absolute DICOM GSDF deviation in JND/GL percentage for diagnostic display type after calibration. Numbers indicate monitor model from table 5. CR – conventional radiology, NM – nuclear medicine, CT – computed tomography. Numbers indicate model number provided in table 5.

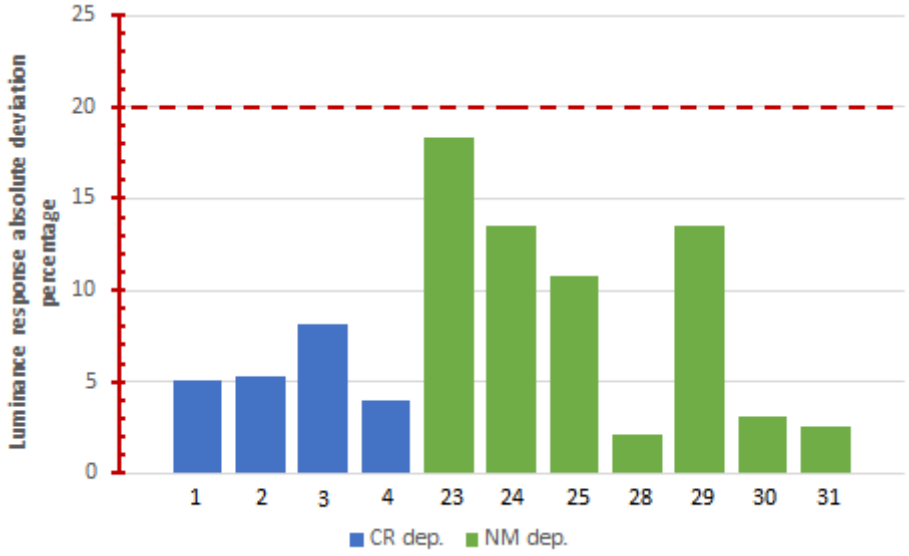

Figure 40. Measured absolute DICOM GSDF deviation in JND/GL percentage for clinical review display type after calibration. CR – conventional radiology, NM – nuclear medicine. Numbers indicate model number provided in table 5.

Comparing all monitor average luminance response deviation percentage per each DDL (as shown in figure 41) it is evident, that each DDL error decreased. The most noticeable change (14.82 % improvement) was in 15th display driving level (DDL 15), where the error was the largest after initial measurements (19.59%). Also, the smallest change (2.05% improvement) was in 120th display driving level, which after initial measurements had the smallest error from the ideal result (5.34%).

