Assessment of display performance for medical imaging systems: Executive summary of AAPM TG18 report

Chairman:

Ehsan Samei[a)]
Duke Advanced Imaging Laboratories, Departments of Radiology, Physics, and Biomedical Engineering, Duke University, DUMC 3302, Durham, North Carolina 27710

TG18 Members:

Aldo Badano
FDA, CDRH

Dev Chakraborty
University of Pennsylvania (currently with University of Pittsburgh)

Ken Compton
Clinton Electronics (currently with National Display Systems)

Craig Cornelius
Eastman Kodak Company (currently a consultant)

Kevin Corrigan
Loyola University

Michael J. Flynn
Henry Ford Health system

Bradley Hemminger
University of North Carolina, Chapel Hill

Nick Hangiandreou
Mayo Clinic, Rochester

Jeffrey Johnson
Sarnoff Corp, NIDL (currently with Siemens Corporate Research)

Donna M. Moxley-Stevens
University of Texas, Houston, M.D. Anderson Cancer Center

William Pavlicek
Mayo Clinic, Scottsdale

Hans Roehrig
University of Arizona

Lois Rutz
Gammex/RMI

Ehsan Samei
Duke University Medical Center

Jeffrey Shepard
University of Texas, Houston, M.D. Anderson Cancer Center

Robert A. Uzenoff
Fujifilm Medical Systems USA

Jihong Wang
University of Texas, Southwestern Medical Center (currently with University of Texas, Houston)

Charles E. Willis
Baylor University, Houston, Texas Children's Hospital (currently with University of Texas, Houston)

(Received 27 August 2004; revised 4 December 2004; accepted for publication 23 December 2004; published 31 March 2005)

Digital imaging provides an effective means to electronically acquire, archive, distribute, and view medical images. Medical imaging display stations are an integral part of these operations. Therefore, it is vitally important to assure that electronic display devices do not compromise image quality and ultimately patient care. The AAPM Task Group 18 (TG18) recently published guidelines and acceptance criteria for acceptance testing and quality control of medical display devices. This paper is an executive summary of the TG18 report. TG18 guidelines include visual, quantitative, and advanced testing methodologies for primary and secondary class display devices. The characteristics, tested in conjunction with specially designed test patterns (i.e., TG18 patterns), include reflection, geometric distortion, luminance, the spatial and angular dependencies of luminance, resolution, noise, glare, chromaticity, and display artifacts. Geometric distortions are evaluated by linear measurements of the TG18-QC test pattern, which should render distortion coefficients less than 2%/5% for primary/secondary displays, respectively. Reflection measurements include specular and diffuse reflection coefficients from which the maximum allowable ambient lighting is determined such that contrast degradation due to display reflection remains below a 20% limit and the level of ambient luminance (L_{amb}) does not unduly compromise luminance ratio (LR) and contrast at low luminance levels. Luminance evaluation relies on visual assessment of low contrast features in the TG18-CT and TG18-MP test patterns, or quantitative measurements at 18 distinct luminance levels of the TG18-LN test patterns. The major acceptable criteria for primary/secondary displays are maximum luminance of greater than 170/100 cd/m^2, LR of greater than 250/100, and contrast conformance to that of the grayscale standard display function (GSDF) of better than 10%/20%, respectively. The angular response is tested to ascertain the viewing cone within which contrast conformance to the GSDF is better than 30%/60% and LR is greater than 175/70 for primary/secondary displays, or alternatively, within which the on-axis contrast thresholds of the TG18-CT test pattern remain discernible. The evaluation of luminance spatial uniformity at two distinct luminance levels across the display faceplate using TG18-UNL test patterns should yield nonuniformity coefficients smaller than 30%. The resolution evaluation includes the visual scoring of the CX test target in the TG18-QC or TG18-CX test patterns, which should yield scores greater than 4/6 for primary/secondary displays. Noise evaluation includes visual evaluation of the contrast threshold in the TG18-AFC test pattern, which should yield a minimum of 3/2 targets visible for primary/secondary displays. The guidelines also include methodologies for more quantitative resolution and noise measurements based on MTF and NPS analyses. The display glare test, based on the visibility of the low-contrast targets of the TG18-GV test pattern or the measurement of the glare ratio (GR), is expected to yield scores greater than 3/1 and GRs greater than 400/150 for primary/secondary displays. Chromaticity, measured across a display faceplate or between two display devices, is expected to render a u',v' color separation of less than 0.01 for primary displays. The report offers further descriptions of prior standardization efforts, current display technologies, testing prerequisites, streamlined procedures and timelines, and TG18 test patterns. [DOI: 10.1118/1.1861159]

Key words: medical display, liquid crystal display, cathode ray tube, image quality, quality assurance, quality control, acceptance testing, picture archiving and communication system (PACS)

I. INTRODUCTION

The adoption of digital detectors and Picture Archiving and Communication Systems (PACS) has provided healthcare institutions an effective means to electronically archive and retrieve radiological images. Medical display workstations, an integral part of PACS, are used to display these images for diagnostic and clinical purposes. Considering the fundamental importance of image quality to the overall effectiveness of a diagnostic imaging practice, it is vitally important to assure that electronic display devices (also termed softcopy displays) do not compromise image quality as a number of studies have suggested.[1–3]

According to the American Association of Physicists in Medicine (AAPM) professional guidelines,[4] the performance assessment of electronic display devices falls within the professional responsibilities of medical physicists. While many prior publications have addressed some aspect of medical display performance,[5–15] prior evaluation and standardization efforts have fallen short of providing an unified approach for testing the performance of display devices such that the tests would take into consideration all the important aspects of display performance, be specific to medical displays, and be relatively easy to implement in a clinical setting.

AAPM Task Group 18 (TG18) recently completed a report which suggests standard guidelines and criteria for acceptance testing and quality control of medical display devices.[16] The intended audience of the report is practicing medical physicists, engineers, researchers, radiology administrative staff, manufacturers of medical displays, radiolo-

gists, and students interested in display quality evaluation. The report is developed such that while addressing the current dominant medical display technologies, cathode-ray tubes (CRTs) and liquid crystal displays (LCDs), many of the tests and concepts could be adapted to future display technologies.

The report is divided into six sections. Section one summarizes prior standardization efforts in the performance evaluation of medical display devices. Section two is a tutorial on the current and emerging medical display technologies. Section three sets forth prerequisites for the assessment of the display performance and includes a description of required instrumentation and TG18 test patterns. Section four is the main body of the report containing the description, quantification methods, and acceptance criteria for each key display characteristic. Sections five and six outline procedures for acceptance testing and quality control of display devices. The report further includes appendices providing guidelines for evaluating the performance of "closed" display systems, requirements for equivalent appearance of monochrome images, a full tabular description of TG18 test patterns, and a selected bibliography.

Considering the significant extent of the TG18 report, this paper aims to provide an executive summary of the report in a more condensed format. This paper focuses mainly on the testing procedures and criteria of the most direct relevance to acceptance testing and quality control procedures. The educational, advanced, and detailed descriptive portions of the report are not included. Interested individuals are referred to the full report for a complete description of the eliminated, summarized, and referenced sections.

II. GENERAL PREREQUISITES FOR DISPLAY ASSESSMENTS

II.A. Classification of Display Devices

In recognition of the currently accepted practice and in accordance with the guidelines set forth by the American College of Radiology[17] and the Food and Drug Administration, display devices for medical imaging are characterized in the TG18 report as either primary or secondary. Primary display systems are those used for the interpretation of medical images. They are typically used in radiology and in certain medical specialties such as orthopedics. Secondary systems are those used for viewing medical images by medical staff or specialists other than radiologists after an interpretive report is rendered. The operator's console monitors commonly used to "adjust" the images before they are sent for interpretation are treated as a primary display in terms of contrast response but secondary otherwise.

II.B. Required Tools

II.B.1. Instrumentation

Although many display tests can be performed visually, a more objective and quantitative evaluation of display performance requires special test tools. The required instruments vary in their complexity and cost depending on the context of the evaluation (research, acceptance testing, or quality control) and how thorough the evaluation needs to be. Table I summarizes the required instruments for display quality evaluation. The readers are advised to consult Sections 3.1 and 4-6 of the TG18 report to determine the subset of the tools and their performance requirements for the particular tests being performed.

II.B.2. Test Patterns

The TG18 report recommends the use of specific test patterns for performance evaluation of display devices in order to facilitate comparisons of measurements. The recommended patterns are designated with a nomenclature of the form TG18-*xyz*, where *x*, *y*, and *z* describe the type and derived variants of a pattern. The patterns are listed in Table II and a few examples are illustrated in Fig. 1. The full description of the patterns are in Sec. 3.2 and Appendix III of the TG18 report.

While the electronic copy of the TG18 report provides the patterns in multiple formats, they may also be generated with the aid of the information provided in the report. When displaying the patterns, no special processing functions should be applied. The 16-bit version of the patterns should be displayed with a window width and level set to cover the range from 0 to 4095 (window width, WW=4096, window level, WL=2048), except for the TG18-PQC, TG18-LN, and TG18-AFC patterns, where a WW of 4080 and WL of 2040 should be used. For 8-bit patterns, the displayed range should be from 0 to 255 (WW=256, WL=128). For some of the patterns, it is also essential to have a one-on-one relationship between the image pixels and the display pixels.

II.B.3. Software

Though not essential, software tools can facilitate the performance assessment of display devices. They include software for semiautomated generation of test patterns, processing software for assessment of resolution and noise, and spreadsheets for recording and manipulating the evaluation results. Some tools are provided along with the electronic copy of the TG18 report. Further information is available in Sec. 3.3 of the report.

II.C. Initial Steps for Display Assessment

II.C.1. Availability of Tools

Before starting the tests, the availability of the applicable tools and test patterns should be verified. Lists of desired tools for acceptance testing and quality control purposes are provided in the following section of this paper. The TG18 test patterns should be stored on the display workstation during installation, or otherwise be accessible from a network archive. This approach ensures that the same pattern will be utilized for all future testing.

TABLE I. Instrumentation used for display quality evaluation.

Instrument	Desired requirements	Purpose
Near-range luminance meter	• Calibration traceable to NIST • 0.05–1000 cd/m^2 luminance range • Better than 5% accuracy • Better than 10^{-2} (ideally 10^{-3}) precision • Aperture range $\leqslant$5 deg • Better than 3% compliance with the Commission Internationale de L'Eclairage (CIE) standard photopic spectral response	Luminance and luminance uniformity measurements
Telescopic luminance meter	• Those listed above for near-range meter • Acceptance angle $\leqslant$1 deg • Ability to focus to an area $\leqslant$6 mm	Luminance, luminance uniformity, reflection, angular response, and glare measurements
Illuminance meter	• Calibration traceable to NIST • 1–1000 lux illuminance range • Better than 5% accuracy • Better than 3% compliance with the CIE standard photopic spectral response • 180 deg cosine (Lambertian) response to better than 5% out to 50° angulation	Reflection and ambient lighting measurements
Colorimeter	• Calibration traceable to NIST • 1–1000 cd/m^2 luminance range • Better than 0.004 (u',v') accuracy	Chromaticity measurements
Digital camera	• Low noise and wide dynamic range • 1–500 cd/m^2 luminance range • $>512\times512$ matrix size • 10- to 12-bit depth • Equipped with a focusable macro lens • Variable frame rate/integration times up to 1 s • Digital interface to a computer • Calibrated for camera luminance, flat-field response, noise, and MTF • Equipped with a stable stand or tripod with directional adjustments	Quantitative resolution and noise measurements
Light source	• Uniform luminance $>$200 cd/m^2 • Small enough to subtend 15° from center of display	Quantitative specular reflection measurement
Illumination device	• See TG18 report Sec. 3.1.3	Quantitative diffuse reflection
Baffle	• Light absorbing characteristics • 5–15 mm opening	Glare and luminance measurements
Cone	• Light absorbing characteristics • 5 mm opening and $\leqslant$60 deg angular divergence	Glare and luminance measurements
Light absorbing cloth or hood	• Light absorbing characteristics	Display evaluation in the areas that have no control over the level of ambient lighting

TABLE I. (*Continued.*)

Instrument	Desired requirements	Purpose
Measuring microscope or magnifier	• Magnification ⩾25–50*x* • Equipped with a metric reticle with ⩽0.05 mm divisions • Focusing capabilities • Allow a working distance of ⩾12.5 mm	Visual resolution measurements
Flashlight	• None	Allow inspections in dark
Lint-free cleaning tissue glass-cleaning solution	• Recommended by the display manufacturer	Used for cleaning the faceplate, if needed
Two rulers and angle measurement device	• 1 m in length	Angular response and specular reflection coefficient measurements
Tape measure	• Flexible and 20–30 cm in length	Geometric distortions measurements

II.C.2. Display Placement

Prior to testing, the proper placement of a display device should be verified and adjustments made as appropriate. In the placement of a display device, the following should be considered:

1. Display devices should always be positioned to minimize specular reflection from direct light sources such as ceiling lights, film illuminators, or surgical lamps. The reflection of such light sources should not be observed on the faceplate of the display in the commonly used viewing orientations.
2. Many display devices, such as CRTs, are affected by magnetic fields; they should not be placed in an area with strong magnetic fields (i.e., vicinity of MRI scanners), unless properly shielded.
3. Displays should be placed ergonomically to avoid neck and back strain at reading level with the center of the display slightly below eye level.

II.C.3. Start-up Procedures

Prior to evaluation, the display device should be warmed up for approximately 30 min. In addition, the general system functionality should be verified by a quick review of the TG18-QC [Fig. 1(a)] test pattern. The pattern should be evaluated for distinct visibility of the 16 luminance steps, the continuity of the continuous luminance bars at the right and left of the pattern, the absence of gross artifacts, and the proper size and positioning of the active display area. Any adjustments to vertical and horizontal size must be made *prior* to performing the luminance measurements.

Dust and smudges on the face of the display will absorb, reflect, or refract emitted light possibly resulting in erroneous test results. In addition, newly installed displays are sometimes covered with a protective plastic layer, which upon removal can leave residual marks on the faceplate. Before testing a display device, the cleanliness of the faceplate should be verified. If the faceplate is not clean, it should be cleaned following the manufacturer's recommendations.

II.C.4. Ambient Lighting Level

The artifacts and loss of image quality associated with reflections from the display surface depend on the level of ambient lighting. As shown in Table III, illumination of display device surfaces in various locations of a medical facility may vary by over two orders of magnitude. The reflection measurement described in a later section of this document delineates a method to determine the maximum ambient light level appropriate for any given display device based on its reflection and luminance characteristics. It is important to verify that the ambient lighting in the room is below this maximum. The condition for the tests should be similar to those under normal use of the equipment. By recording ambient light levels at a reference point at the center of the faceplate and noting the location and orientation of the display devices at acceptance testing, it will be possible to optimize repeatability of testing conditions in the future. If a display device is equipped with a photocell for ambient light detection, its use should be in compliance with the Digital Imaging and Communication in Medicine (DICOM) grayscale standard display function (GSDF) as further discussed below.

II.C.5. Pretest Luminance Settings

Before the performance of a display system can be assessed, proper display area size should be established, and the maximum luminance L_{max} and the minimum luminance L_{min} must be checked to verify that the device is properly configured. The desired values should be determined based on the desired luminance ratio, the reflection characteristics of the system, and the ambient lighting level (see the reflection and luminance sections below). Using a luminance meter, the luminance values should be recorded using the TG18-LN8-01 (or TG18-LN12-01) test pattern for L_{min} and

TABLE II. Test patterns recommended for display quality evaluation. The patterns are divided into six sets. Most patterns are available in 1024×1024 (1 k) size and in either DICOM or tiff format. Some patterns are available in 2048×2048 (2 k) size.

Set	Series	Type	Images	Description
Multipurpose (1 k and 2 k)	TG18-QC	Vis./Qnt	1	Resolution, luminance, distortion, artifacts
	TG18-BR	Visual	1	Briggs pattern, low contrast detail vs luminance
	TG18-PQC	Vis./Qnt.	1	Resolution, luminance, contrast transfer for prints
Luminance (1 k only)	TG18-CT	Visual	1	Luminance response
	TG18-LN	Quant.	18	DICOM grayscale calibration series
	TG18-UN	Visual	2	Luminance and color uniformity, and angular response
	TG18-UNL	Quant.	2	Same as above with defining lines
	TG18-AD	Visual	1	Contrast threshold at low luminance for evaluating display reflection
	TG18-MP	Visual	1	Luminance response (bit depth resolution)
Resolution (1 k and 2 k)	TG18-RH	Quant.	3	Five horizontal lines at three luminance levels for LSF evaluation
	TG18-RV	Quant.	3	Five vertical lines at three luminance levels for LSF evaluation
	TG18-PX	Quant.	1	Array of single pixels for spot size
	TG18-CX	Visual	1	Array of Cx patterns and a scoring reference for resolution uniformity
	TG18-LPH	Visual	3	Horizontal bars at 1 pixel width, 1/16 modulation, three luminance levels
	TG18-LPV	Visual	3	Vertical bars at 1 pixel width, 1/16 modulation, three luminance levels
Noise (1 k only)	TG18-AFC	Visual	1	4AFC contrast-detail pattern, four CD values
	TG18-NS	Quant.	3	Similar to RV/RH, five uniform regions for noise evaluation
Glare (1 k only)	TG18-GV	Visual	2	Dark spot pattern with low contrast object
	TG18-GQ	Quant.	3	Dark spot pattern for glare ratio measurement
	TG18-GA	Quant.	8	Variable size dark spot patterns
Anatomical (2 k only)	TG18-CH	Visual	1	Reference anatomical PA chest pattern
	TG18-KN	Visual	1	Reference anatomical knee pattern
	TG18-MM	Visual	2	Reference anatomical mammogram pattern

TG18-LN8-18 (or TG18-LN12-18) for L_{max}, respectively. For these measurements, ambient illumination should be reduced to negligible levels using a dark cloth shroud if necessary. If the measured values for L_{max} and L_{min} are not appropriate, the proper values should be established using the brightness and contrast controls of the display. Otherwise, the display device should be serviced before testing its performance. The TG18 report further recommends compliance of medical display systems with the DICOM GSDF.[15] Before initiating the testing procedures, the device should be calibrated or otherwise its calibration verified within its operating luminance range defined by L_{max} and L_{min}.

II.C.6. Personnel

The acceptance and quality control (QC) testing of a display system must be performed by an individual(s) having appropriate technical and clinical competencies. Even though the vendor is expected to perform some testing before turning a display system over to the user, the user must independently test the system(s). For acceptance testing and annual QC evaluation, the tests should be performed by a medical physicists trained in display performance assessments. Other staff including biomedical engineers, in-house service electronic technicians, or trained x-ray technologists can perform some of the tests described herein; however, in such situations, a qualified medical physicist should accept full oversight responsibilities and final approval of the results. For monthly or quarterly QC, the tests can be delegated to such qualified professionals as well as long as they work under the direct supervision of the medical physicist. The daily QC of a display system should be performed by the operator/user of the system. Radiology staff using electronic displays should