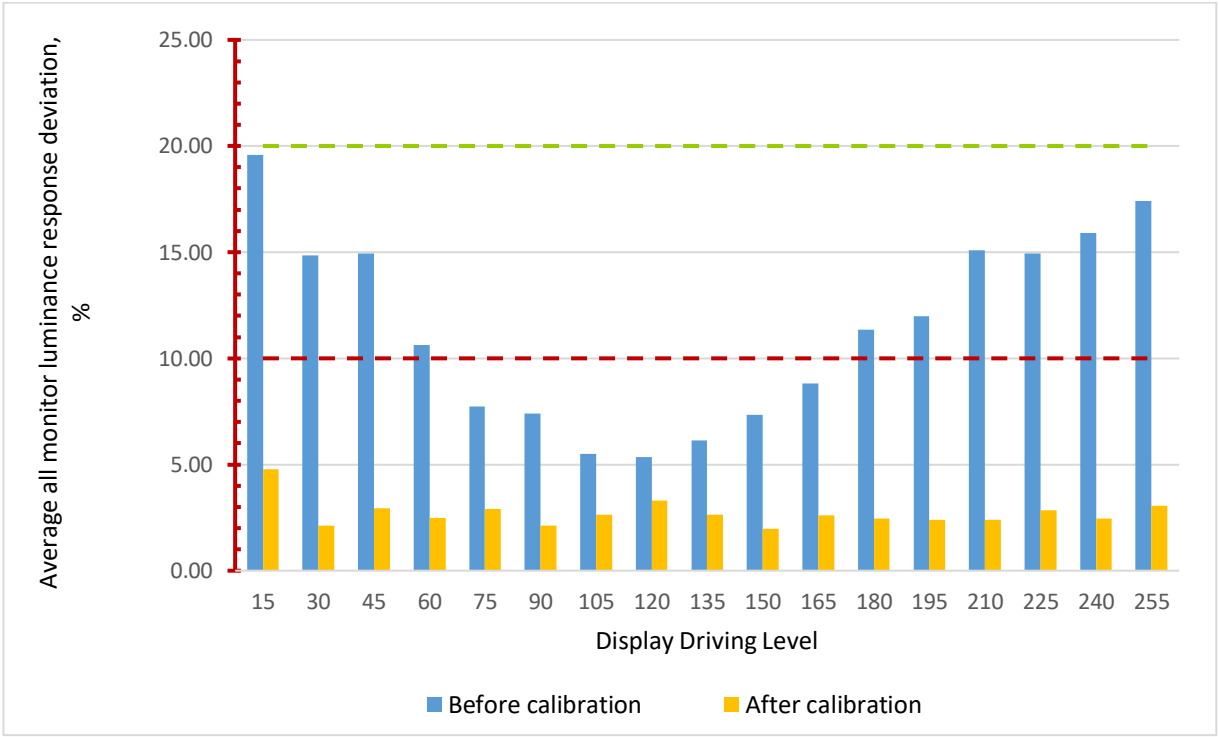

Figure 41. Average all monitor luminance response deviation (measured in percent) per each DDL (display driving level) before and after the calibration. Red dashed lane indicates the maximum allowed average error for diagnostic type displays; green dashed line indicates maximum allowed error for clinical review (secondary) type of displays.

3.3.6. Luminance and contrast ratio

For the clinical review display type, after the calibration the variation in contrast and luminance ratio increased - CR ranges between 1263 and 3699, average value = 3506, LR ranges between 656.2 and 1193.8, average value = 928.7, diagnostic type display CR ranges between 166.6 and 5790, average value = 2077.6, LR range is 312.5 – 2383.1, average value = 1084.8.

In total, 4 diagnostic monitors pass this criterion and none from the clinical review display category.

Although it seems like the overall situation for contrast ratio got worse, after comparing these parameters using Student Two Sample t-test, the true difference in means of the contrast ratio before and after calibration was insignificant (t = 1.2201, df = 48.349, p-value = 0.2283).

For the luminance ratio, the parameter in overall got worse, as true difference in means got worse (t = -2.831, df = 68.168, p-value = 0.006093).

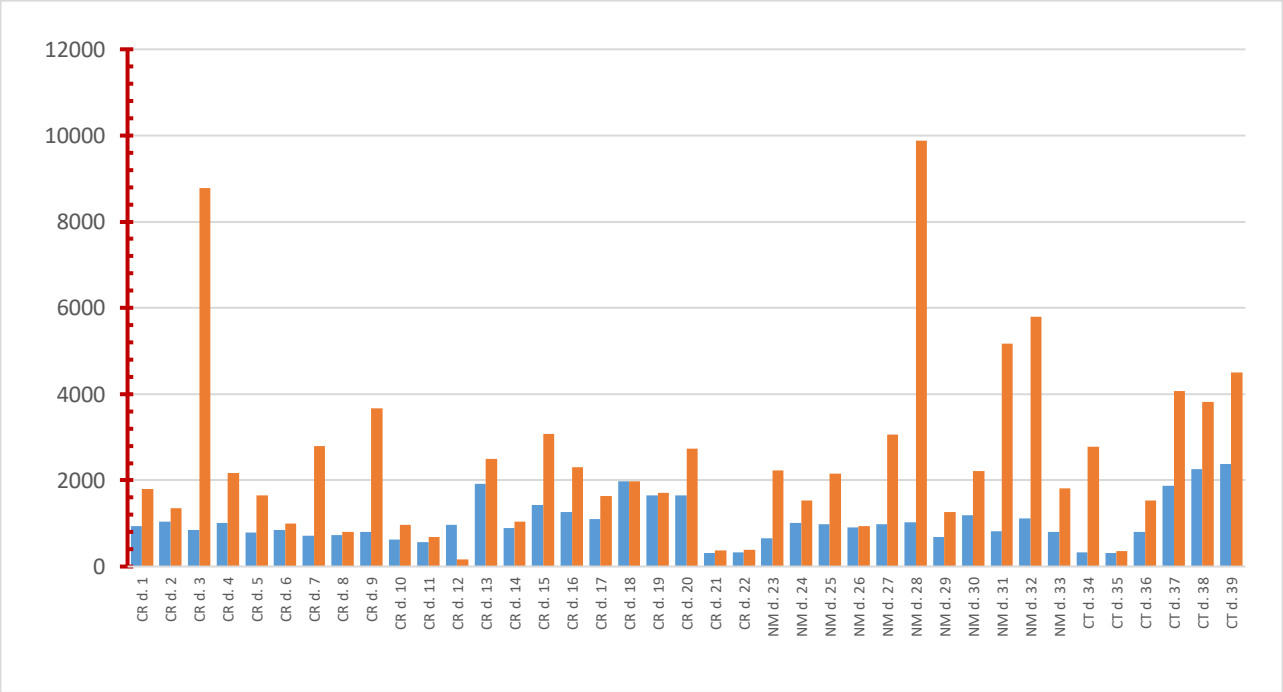

Figure 42. Luminance ratio (blue bars) and contrast ratio (orange bars) measurements after calibration. Numbers indicate monitor model (provided in the table 5), CR d. – conventional radiology department, NM d. – nuclear medicine department, CT d. – computed tomography department.

3.3.7. Quantitative re-evaluation (Veiling glare and Noise)

After display calibration, noise parameters got worse for 3 displays (evaluation performed the same as before the calibration) (1 clinical review type in the conventional radiology department and 2 diagnostic type in nuclear medicine department) and improved for two clinical review type displays nuclear medicine department. For the veiling glare, in total 4 monitors showed worse results (1 element from the TG18-GV disappeared for 2 clinical review monitors (both were in conventional radiology department) and for two diagnostic monitors, located in nuclear medicine department) and eight monitors (6 of them were diagnostic type) showed improved results (for more details, see table 8). Evaluating veiling glare with TG18-GVN (black background), 6 monitors improved in total (2 of which were clinical review type, for more details, see table 8).

Table 8. Veiling glare and noise element visibility after calibration. Brackets indicate the change of element visibility after calibration. CR d. Means that the measurements were performed in conventional radiology department, NM d. – nuclear medicine department, CT d. – computed tomography department.