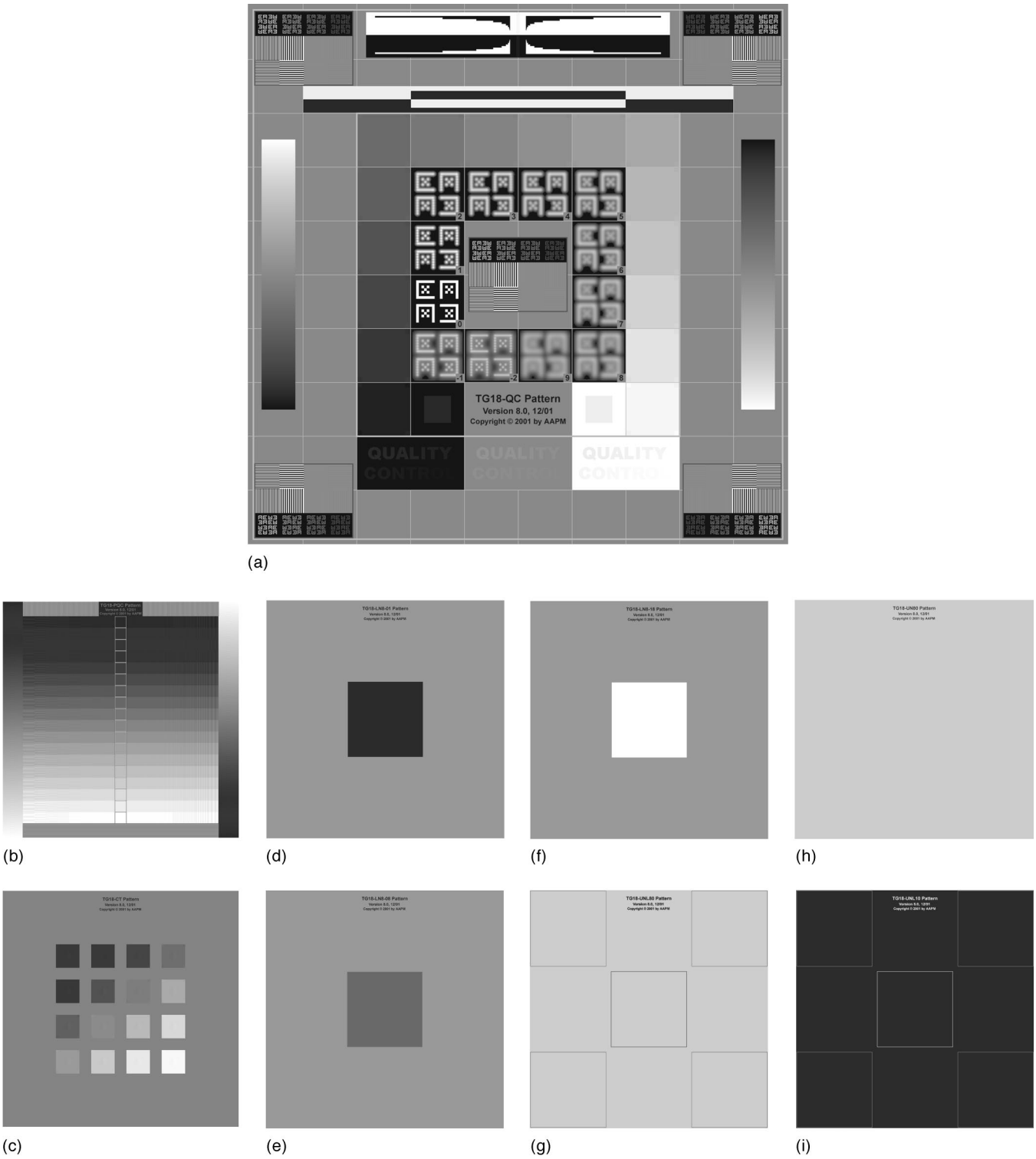


Fig. 1. Examples of TG18 test patterns: TG18-QC (a), TG18-PQC (b), TG18-CT (c), TG18-LN8-01 (d), TG18-LN8-08 (e), TG18-LN8-18 (f), TG18-UNL80 (g), TG18-UN80 (h), TG18-UNL10 (i), TG18-MP (j), TG18-RV89 (k), TG18-RH50 (l), TG18-CX (m), TG18-AFC (n), TG18-GV (o), TG18-GA30 (p), TG18-GQB (q), TG18-CH (r), TG18-KN (s), TG18-MM1 (t), and TG18-MM2 (u).

be familiar with the daily testing procedure and expected results. All personnel responsible for performing QC tests will require initial training specific to their level of responsibility and periodic retraining and mentoring by medical physics staff.

II.C.7. Specific Prerequisites for Acceptance Testing

Acceptance testing requires close communication with the vendor for understanding and documenting the operational

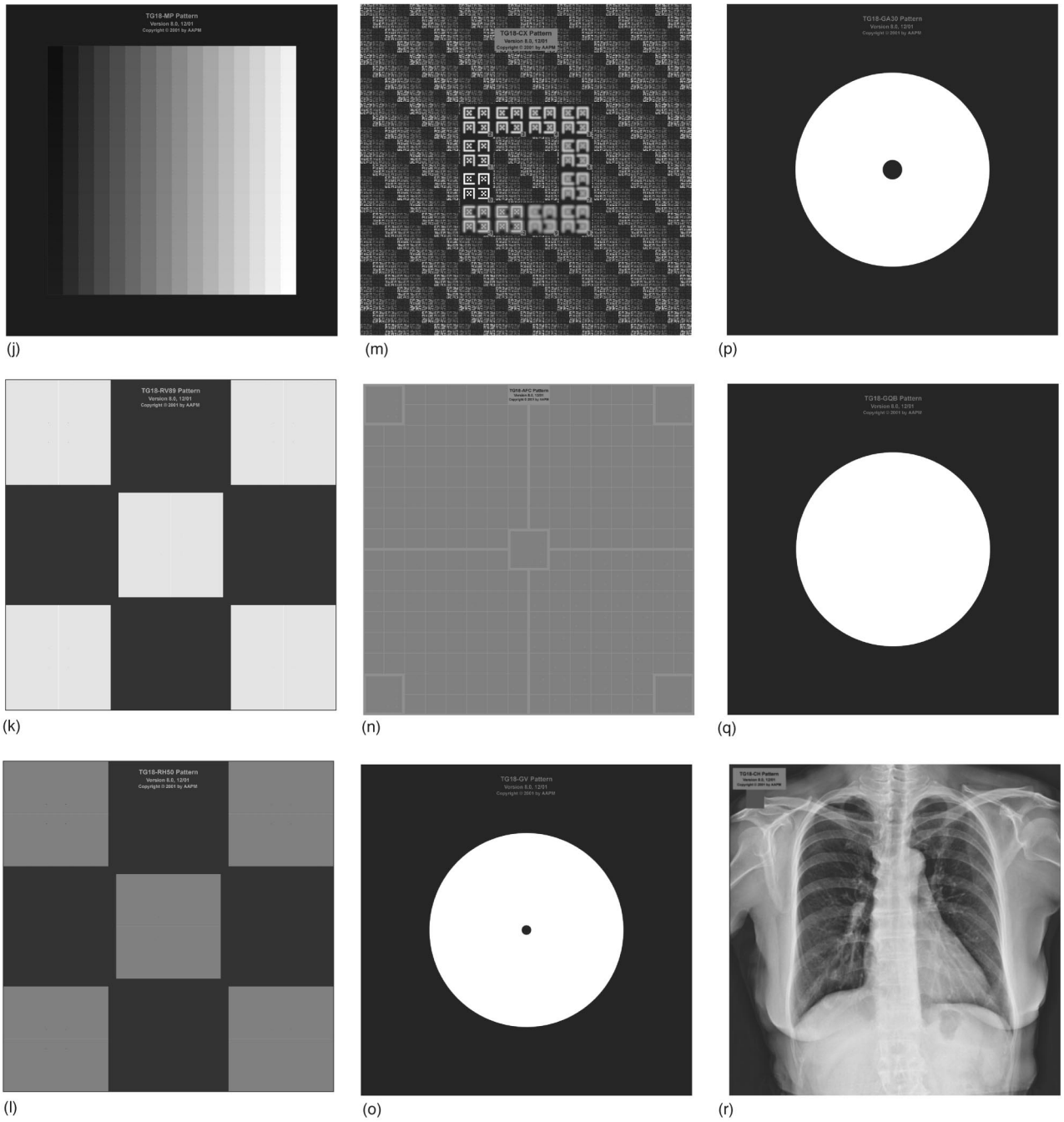
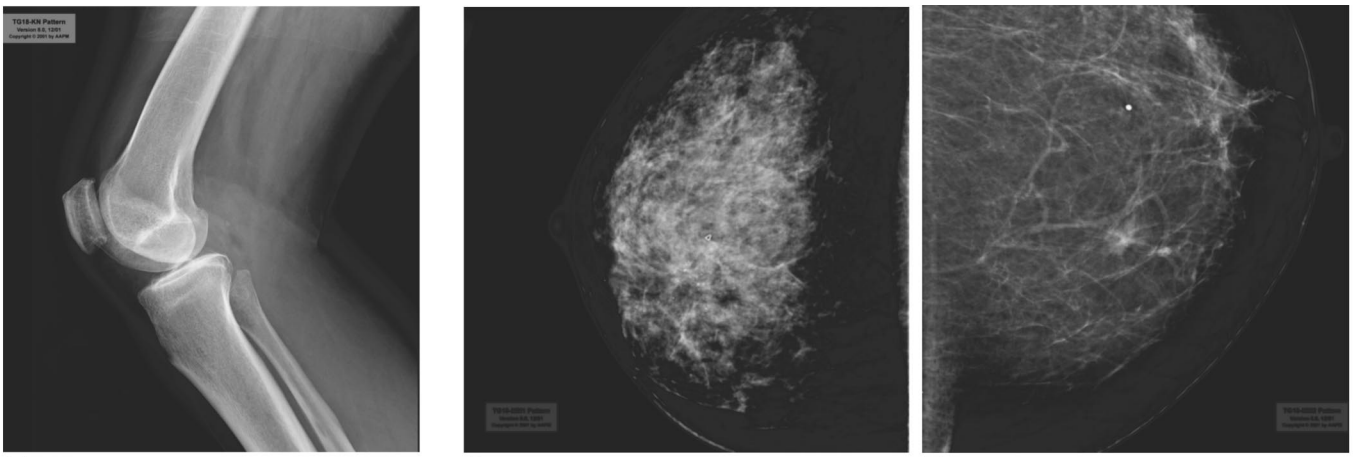


Fig. 1. (*Continued*).

features and dedicated QC utilities of the system. Any recommended service and/or calibration schedule, including the services provided, tests performed, and the service/calibration intervals, must be obtained from the manufacturer, ideally as part of the purchasing process. Prior to acceptance testing, the characteristics of the display systems delivered should be verified against those specified in the purchase agreement. A database should be established which includes information such as display type, size, resolution, manufacturer, model, serial number, manufacture date, room number, display identification (if applicable), associated display hardware (e.g., display controller) and test patterns available on the systems. All delivered documentation from the vendor should also be reviewed with special attention to the testing results performed at the factory.

II.C.8. Specific Prerequisites for Quality Control

The initial acceptance testing data are used to establish and maintain expected performance. Data acquired during routine QC testing must be compared to the limits established around the baseline values. It is also essential to utilize the same pattern for repeat evaluations of a given display device. The use of worksheets and checklists will help in



(s) (t) (u)

Fig. 1. (*Continued*).

establishing and monitoring the baselines. It is strongly recommended to record and maintain this information in electronic databases. Most commercial calibration packages support automated recording, tracking, and analysis of display QC results.

III. ASSESSMENT OF DISPLAY PERFORMANCE

The performance assessment of a display device in a clinical setting might be performed in the context of acceptance testing, prior to first clinical use, or quality control, throughout the life of the device. Tables IV and V provide a list of the tests, the required tools, and the expected performance for the two types of procedures with specific reference to the TG18 report. Depending on the interest and resources, additional advanced tests are further encouraged. For QC tests, hardware features and reproducible performance can reduce the need for very frequent testing. However, it is recommended that initially the tests be performed more frequently. If stability is maintained, a determination can be made to decrease the frequency of testing.