Department (monitor type)	Vendor	Model	TG18-AFC	TG18-GV	TG18-GVN
CR d. (Clinical)	Dell	U2415	3(-1)	3(-1)	5
CR d. (Clinical)	Dell	P2417H	3	4(+1)	5
CR d. (Clinical)	HP	EliteDisplayE221C	3	4(+2)	5
CR d. (Clinical)	Dell	U2415	3	3	5(+1)
CR d. (Diagnostic)	NEC	LCD1990SXi	3	4	5
CR d. (Diagnostic)	NEC	LCD1990SXi	3	4(+4)	5(+5)
CR d. (Diagnostic)	NEC	LCD1990SXi	3	5(+5)	5(+5)
CR d. (Diagnostic)	NEC	LCD1990SXi	3	4	5
CR d. (Diagnostic)	NEC	PA272W	3	3	5
CR d. (Diagnostic)	Eizo	Radiforce MX193	3	5(+1)	5
CR d. (Diagnostic)	Eizo	Radiforce MX193	3	5(+1)	5
CR d. (Diagnostic)	Barco	MDRC-2324	3	3	5
CR d. (Diagnostic)	Barco	MDCC-6530	4	4	5
CR d. (Diagnostic)	Barco	MDRC-2324	3	3(-1)	5
CR d. (Diagnostic)	Barco	MDCC-6530	3	3	5
CR d. (Diagnostic)	Eizo	Radiforce GX540	3	5	5
CR d. (Diagnostic)	Eizo	Radiforce GX540	3	4	5
CR d. (Diagnostic)	Wide	2210E	3	5	5
CR d. (Diagnostic)	Wide	2210E	3	5	5
CR d. (Diagnostic)	Wide	JMW1100KB2449F03	3	5	5
CR d. (Diagnostic)	JVC	CL-R211	3	3	5
CR d. (Diagnostic)	JVC	CL-R211	3	3	3
NM d. (Clinical)	HP	LA2306x	3(+1)	3	5(+2)
NM d. (Clinical)	Dell	P2412H	3	3	5
NM d. (Clinical)	Dell	P2412H	4(+1)	4	5
NM d. (Diagnostic)	Eizo	EV2456	3	4(-1)	5
NM d. (Diagnostic)	Eizo	EV2456	3(-1)	5	5
NM d. (Clinical)	Dell	P2417H	3	3(-1)	5
NM d. (Clinical)	Dell	U2412M	3	3	5
NM d. (Clinical)	Dell	P2719H	3	3	5
NM d. (Clinical)	Dell	U2412M	3	3	5
NM d. (Diagnostic)	Lenovo	T27Q	3(-1)	3	5
NM d. (Diagnostic)	Lenovo	T27Q	3	3	5
CT d. (Diagnostic)	JVC	CL-R211	3	3	5
CT d. (Diagnostic)	JVC	CL-R211	4	5	5
CT d. (Diagnostic)	Eizo	FlexScan S1923	3	3(+3)	5(+1)
CT d. (Diagnostic)	Eizo	FlexScan S1923	3	4(+4)	5(+3)
CT d. (Diagnostic)	NDS	Dome E2	4	3	5
CT d. (Diagnostic)	NDS	Dome E2	3	3	5

Table 9. Passed acceptance criteria before and after the calibration.

	Before calibration		After calibration	
	Clinical	Diagnostic	Clinical	Diagnostic
Illuminance	11 out of 39		19 out of 39	
Suggested passing interval by AAPM, lx	25-75 lx			
Ambient light	3 out of 11	5 out of 28	3 out of 11	5 out of 28
AAPM suggested passing criteria ratio with L'min (minimum luminance)	≤1/4 of L'min			
L' min (minimum luminance)	0 out of 11	11 out of 28	0 out of 11	1 out of 28
AAPM suggested passing criteria, cd/m^2	Diagnostic (primary): >1 cd/m^2, Clinical review(secondary): >0.8 cd/m^2			
L' max (maximum luminance)	2 out of 11	11 out of 28	4 out of 11	10 out of 28
AAPM suggested passing criteria, cd/m^2	Diagnostic (primary): >300 cd/m^2, Clinical review(secondary): >250 cd/m^2			

Table 9 (continued). Passed acceptance criteria before and after the calibration.

	Before calibration		**After calibration**	
	Clinical	**Diagnostic**	**Clinical**	**Diagnostic**
DICOM deviation from GSDF	0 out of 11	13 out of 28	11 out of 11	24 out of 28
AAPM suggested passing criteria threshold value, %	Diagnostic (primary): < 10%, Clinical review(secondary): < 20%			
Luminance ratio	2 out of 11	9 out of 28	0 out of 11	0 out of 28
AAPM recommended value	Diagnostic (primary) : 250-450; Clinical review (secondary): 250-450			
Uniformity	37 out of 39 (maximum allowed luminance deviation from the median value of AAPM should not exceed 30%)			

Conclusions

In this work, 39 monitors, 28 of which were diagnostic (primary) and 11 were clinical review (secondary) type were evaluated, and their performance parameters before and after calibration were compared. The measurements were performed in three different departments (nuclear medicine, conventional radiology and computed tomography).