The sections below provide the assessment methodologies. It is generally ideal to perform the tests in the order in which they are discussed as some of the latter tests may be influenced by parameters that are addressed in earlier tests. Full descriptions of the specific display characteristics as well as advanced testing procedures are provided in the TG18 report,[16] to which the interested readers are referred.

III.A. Geometric Distortions

Geometric distortions of displayed images are often a concern in cathode-ray tube (CRT) display devices. The distortions can be in concave, convex, skewed, or other nonlinear forms. The magnitude and type of such distortions should be evaluated and, if deemed inappropriate, adjusted to meet certain minimum requirements as noted below.

III.A.1. Visual Evaluation

The geometric distortion of a display system is ascertained visually using the TG18-QC or the TG18-LPV/LPH test pattern. The patterns should be maximized to fill the entire usable display area. For displays with rectangular display areas, the patterns should cover at least the narrower dimension of the display area and be placed at the center of the area used for image viewing. The pattern(s) should be examined from a viewing distance of 30 cm.

The patterns should appear straight without significant geometric distortions, and should be properly scaled to the aspect ratio of the video source pixel format so that the grid of the TG18-QC pattern appears square. The lines should appear straight indicative of proper linearity without any curvature or waviness. Some small barrel and pincushion distortions are normal for CRT devices but should not be excessive. For the TG18-LPV and TG18-LPH patterns, in addition to straightness, the lines should appear equally spaced.

III.A.2. Quantitative Evaluation

Spatial accuracy for geometric distortions can be quantified using the TG18-QC test pattern, maximized to fill the entire display area. Using a straight edge as a guide for a best fit and with the aid of a flexible plastic ruler, distances should

TABLE III. Typical ambient lighting levels.

Area	Illumination (lux)
Operating rooms	300–400
Emergency medicine	150–300
Hospital clinical viewing stations	200–250
Staff offices	50–180
Diagnostic reading stations (CT/MR/NM)	15–60
Diagnostic reading stations (x-rays)	2–10

TABLE IV. Tests, tools, and acceptance criteria for acceptance testing and annual quality control of electronic display systems. The section notations refer to the TG18 report.

Test	Major required tools		Procedure	Acceptance criteria (for two classes of displays)		Suggested action (if unacceptable)
	Equipment	Patterns		Primary	Secondary	
Geometric distortions	Flexible ruler or transparent template	TG18-QC	See Sec. 4.1.4	Deviation $\leqslant 2\%$	Deviation $\leqslant 5\%$	Readjustment, repair, or replacement for repeated failures
Reflection[a]	Measuring ruler, light sources, luminance and illuminance meters, illuminator	TG18-AD	See Secs. 4.2.3 and 4.2.4	$L_{\min} \geqslant 1.5 L_{\mathrm{amb}}$ (ideally $\geqslant 4 L_{\mathrm{amb}}$)	$L_{\min} \geqslant 1.5 L_{\mathrm{amb}}$ (ideally $\geqslant 4 L_{\mathrm{amb}}$)	Results are used to adjust the level of ambient lighting
Luminance response	Luminance and illuminance meters	TG18-LN TG18-CT TG18-MP	See Secs. 4.3.4 and 4.3.3	$L_{\max} \geqslant 170$ cd/m^2 LR $\geqslant 250$ $\kappa_\delta \leqslant 10\%$ $\Delta L_{\max} \leqslant 10\%$	$L_{\max} \geqslant 100$ cd/m^2 LR $\geqslant 100$ $\Delta L_{\max} \leqslant 10\%$ $\kappa_\delta \leqslant 20\%$	Readjustment, recalibration, repair, or replacement for repeated failures
Luminance dependencies[b]	Luminance meter, luminance angular response measurement tool	TG18-UNL TG18-LN TG18-CT	See Secs. 4.4.3 and 4.4.4	Nonunif. $\leqslant 30\%$ $\mathrm{LR}'_{\delta\theta} \geqslant 175$ $\kappa_{\delta\theta} \leqslant 30\%$	Nonunif. $\leqslant 30\%$ $\mathrm{LR}'_{\delta\theta} \geqslant 70$ $\kappa_{\delta\theta} \leqslant 60\%$	Readjustment, repair or replacement for repeated failures; Angular results used to define acceptable viewing angle cone
Resolution[c]	Luminance meter magnifier	TG18-QC TG18-CX TG18-PX	See Secs. 4.5.3 and 4.5.4.1.2	$0 \leqslant \mathrm{Cx} \leqslant 4$ RAR = 0.9 − 1.1 AR $\leqslant 15$	$0 \leqslant \mathrm{Cx} \leqslant 6$	Focus adjustment, repair, or replacement for repeated failures
Noise[c]	None	TG18-AFC	See Sec. 4.6.3	All targets visible except the smallest	Two largest sizes visible	Reverification of luminance response, otherwise replacement
Veiling glare	Baffled funnel, telescopic photometer	TG18-GV TG18-GVN TG18-GQs	See Secs. 4.7.3 and 4.7.4	$\geqslant 3$ targets visible, GR $\geqslant 400$	$\geqslant 1$ target visible, GR $\geqslant 150$	Reverification of luminance response, otherwise replacement
Chromaticity	Colorimeter	TG18-UNL80	See Sec. 4.8.4	$\Delta(u', v') \leqslant 0.01$	None	Replacement

Note: Acronyms: L_{amb}= ambient luminance, $L_{\min}$= minimum luminance, $L_{\max}$= maximum luminance, LR=luminance ratio, κ_δ= maximum deviation between measured and GSDF contrast, Cx= Cx score, RAR= resolution-addressability ratio, AR= astigmatism ratio, GR= glare ratio.
[a]In the absence of illumination devices, this acceptance testing can be performed only visually using TG18-AD and the method described in Sec. 4.2.3.1.
[b]Angular tests are not required as a part of annual quality control.
[c]More objective resolution and noise measurements can be performed as described in Secs. 4.5.4 and 4.6.4 using a digital camera.

be measured in square areas in the horizontal and vertical directions in each of the four quadrants of the pattern and within the whole pattern (Fig. 2). It is important to assure the locations of the cross hatches be viewed perpendicular to the display's faceplate. In each quadrant, between quadrants, and within the whole pattern, the maximum percent deviations between the measurements in each direction and between the measurements in the horizontal and vertical directions should be determined. The percentages should be calculated in relation to the smallest of the values being compared.

The measured spatial deviations shall be less than 2% and 5% for primary or secondary displays, respectively. If a display device does not meet these criteria, adjustments should be made to the distortion control of the device. Often, as the area of the display is increased or decreased, the luminance will also increase or decrease in a nonlinear fashion. Therefore, it is important to make and finalize such adjustments prior to testing and adjustment of the display luminance characteristics. In addition, if a display workstation contains more than one display device, it is important to have the vertical and horizontal sizes of the active areas carefully matched within 2%. This facilitates the subsequent matching of their luminance response characteristics.

TABLE V. (a) Tests for daily quality control of electronic display system, performed by the display user. (b) Tests for monthly/quarterly quality control of electronic display systems performed by a medical physicist or by a QC technologist under the supervision of a medical physicist. The section notations refer to the TG18 report. For acronyms see Table IV.

Test	Major required tools Equipment	Patterns	Procedure	Acceptance criteria (for two classes of displays) Primary	Secondary	Suggested action (if unacceptable)
			(a)			
Overall visual assessment	None	TG18-QC or anat. images	See Secs. 4.10.1 or 4.10.6	See Secs. 4.10.1/4.10.6	See Secs. 4.10.1/4.10.6	Further /closer evaluation
			(b)			
Geometric distortions	None	TG18-QC	See Sec. 4.1.3.1	See 4.1.3.2	See 4.1.3.2	Further/closer evaluation
Reflection	Luminance and illuminance meters	TG18-AD	See Secs. 4.2.3 and 4.2.4	$L_{min} \geqslant 1.5L_{amb}$ (ideally $\geqslant 4L_{amb}$)	$L_{min} \geqslant 1.5L_{amb}$ (ideally $\geqslant 4L_{amb}$)	Readjust the level of ambient lighting
Luminance response	Luminance and illuminance meters	TG18-LN TG18-CT TG18-MP	See Secs. 4.3.4 and 4.3.3	$L_{max} \geqslant 170$ cd/m^2 LR$\geqslant 250$ $\Delta L_{max} \leqslant 10\%$ $\kappa_\delta \leqslant 10\%$	$L_{max} \geqslant 100$ cd/m^2 LR$\geqslant 100$ $\Delta L_{max} \leqslant 10\%$ $\kappa_\delta \leqslant 20\%$	Readjustment, recalibration, repair, or replacement for repeated failures
Luminance dependencies	Luminance meter	TG18-UN TG18-UNL	See Secs. 4.4.3 and 4.4.4	Nonunif. $\leqslant 30\%$	Nonunif. $\leqslant 30\%$	Readjustment, repair, or replacement for repeated failures
Resolution	Magnifier	TG18-QC TG18-CX	See Sec. 4.5.3	$0 \leqslant Cx \leqslant 4$	$0 \leqslant Cx \leqslant 6$	Focus adjustment, repair, or replacement for repeated failures

III.B. Display Reflection

Electronic display devices have specular and diffuse reflection that can reduce image contrast and affect image quality. Ambient light reflections are more pronounced in display devices with thick faceplates (e.g., CRTs) compared to those with thinner faceplates (e.g., LCDs). They are generally reduced by the application of antireflective (AR) coating on the faceplate and/or the addition of light absorbers within the faceplate of the display, but these means do not completely eliminate reflections. The reflection characteristics of a medical display device should be evaluated in order to establish the maximum allowable level of ambient lighting at which the device can be operated without overly compromising the desired luminance performance and contrast threshold.

III.B.1. Visual Evaluation

III.B.1.a. Specular Reflection Characteristics An effective and simple visual test for specular reflection of a display device is to observe the device in the power-save mode or turned off. The ambient lighting in the room should be maintained at levels normally used. The display's faceplate should be examined at a distance of about 30–60 cm within an angular view of ±15 deg for the presence of specularly reflected light sources or illuminated objects. Patterns of high contrast on the viewer's clothing are common sources of reflected features. In general, no specularly reflected patterns of high contrast objects should be seen. If light sources such as that from a film illuminator or window are seen, the position of the display device in the room is not appropriate. If high contrast patterns such as an identification badge on a white shirt or a picture frame on a light wall are seen, the ambient illumination in the room should be reduced.

III.B.1.b. Diffuse Reflection Characteristics The effect of diffusely reflected light on image contrast may be observed

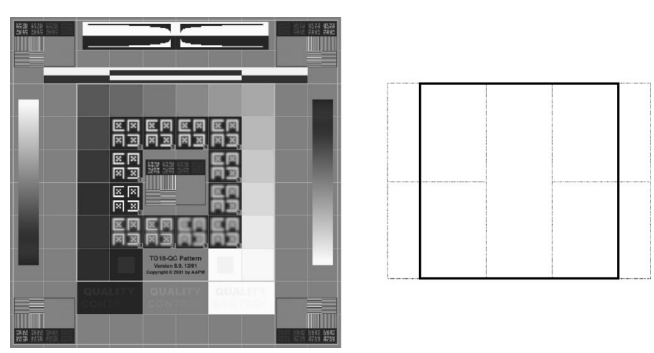


FIG. 2. Spatial measurements for the quantitative evaluation of geometric distortions using the TG18-QC test pattern. The small squares with dashed lines (- - -) define the four quadrants of the pattern, and the large square at the center encompassing the luminance patches is the one to be used for geometric distortion characterization within the whole image.

by alternately viewing the low-contrast patterns in the TG18-AD test pattern in near total darkness and in normal ambient lighting, determining the threshold of visibility in each case. A dark cloth placed over both the display device and the viewer may be helpful for establishing near total darkness. The pattern should be examined from a viewing distance of 30 cm. The threshold of visibility should not be different when viewed in total darkness and when viewed in ambient lighting conditions. If the ambient lighting renders the "dark-threshold" not observable, the ambient illuminance on the display surface is causing excess contrast reduction, and the room ambient lighting needs to be reduced.

III.B.2. Quantitative Evaluation

III.B.2.a. Specular Reflection Characteristics The specular reflection coefficient for a display device can be measured with a small-diameter source of diffuse white light as described in Sec. 3.1.3 of the TG18 report. The display should be in the power-save mode or turned off. The light source, subtending 15° from the center of the display, should be positioned d_1 centimeters from the center of the display and be pointed toward the center at an angle of 15° from the surface normal. The reflected luminance of the light source should then be measured with a telescopic photometer from a distance of d_2 centimeters from the center of the display and similarly angled at 15° to the normal. Finally, the directly viewed luminance of the light source should be measured with the same photometer from a distance of d_1+d_2 centimeters. The specular reflection coefficient R_s is the ratio of the reflected spot luminance to the directly viewed spot luminance. All measurements should be made in a dark room.

As the artifacts associated with specular reflections depend on the ambient lighting, the measured specular reflection coefficient should be used to establish the maximum allowable ambient lighting E as

$$E \leqslant (\pi C_t L_{\min})/(0.9\, R_s), \tag{1}$$

where the contrast threshold $C_t = \Delta L/L$ (see Fig. 3 and Sec. 4.3.1 of TG18 report), corresponds to its value at the minimum luminance $L_{\min}$. For convenience, this relationship is tabulated in Table VI so that the maximum room lighting can be identified if R_s and $L_{\min}$ are known. As an example, for a typical CRT with antireflective (AR) coating ($R_s = 0.004$) operated at minimum luminance values of 0.5, 1, 1.5, and 2.0 cd/m^2, the ambient lighting based on specular reflection consideration should be less than approximately 14, 21, 28, and 31 lux, respectively. Note that in the adjustment and measurement of the appropriate level of ambient lighting, illuminance in the room should be measured with the illuminance meter placed at the center of the display and facing outward, so the proper amount of light incident on the faceplate can be assessed.

III.B.2.b. Diffuse Reflection Characteristics The luminance from diffuse reflections adds to that produced by the display device. The ambient illumination produces a luminance of $L_{\text{amb}} = R_d E$, where E is ambient illuminance on the display surface, and R_d is the diffuse reflection coefficient in units of cd/m^2 per lux or 1/sr. In the dark areas of a low-contrast image, the change in luminance ΔL_t will produce a relative contrast of $\Delta L_t/(L_{\min}+L_{\text{amb}})$. For some devices, the luminance response can be calibrated to account for the presence of a known amount of luminance from ambient lighting L_{amb} and produce equivalent contrast transfer in both dark and bright regions. However, if L_{amb} is sufficiently large in relation to $L_{\min}$, even if the device has a high contrast ratio, the overall luminance ratio of the device is compromised.

The diffuse reflection coefficient may be measured using standardized illumination of the display surface with the illuminator device described in Sec. 3.1.3 of the TG18 report (Fig. 4). The illuminance should then be measured in the center of the display device using a probe placed on the center of the display surface. The sensitive area of the meter should be held vertically to measure the illuminance incident on the display faceplate. The induced luminance at the center of the display surface should then be measured with a telescopic luminance meter as illustrated in Fig. 4. The luminance measurement should be made through the small aperture at the back of the containment device so as to not perturb the reflective characteristics of the containment struc-

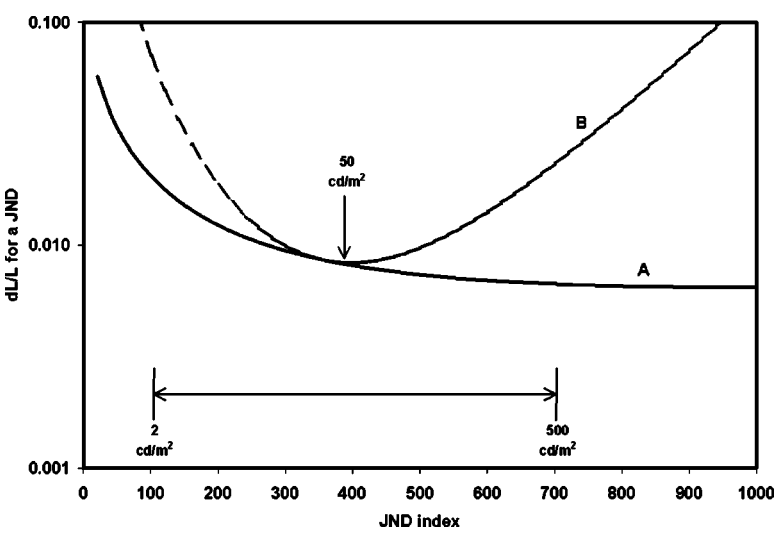


FIG. 3. Contrast threshold for varied visual adaptation (a) and fixed (b) visual adaptation (Ref. 19). The contrast threshold dL/L for a just noticeable difference (JND) depends on whether the observer has fixed (b) or varied (a) adaptation to the light and dark regions of an overall scene. dL/L is the peak-to-peak modulation of a small sinusoidal test pattern.

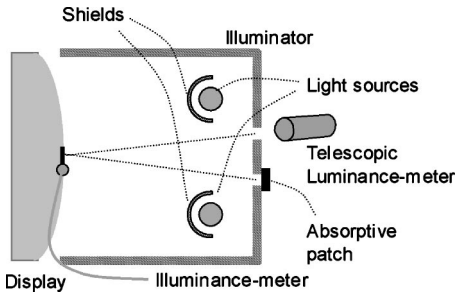


FIG. 4. Typical illuminating device used for the measurement of the diffuse reflection coefficient of a display device.

ture. The viewing aperture must be located from 8° to 12° off to the side from the normal so as to not interfere with the measurement result. The diffuse reflection coefficient R_d is computed as the ratio of the luminance to the illuminance in units of sr^{-1}.

As diffuse reflection reduces the contrast, for primary class display devices, the level of ambient illuminance should be set to insure that the contrast in dark regions observed with ambient illumination will be at least 80% of the contrast observed in near total darkness. This requirement translates to $L_{amb} < 0.25\ L_{min}$, or

$$E \leqslant (0.25\ L_{min})/R_d. \tag{2}$$

Table VII identifies the ambient lighting for which L_{amb} is 0.25 of L_{min} as a function of R_d and L_{min}. As an example, for a typical CRT with AR coating ($R_d = 0.02\ sr^{-1}$) operated at minimum luminance values of 0.5, 1, 1.5, and 2.0 cd/m^2, the ambient lighting based on diffuse reflection consideration should be less than approximately 7, 12, 19, and 25 lux, respectively. In situations where the level of ambient lighting can be strictly controlled and taken into account in the luminance calibration of the display device, a larger L_{amb} can be tolerated ($L_{amb} < L_{min}/1.5$) as noted in the next section.

III.C. Luminance Response

The human visual system perceives brightness in a nonlinear fashion.[18] Ideally, the luminance response of a display device should match this nonlinear response such that image values are displayed in equally perceptible luminance increments. While limited due to variations in the contrast sensitivity of the human eye in scenes with wide ranges of luminance levels (e.g., medical images),[19,20] the DICOM grayscale standard display function (GSDF)[15] offers a way to approach this goal by applying a specific look-up-table to the display values, such that the display values present equally discriminable levels of brightness.

The intrinsic luminance response (i.e., luminance versus display value) of most display devices is markedly different from the GSDF. It usually follows a power-law relationship for CRTs, and a linear one for LCDs.[5] In addition, the luminance response may vary over time. In CRTs, for example, the phosphor efficiency decreases as the device ages. Modern display devices also have utilities that automatically calibrate the luminance response of the device to GSDF. However, the functionality and accuracy of these utilities should be independently verified by the user.

III.C.1. Visual Evaluation

The luminance response of a display device is visually inspected using the TG18-CT test pattern. The pattern should be evaluated from a viewing distance of 30 cm for visibility of the central half-moon targets and the four low-contrast objects at the corners of each of the 16 different luminance regions. Since this pattern is viewed in one state of visual adaptation, it is expected that the contrast transfer will be better at the overall brightness for which the visual system is adapted as opposed to the darkest or the brightest regions. With experience, the visual characteristics of this test pattern can be recognized for a system with quantitatively correct luminance response. In general, the low contrast targets should be visible in all regions. A common failure is not to be able to see the targets in one or two of the dark regions.

The bit-depth resolution of the display should be evaluated using the TG18-MP test pattern. The evaluation includes ascertaining the horizontal contouring bands, their relative locations, and grayscale reversals. The pattern should be examined from a viewing distance of 30 cm. In general, the relative location of contouring bands and any luminance levels should not be farther than the distance between the 8-bit markers (long markers). No contrast reversal should be discernible.