1. Display performance was optimized by using PerfectLum 4.0™ software and Xrite i1Pro photometer. The photometer measurements of display driving level values were successfully changed for all monitors by the PerfectLum 4.0™ software and stored in the graphics processing unit memory.

2. After initial measurements, none of the clinical review type monitors passed luminance response test and only 13 out of 28 diagnostic displays passed this criterion. Minimum luminance (L‘min) before the calibration was passed by 11 monitors and maximum (L‘max) – by 13. Luminance uniformity was passed by 37 monitors out of 39. After the calibration, all clinical review type monitors passed luminance response test and 4 out of 28 diagnostic displays failed this criterion. Minimum luminance before the calibration was passed by 1 monitors and maximum – by 14. Calibrated monitors performed noticeably better only in luminance response (quantitative) and veiling glare (qualitative) parameters. For luminance response average all monitor error (measured in percent) per each DDL value, the 15th DDL value improved the most, as it had biggest overall error over all measured monitors before the calibration. Other display parameters depended on other factors: ambient luminance depended on illuminance, minimum luminance parameter depended on monitor age, as the minimum luminance value due to too dim monitor values was affected by the ambient luminance too much, therefore, L‘min was also related to illuminance. Maximum luminance parameter mostly depended with display‘s total working time and technical specifications.

3. Ambient minimum and maximum luminance criteria failed because of illuminance variations and monitor age. For recommendations, it is essential to anually measure monitor maximum luminance (L'_{max}), as well as invest in better equipment for blocking external light in order to improve illuminance control, which might affect ambient (L_{amb}) and minimum (L'_{min}) luminance results. When it comes to annual measurements, it is important to check, whether monitor luminance response results satisfy AAPM recommended values. Also, it is important to plan, how monitors should be placed in the room (especially where diagnostic displays are) to reduce the chance of the reflection occurrence from other monitors.

List of references

1. Willemink M. J., Noël P. B. (2018). The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence. European Radiology.
2. Tabatabaei M. S., Langarizadeh M., Tavakol K. (2017). An Evaluation Protocol for Picture Archiving and Communication System: a Systematic Review. ACTA INFORM MED. 2017 DEC; 25(4): 250-253.
3. Bevins N. B., Flynn M. J., Silosky M.S., Marsh R. M., Walz-Flannigan A. I., Badano A. (2019). Display Quality Assurance. The Report of AAPM Task Group 270. ISBN: 978-1-936366-66-8 ISSN: 0271-7344 American Association of Physicists in Medicine.
4. Gennaro G., Avramova-Cholakova S., Azzalini A., Luisa Chapel M., Chevalier M., Ciraj O., Torresin A. (2018). Quality Controls in Digital Mammography protocol of the EFOMP Mammo Working group. Physica Medica, 48, 55–64.
5. Bevins N. B., Silosky M. S., Badano A., Marsh R. M., Flynn M. J., Walz-Flannigan A. I. (2020). Practical application of AAPM Report 270 in display quality assurance: A report of Task Group 270. Medical Physics.
6. Samei E., Badano A., Chakraborty D., Compton K., Cornelius C., Corrigan K., Willis C. E. (2005). Assessment of display performance for medical imaging systems: Executive summary of AAPM TG18 report. Medical Physics, 32(4), 1205–1225.
7. Pollard B. J., Samei E., Chawla A. S., Beam C., Heyneman L. E., Koweek L. M. H., McAdams H. P. (2012). The Effects of Ambient Lighting in Chest Radiology Reading Rooms. Journal of Digital Imaging, 25(4), 520–526.
8. Silosky M. S., Marsh R. M. (2018). Performance characteristics and quality assurance considerations for displays used in interventional radiology and cardiac catheterization facilities. Journal of Applied Clinical Medical Physics, 19(5), 708–717.
9. Krupinski E. A. (2008). Medical Grade vs Off-the-Shelf Color Displays: Influence on Observer Performance and Visual Search. Journal of Digital Imaging, 22(4), 363–368.
10. Wheeler T. R. H., Clark M. G. (1992). CRT Technology. Color in Electronic Displays, 221–256.
11. Chen H.W., Lee J.H., Lin B.Y., Chen S., Wu S. T. (2017). Liquid crystal display and organic light-emitting diode display: present status and future perspectives. Light: Science & Applications, 7(3), 17168.
12. Yoshida H. (2012). Vertically Aligned Nematic (VAN) LCD Technology. Handbook of Visual Display Technology, 1485–1505.
13. Hong H. (2012). In-Plane Switching (IPS) Technology. Handbook of Visual Display Technology, 1469–1483.
14. Ekpo E. U., McEntee M. F. (2015). An Evaluation of Performance Characteristics of Primary Display Devices. Journal of Digital Imaging, 29(2), 175–182.
15. Digital Imaging and Communications in Medicine (DICOM) Part 14: Grayscale Standard Display Function, National Electrical Manufacturers Association, 2011.
16. Hertel D., Penczek J. (2020). Evaluating Display Reflections in Reflective Displays and Beyond. Information Display, 36(2), 14–24.
17. Liu P., Badano A. (2014). Characterizing the reflectivity of handheld display devices. Medical Physics, 41(8Part1), 081909.