TABLE VI. Maximum allowable ambient illuminance based on specular reflection: For a display device with a specific minimum luminance L_{min} and a specific specular reflection coefficient R_s the ambient illumination which maintains specular reflections from high contrast objects below the visual contrast threshold (C_t) is tabulated.

		Maximum room illuminance (lux)				
L_{max}–L_{min} (cd/m^2)	C_t	$R_s = 0.002$	$R_s = 0.004$	$R_s = 0.008$	$R_s = 0.020$	$R_s = 0.040$
5000–20	0.010	349	175	87	35	17
2500–10	0.011	192	96	48	19	10
1000–4	0.015	105	52	26	10	5
500–2	0.018	63	31	16	6	3
250–1	0.024	42	21	10	4	2

III.C.2. Quantitative Evaluation

In the quantitative method, luminance $L(p)$ is measured using a calibrated luminance meter at the center of the 18 TG18-LN test patterns, corresponding to 18 distinct digital driving levels p. The measurement of $L(p)$ using patterns other than the TG18-LN patterns may result in different values due to the influence of veiling glare. The effect of ambient illumination should be reduced to negligible levels, by using a dark cloth if necessary. If a telescopic luminance meter is used, in order to minimize the influence of meter's flare on the low-luminance measurements, the measurements may need to be made through a cone or baffle to shield the instrument from the surrounding light. For display devices with non-Lambertian light distribution, such as a LCD, if the measurements are made with a near range luminance meter, the meter should either have an aperture angle smaller than 5 deg or display-specific correction factors should be applied.[21]

The ambient luminance on the display faceplate (L_{amb}) should either be estimated from the measured R_d values as $L_{\mathrm{amb}}=ER_d$ or measured directly. In the case of direct measurement, the display device should be put in the power-save or blank screen-save mode (otherwise turned off). A telescopic luminance meter normal to the display surface is used with a light-absorbing mask placed behind the meter to minimize specular reflection from the display. Otherwise the room lighting is set to the conditions established for the normal use of the equipment (see Sec. III B above). The values for $L'(p)$ including L'_{max} and L'_{max} are then computed by the addition of L_{amb} to the measured $L(p)$ values.

The recommended value for L'_{max} is typically specified by the vendor as the highest value that can be used without compromising other performance characteristics, such as lifetime or resolution. L'_{max} should be greater than 171 cd/m^2 for primary displays[17] and 100 cd/m^2 for secondary displays, and should be within 10% of the desired value for both classes of display. Furthermore, for workstations with multiple monitors, L'_{max} should not differ by more than 10% among monitors. L'_{max} should be such that the desired luminance ratio $LR'=L'_{\mathrm{max}}/L'_{\mathrm{min}}$ is obtained. If the manufacturer's recommendations are not available, it is recommended that the luminance ratio of a display device be set equal to or greater than 250 for all primary class devices.[19] For secondary class devices, LR$'$ should be no less than 100. In general, L'_{min} should be within 10% of the nominally desired values for both classes of display.

As ambient lighting can impact the low luminance response of a display device and reduce the device's effective luminance ratio, a limit on L_{amb} is further indicated. For both classes of display devices, L_{amb} should ideally be less than $0.25L_{\mathrm{min}}$ (or $0.2L'_{\mathrm{min}}$). In situations where the level of ambient lighting can be strictly controlled and taken into account in the luminance calibration of the display device, a larger L_{amb} can be tolerated, but L_{amb} should always be less than $L_{\mathrm{min}}/1.5$ (or $L'_{\mathrm{min}}/2.5$). If necessary, arrangements should be made to reduce the room lighting in order to achieve a sufficiently small L_{amb}.

In evaluating the luminance response of the display between the maximum and minimum extremes, the measured luminance values should be related to the DICOM GSDF luminance response in terms of the contrast response, i.e., the slope of the measured luminance response. To do so, using the DICOM's table of just noticeable difference (JND) indices versus luminance, the JND indices for the measured L'_{min} and L'_{max} should first be identified. The JND indices for the intermediate L' values should then be evenly spaced within the JND range and linearly related to the actual p values used as

$$J_i = J_{\mathrm{min}} + \frac{P_i(J_{\mathrm{max}} - J_{\mathrm{min}})}{\Delta P}, \tag{3}$$

where J indicates the JND indices (e.g., Fig. 5). The measured data are then expressed as the observed contrast, δ_i, at each luminance step L'_i, as a function of mean JND index value associated with that step

$$\delta_i = \frac{2(L'_i - L'_{i-1})}{(L'_i + L'_{i-1})(J_i - J_{i-1})} @ 0.5(J_i + J_{i-1}). \tag{4}$$

The expected response from DICOM GSDF luminance values δ_i^d is also similarly computed using the following equation:

TABLE VII. Maximum room lighting based on diffuse reflection: For a display device with a specific minimum luminance L_{min} and a specific diffuse reflection coefficient R_d in units of cd/m^2 per lux or 1/sr, the ambient illumination which maintains 80% contrast in dark regions is tabulated. The maximum room illuminance is calculated as $0.25L_{\mathrm{min}}/R_d$.

	Maximum room illuminance (lux)				
L_{max}–L_{min} (cd/m^2)	$R_s=0.005$	$R_s=0.010$	$R_s=0.020$	$R_s=0.040$	$R_s=0.060$
5000–20	1000	500	250	125	83
2500–10	500	250	125	62	42
1000–4	200	100	50	25	17
500–2	100	50	25	12	8
250–1	50	25	12	6	4

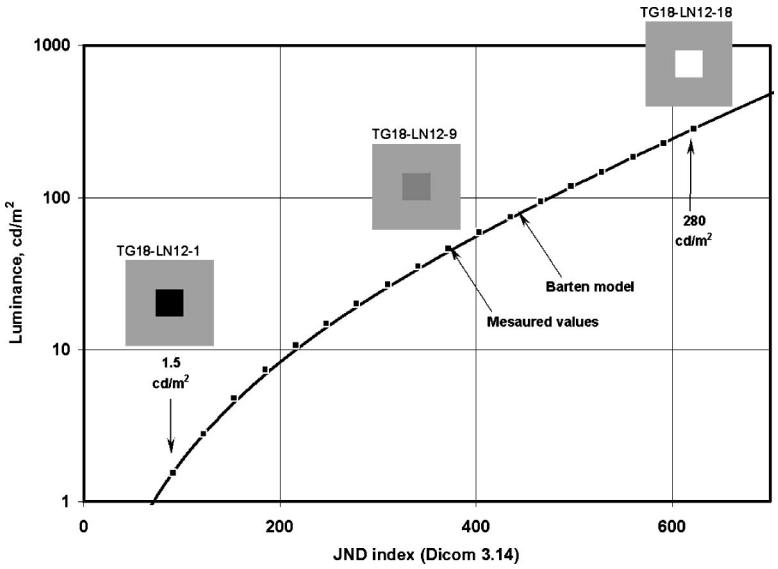


FIG. 5. Example of the measured luminance for 18 display levels is plotted in relation to the DICOM GSDF. The *p*-values used to measure luminance have been linearly scaled to JND indices with the values at L'_{max} and L'_{min} set to be equal to the JND corresponding indices.

$$\delta_i^d = \frac{2(L_i^d - L_{i-1}^d)}{(L_i^d + L_{i-1}^d)(J_i - J_{i-1})} @0.5(J_i + J_{i-1}). \quad (5)$$

The difference between the measured and GSDF contrast responses at any given point $\kappa_\delta = \mathrm{Max}(|\delta_i - \delta_i^d|)$, should be less than 10% and 20% for the primary and secondary class display devices, respectively (Fig. 6). This criterion applies specifically to contrast evaluated from the 18 measurements of luminance made at uniformly spaced *p*-value intervals. The failure of a display device to meet the above criteria should prompt adjustment, recalibration, repair, or replacement of the device.

III.D. Luminance Dependencies

The luminance response evaluations described above only pertain to the luminance characteristics of a display device at one location on the display faceplate viewed perpendicularly. However, display devices often exhibit spatial luminance non-uniformities and variation in contrast as a function of viewing angle, both of which should be characterized as a part of display evaluation protocol.

III.D.1. Visual Evaluation

III.D.1.a. Nonuniformity The visual method for assessing display luminance uniformity involves the TG18-UN10 and TG18-UN80 test patterns. The patterns are displayed and the uniformity across the displayed pattern is visually assessed from a viewing distance of 30 cm. The patterns should be free of gross nonuniformities from center to the edges. No luminance variations with dimensions on the order of 1 cm or larger should be observed.

III.D.1.b. Angular Dependence Angular response may be evaluated visually using the TG18-CT test pattern. The pattern should first be viewed on axis to determine the visibility of all half-moon targets. The viewing angle at which any of the on-axis contrast thresholds are rendered invisible should then be determined by changing the viewing orientation in

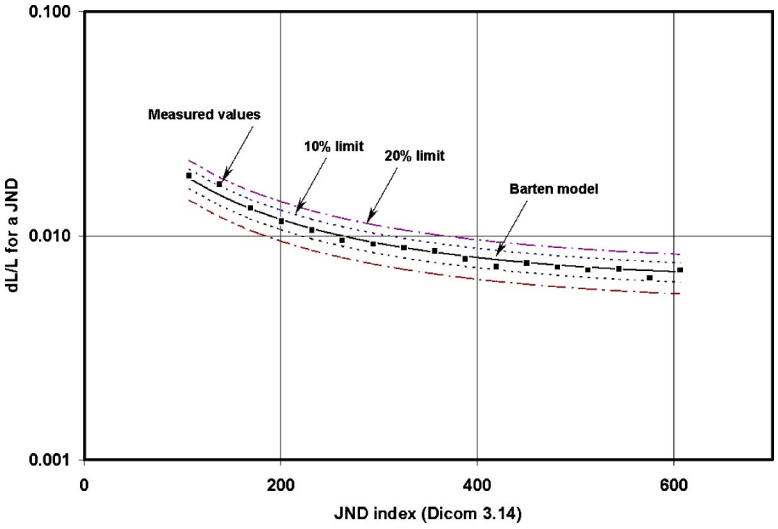


FIG. 6. Example of the contrast response computed from 18 gray levels is related to the expected contrast response associated with the DICOM GSDF with 10% tolerance limits indicated.

polar and azimuthal orientations. Alternatively, a uniform test pattern with uniformly embedded test targets may be used. The viewer distance at which all targets along the axial or diagonal axes are visible may be used as an indication of the angular response performance of the display in terms of the viewing angle cone within which the performance is acceptable. The acceptable viewing angle cone should be clearly labeled on the display device.

III.D.2. Quantitative Evaluation

III.D.2.a. Nonuniformity Using the TG18-UNL10 and TG18-UNL80 test patterns, luminance is measured at five locations over the faceplate of the display device (center and four corners) using a calibrated luminance meter with attentions to the precautions noted in Sec. III C. The maximum luminance deviation for each display pattern is calculated as the percent difference between the maximum and minimum luminance values relative to their average value, $200^{*}(L_{\max}-L_{\min})/(L_{\max}+L_{\min})$. The value for an individual display device should be less than 30%.

III.D.2.b. Angular Dependence The luminance of a LCD display may be quantitatively evaluated as a function of viewing angle. This can be done with two basic approaches: the *conoscopic* and the *gonioscopic* methods, as noted in the TG18 report. A basic quantitative test should include the evaluation of luminance ratio as a function of viewing angle using the TG18-LN test patterns. For these measurements, it is useful to have a subjective understanding of the viewing angle dependence to determine the specific horizontal and vertical angles at which quantitative measurements should be made.

Ideally, the angular response of a display should not reduce the luminance ratio by more than 30%. Thus, an acceptable viewing angle is defined as an angular cone within which LR′ is greater than 175 (250×0.7) for primary displays and 70 (100×0.7) for secondary displays.[22] If the luminance in midluminance values is measured, the angular luminance results should be evaluated the same as the on-axis measurements described above in terms of conformance to the GSDF. The contrast response for any viewing angle should not be greater than three times the expected limits on axis ($\kappa_\delta\leqslant3\times10\%=30\%$ for primary displays and $\kappa_\delta\leqslant3\times20\%=60\%$ for secondary displays). For a display device, both LR′ and κ_δ requirements should be met.

The viewing angle limitation for medical use of a device should be clearly labeled on the device for optimum viewing. If multiple devices of the same design are used, it is sufficient to assess the viewing angle limits on one device. For such systems, the acceptable viewing angle cone should be used to arrange the monitors for minimum contrast reduction due to the angular dependencies of luminance.

III.E. Display Resolution

Resolution is the ability of a display device to present the spatial details of a displayed image. This ability is related to both the number of pixels *and* the actual spatial extent of each pixel. Because of various optical and electronic processes, a display pixel can have a breadth that is larger than its nominal value, degrading the display resolution from its ideal level.

III.E.1. Visual Evaluation

Display resolution can be evaluated by visually assessing the appearance of the "Cx" patterns in the TG18-QC or the TG18-CX test patterns. The patterns should be displayed so that each image pixel is mapped to one display pixel. Most image viewers have the function to accomplish this display mode. In order not to be limited by the modulation transfer function (MTF) of the eye, the use of a magnifying glass is recommended. In the TG18-QC pattern, the examiner should inspect the displayed "Cx" targets at the center and four corners of the pattern and score the appearance using the provided scoring scale. The line-pair patterns at Nyquist and half-Nyquist frequencies in the horizontal and vertical directions should also be evaluated in terms of visibility of the lines. The average brightness of the patterns should also be evaluated using the grayscale step pattern as a reference. The difference in visibility of test patterns between horizontal and vertical patterns should be noted. The relative width of the black and white lines in these patches should also be examined using a magnifier. The resolution uniformity may be ascertained across the display area using the TG18-CX test pattern and a magnifier in the same way that the "Cx" elements in the TG18-QC pattern are evaluated.

In the visual inspection of the TG18-QC and TG18-CX patterns on primary class display systems, the Cx elements should be scored between 0 and 4 at all locations. This limit coincides with a resolution-addressability ratio (RAR) $\leqslant1.15$.[8] For secondary class displays, the Cx scores should be between 0 and 6 (RAR$\leqslant1.47$). For both classes, the horizontal and vertical line-pair patterns at Nyquist frequency should be discernible at all locations and for all directions. The TG18 report further includes a method to determine the extent of the display pixels (i.e., RAR) using the TG18-PX test pattern.

III.E.2. Quantitative Evaluation

Quantification of the MTF requires the use of a displayed-image digitizing system, such as a digital camera, to digitally capture a portion of the display and to analyze the resulting images. The lens flare should be reduced with the use of a high f number and the aid of a cone or funnel device. The magnification of the lens should result in over-sampling of the display with at least 64 camera pixels covering one display pixel. The camera needs to be well focused on the screen of the display under test. This is best done when the lens aperture is opened to its maximum level to achieve low depth-of-focus. Afterward, the lens aperture is set to its smallest level in order to achieve a large depth-of-focus and minimum flare.

The TG18-RV, TG18-RH, and TG18-NS patterns provide line inputs as target patterns for the MTF measurements. These patterns allow the assessment of MTF in the horizontal and vertical directions at three luminance levels and five

locations on the display area. At each location, the camera should be securely positioned in the normal direction in front of the target area of the display and focused on the line. The magnification should be determined in accordance with the display pixel size, camera matrix size, and the desired oversampling. The camera field of view should include the pixel markers in the pattern. While the camera should be placed in normal direction with respect to the faceplate, it needs to be rotated parallel to the faceplate such that the camera pixel array is angled at 2–5 deg with respect to that of the displayed image. Images from all six patterns should be captured before moving the camera to the next location. The exposure time should be selected such that the digital signal of the camera exceeds the dark signal by a factor of 100. Furthermore, the exposure time should be long enough to permit integration over multiple display frames, but short enough with respect to instabilities of the scanning and deflection circuits. Ultimately the integration time should be appropriate with respect to the integration time of the human eye, for which the experiments are conducted. Integration times between 0.2 and 1 s are appropriate to use. The measurements should be made in a darkened room.

The 30 images should be acquired without any image compression. The data should be transferred to a computer for data processing. The captured line patterns should be reduced to orthogonal MTFs using Fourier analysis. There are several processing steps in the calculations, and the results are expected to vary slightly with the methods. For standardization and simplicity, the following steps are suggested:[23]

1. Determine the size that the image pixels represent in terms of the spatial dimension on the display using the known physical distance of the pixel markers on the patterns and the measured pixel distance of the markers in the captured images.
2. Linearize the image data with respect to display luminance using the luminance response of the display.
3. Add the mean value of the image from the TG18-NS to that of the TG18-RV (or TG18-RH) pattern, and subtract the TG18-NS image pixel by pixel from the TG18-RV (or TG18-RH) image in order to remove display pixel structure. Averages of multiple images may be used for more complete removal of structured noise. The subtracted image is used for further processing.
4. Identify a central rectangular region of interest (ROI) extending along the image of the line.
5. Determine the angle of the line.
6. Reproject the two-dimensional (2D) data within the ROI along the direction of the line into subpixel bins to obtain the composite line spread function (LSF).
7. Smooth the LSF if it expresses excessive noise.
8. Find the Fourier transform of the LSF, and normalize the resulting MTF.
9. Divide the MTF by the *sinc* function associated with the width of the LSF subpixel bins, and correct for the previously characterized MTF of the camera system (see Sec. 3.1.2 of the report).

Note that in some cases the LSF might be asymmetric. In those cases, each side of the LSF is used to form two symmetric LSFs. The resultant MTFs are reported, along with their average, as representative of the display resolution. Values of the measured MTF at the Nyquist frequency should be at least 35% for primary display devices and 25% for secondary devices. Measured responses outside the acceptable range should prompt corrective actions in the form of focus or dithering adjustments, repair, or replacement of the device.

III.F. Display Noise

Display noise refers to statistical fluctuations in the image that either vary spatially, so-called spatial noise, or vary in time, so-called temporal noise. Temporal noise, which is usually dominant in the dark regions of displayed images, is difficult to characterize outside of a laboratory setting and its perceptual influence is less well understood. Spatial noise is dominant in the brighter areas of displayed images. In CRTs, phosphor granularity is the main contributor to spatial noise; while in LCDs, the dominant noise is that associated with the pixelated background.

III.F.1. Visual Evaluation

The visual method to quantify the spatial noise of a display system is based on the method to determine just noticeable luminance differences as a function of size using the TG18-AFC test pattern. Each quadrant of the test pattern contains a large number of regions with varying target position. In each quadrant, the contrast and size of the target are constant. The contrast-size values for the four quadrants are 20-2, 30-3, 40-4, and 60-6. The observer should view the patterns from a viewing distance of 30 cm. The quadrants can be subjectively evaluated to establish the contrast-size relationships for which the observer can confidently place the position of all targets. The target visibility in each of the target regions may also be quantified by counting the number of targets readily visible in each of the quadrants and computing the percent correct.

The visual evaluation should render all the targets except the smallest one visible for primary class displays and the two largest sizes visible for secondary class displays. Since the mean value and the standard deviation of the background are each linearly dependent on the luminance, their ratio, i.e., signal-to-noise, remains independent of luminance.[24,25] Therefore, the results of the noise evaluation are independent of the absolute luminance value of the pattern's background. However, the failure of a device in this test can also be an indication of an improper luminance response, the possibility of which can be eliminated by first verifying the proper luminance response of the device.

III.F.2. Quantitative Evaluation

Spatial noise of a display system can be quantified by either single-pixel signal-to-noise ratios[24] or by the normalized noise power spectrum (NPS). Both methods require the

use of a scientific-grade digital camera to capture an image of a uniform pattern displayed on the device. The camera lens should be set to a high f number in order to reduce veiling glare in the camera. Also, the magnification of the lens should result in over-sampling of the display in a way that allows sampling of spatial frequencies up to 40 cycles per degree, which is the resolution limit of the human visual system at the maximum luminance of most electronic displays.[26] The camera images should also be flat-field-corrected, compensated for gain variations, and restored for the degradation of the MTF of the camera optics based on the prior performance evaluation of the camera system, noted earlier.

The central region of the TG18-NS test patterns can be used as the target uniform pattern for measurements at three luminance levels. The camera should be securely positioned in front of the target area of the display and focused. The field of view should include the pixel markers in the pattern. The magnification should be determined in accordance with the display pixel size, camera matrix size, and the desired over-sampling. To eliminate the effects of temporal fluctuations in the luminance output, images should be captured with an integration time of about one second. The measurements should be performed in a darkened room. The images should be transferred uncompressed to a computer for data processing.