18. Ikushima Y., Yabuuchi H., Morishita J., Honda H. (2013). Analysis of Dominant Factors Affecting Fatigue Caused by Soft-Copy Reading. Academic Radiology, 20(11), 1448–1456.
19. Roehrig H., Krupinski E. A., Chawla A. S., Fan J., Gandhi K. (2003). Spatial noise and threshold contrasts in LCD displays. Medical Imaging 2003: Image Perception, Observer Performance, and Technology Assessment.
20. Badano A., Gagne R. M., Jennings R. J., Drilling S. E., Imhoff B. R., Muka E. (2004). Noise in flat-panel displays with subpixel structure. Medical Physics, 31(4), 715–723.
21. Flynn M. J., Badano A. (1999). Image quality degradation by light scattering in display devices. Journal of Digital Imaging, 12(2), 50–59.
22. Martínez-Domingo M.A. (2017). Multispectral High Dynamic Range Polarimetric Imaging applied to scene segmentation and object classification, PhD Physics and Space Sciences.
23. Gi-Gyeong S., Dong-Wook S., Jae-Ho J., Hui-Doo K., Kyung-Nam R. (2008) A Study on Production and Its Usefulness of AAPM TG18 Guiding Instrument for Diagnostic Monitor QC. Journal of the Korean Society for Digital Imaging in Medicine, 10(1), 21-28.
24. Blume H. R., Ho A. M. K., Stevens F., Steven P. M. (2001). Practical aspects of grayscale calibration of display systems. Medical Imaging 2001: PACS and Integrated Medical Information Systems: Design and Evaluation.
25. Ogawa E., Shimura K. (1999). Appearance matching of radiographic images using lightness index. Medical Imaging 1999: Image Perception and Performance.
26. Tyler C. W., Chan H., Liu L., McBride B., Kontsevich L. L. (1992). Bit stealing: how to get 1786 or more gray levels from an 8-bit color monitor. Human Vision, Visual Processing, and Digital Display III.
27. McEntee M. F., Toomey R. J. (2018). Display Optimization from a Perception Perspective. The Handbook of Medical Image Perception and Techniques, 452–469.
28. Jung H., Kim H.-J., Kang W.-S., Yoo S. K., Fujioka K., Hasegawa M., Samei E. (2004). Assessment of flat panel LCD primary class display performance based on AAPM TG 18 acceptance protocol. Medical Physics, 31(7), 2155–2164.
29. X-Rite i1Display Pro (vpixx.com)