The quantification of the display noise by the single-pixel signal-to-noise ratio is noted in the TG18 report. For the NPS determination, the captured uniform patterns are processed by Fourier analysis. There are multiple processing steps involved and the methods can vary the results slightly. For standardization and simplicity, the following steps are suggested:

1. Determine the size that the image pixels represent in terms of the spatial dimension on the display using the known physical distance of the pixel markers on the pattern and the measured pixel distance of the markers in the captured image.
2. Linearize the image data with respect to display luminance.
3. Divide the central 3/4 region of the captured image into multiple, nonoverlapping regions, 128×128 or 256×256 in size. The size of these regions determine the sampling interval of the resulting NPS. Depending on the exact level of magnification (oversampling) and the matrix size of the camera, between nine to 64 regions may be identified. It is recommended that at least 20 regions be used for the assessment of the NPS. To achieve this, it might be necessary to acquire multiple images from the central patch of the TG18-NS pattern by orienting the camera toward another, nonoverlapping area of the central area of the displayed pattern.
4. Apply a two-dimensional fast Fourier transform on each region to yield the 2D NPS.
5. Average the 2D NPS from all regions.
6. Correct for the camera noise. Based on the assumption that the camera noise and the display spatial noise are uncorrelated, the NPS based on sampled camera images without exposure using the same integration time may be subtracted from the results.
7. Derive the orthogonal NPS from the calculated 2D NPS by band averaging, excluding the data on the orthogonal axes.

Since there are currently only a few examples of actual NPS measurements made, and since no correlation of the measurements and diagnostic accuracy is ascertained, no fixed criteria are recommended at this time. However, noise values associated with the display device should not exceed those of typical radiological images that are viewed with the system.

III.G. Veiling Glare

Veiling glare is a light-spreading phenomenon in a display device that leads to the degradation of image contrast in the presence of strong surrounding brightness. In CRTs, veiling glare is caused by internal light-scattering processes in the device's faceplate, light leakage, and electron backscattering. In LCDs, electronic cross-talk can be viewed as a form of veiling glare.

III.G.1. Visual Evaluation

The visual assessment of veiling glare can be accomplished using the TG18-GV and TG18-GVN test patterns. The display size must be adjusted so that the diameter of the white region is 20 cm. The observer should discern the visibility of the low-contrast objects in sequential viewing of the TG18-GVN and TG18-GV patterns. Because the human visual systems will change adaptation if it views the bright field, it is imperative that the bright field is fully blocked from view and that no reflected light from the bright field be observable. This may be accomplished by the use of a mask or cone, which shields the human eye from the surround luminance of the pattern. No significant reduction in the contrast of the target objects should be observed between the two patterns. At least three objects should be readily visible in either pattern for primary class display devices. The corresponding object for secondary class display devices is at least one (5th) target.

III.G.2. Quantitative Evaluation

The quantitative evaluation of veiling glare is accomplished using a highly collimated luminance meter and the TG18-GQ, TG18-GQB, and TG18-GQN test patterns. The display size must be adjusted so that the diameter of the white region is 20 cm. Furthermore, the bright luminance surrounding the central measurement point at the center of the test patterns should be blocked using either a baffled luminance meter or a telescopic luminance meter with a light-blocking baffled funnel or cone. Using either of these devices, the luminance in the center of the central dark region of the TG18-GQ pattern L, the white luminance in the center of the white region of the TG18-GQB pattern L_B, and

the background luminance value in the center of the TG18-GQN pattern L_N, are recorded. The glare ratio for the display is then computed as

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

The veiling glare for a high fidelity display system should not change the contrast of a target pattern by more than 20% with and without a bright surrounding. Thus, the luminance from veiling glare should not be more than 25% of the minimum luminance for the normal operating settings of the display. Since the ratio of the maximum luminance to the minimum luminance should be about 250, this implies a glare ratio of 1000, which is typical of measurements made for transilluminated film. However, the recommended test pattern presents a scene with significantly more veiling glare in the target region than is encountered in medical imaging scenes. Though not as strict criteria which may not be achievable by certain display technologies, TG18 recommends a glare ratio greater than 400 and 150 for primary and secondary display devices, respectively.[27,28]

III.H. Display Chromaticity

In display devices, chromaticity refers to the intrinsic magnitude and uniformity of color tint of the device when displaying a monochrome image. In monochrome CRTs, color tint is dictated by the phosphor type, and can vary slightly from monitor to monitor. In LCDs, color tint is dictated by the color temperature of the backlight. Color tint is usually considered a preference issue. However, it can be a cause of distraction, especially in multiple monitor workstations where the color tints are mismatched.

III.H.1. Visual Evaluation

The visual assessment of color uniformity is performed using the TG18-UN80 test pattern. The pattern is displayed on all the display devices associated with a workstation, and the relative color uniformity of the displayed pattern across the display area of each display device and across different display devices is discerned. No significantly perceivable color differences should be present among display devices and across the display area of each device for primary class devices. No requirements are specified for secondary class displays.

III.H.2. Quantitative Evaluation

The TG18-UNL80 test pattern is displayed on all the display devices associated with a workstation. A colorimeter is then used to measure the (u',v') color coordinates at the center and at the four corners of the display area of each display device, and these coordinates averaged to produce a mean (u',v') chromaticity measurement for the display device. The measurements on all display devices are used to compute the color uniformity index as the maximum distance in (u',v') space between any possible pair of average (u',v') points using $D=((u_1'-u_2')^2+(v_1'-v_2')^2)^{1/2}$. If the colorimeter used outputs the color coordinate in the older (x,y) space, the values can be converted to (u',v') space using the following transformations:

$$u' = 4x/(-2x + 12y + 3), \quad v' = 9y/(-2x + 12y + 3); \tag{7}$$

or

$$x = 27u'/(18u' - 48v' + 36),$$

$$y = 12v'/(18u' - 48v' + 36). \tag{8}$$

Based on clinical experience, a color uniformity parameter of 0.01 or less is necessary to assure acceptable color matching of primary class grayscale display devices of a workstation.[29] The distance between any pair of color coordinates across the display area of each device should also not exceed this limit. No quantitative requirements are specified for secondary class displays.

III.I. Miscellaneous Tests

In addition to the primary display attributes described above, there are a number of secondary attributes that may need to be addressed in a full display performance evaluation. Those include video artifacts, moiré artifacts, color artifacts, physical defects, flicker, and electronic cross talk. Brief descriptions and assessment methods for these characteristics are outlined in the TG18 report.

III.J. Overall Evaluations

In addition to the testing of a display device for a specific performance characteristic, the overall quality of a system can be assessed using a comprehensive visual/quantitative approach. Overall assessment can be based on any of the TG18-recommended multipurpose test patterns. Each pattern should be displayed with one display pixel representing each image pixel and examined from a viewing distance of 30 cm. The findings can be correlated with the results of more focused testing methods specified above and serve as a basis for quality control assessments. The frequency of such an evaluation is discussed in Sec. 6 of the full report. Evaluations based on TG18-QC and TG18 anatomical patterns are outlined below.

III.J.1. Evaluations using TG18-QC Pattern

The appearance of the elements in the TG18-QC test pattern [Fig. 1(a)] can be used to assess the overall performance of a display system. The following are recommended:

1. General image quality and artifacts: Evaluate the overall appearance of the pattern. Note any non-uniformities or artifacts, especially at black-to-white and white-to-black transitions. Verify that the ramp bars appear continuous without any contour lines.
2. Geometric distortion: Verify that the borders and lines of the pattern are visible and straight and that the pattern appears to be centered in the active area of the display device. If desired, measure any distortions.

TABLE VIII. Criteria for evaluating the TG18 anatomical images.

Test pattern	Evaluation criteria
TG18-CH	Degree of difficulty for exam
	Overall contrast
	Overall sharpness
	Symmetrical reproduction of the thorax, as shown by the central position of a spinous process between the medial ends of the clavicles
	Medial borders of the scapulae
	Reproduction of the whole rib cage above the diaphragm
	Visually sharp reproduction of the vascular pattern of the lungs, particularly the peripheral vessels
	Sharp reproduction of the trachea and proximal bronchi
	Sharp reproduction of the borders of the heart and the aorta
	Sharp reproduction of the diaphragm
	Visibility of the retrocardiac lung and the mediastinum
	Visibility of subdiaphragmatic features
	Visibility of the spine through the heart shadow
	Visibility of small details in the whole lung, including the retrocardiac areas
	Visibility of linear and reticular details out to the lung periphery
TG18-KN	Degree of difficulty for exam
	Overall contrast
	Overall sharpness
	Reproduction of trabecular detail
	Reproduction of bony and soft tissue
TG18-MM1 and TG18-MM2	Degree of difficulty for exam
	Overall contrast and brightness
	Overall sharpness (no blur)
	Sharp appearance of Cooper's ligaments
	Structure of the clip and the presence of the gap at its apex (TG18-MM1 only)
	Appearance and visibility of subtle microcalcifications (TG18-MM1 only)
	Visibility of structures at the margins of the breast (TG18-MM1 only)

3. Luminance, reflection, noise, and glare: Verify that all 16 luminance patches are distinctly visible. Measure their luminance using a luminance meter if desired, and evaluate the results in comparison to the DICOM GSDF. Verify that the 5% and 95% patches are visible. Evaluate the appearance of low contrast letters and the targets at the corners of all luminance patches with and without ambient lighting.
4. Resolution: Evaluate the Cx patterns at the center and corners of the pattern and grade them compared to the reference score. Also verify the visibility of the line-pair patterns at the Nyquist frequency at the center and corners of the pattern, and if desired, measure the luminance difference between the vertical and horizontal high-modulation patterns.

III.J.2. Evaluations using Anatomical Images

A radiologist should evaluate the overall clinical image quality of the display using patient images. The TG18 report suggests four specific anatomical images for this purpose: TG18-CH, TG18-KN, TG18-MM1, and TG18-MM2 [Figs. 1(r)–(u)]. These correspond to a chest radiograph, a knee radiograph, and two digital mammograms. Clinical criteria for evaluating these images are given in Table VIII. The images may be scored according to these criteria corresponding to the different image features. The radiologist who wishes to evaluate his/her display should independently rate the image features according to the criteria in Table VIII, then compare their ratings to those obtained with a high-quality transilluminated film print of the patterns. Significant discrepancies need to be brought to the attention of the responsible medical physicist or service engineer.

IV. CONCLUSIONS

Electronic display is a key component of medical imaging systems as it serves as the final element of the imaging chain. Due to hardware variability and degradation over time, it is important to assure that a medical display system is appropriate for the intended medical application and that its performance is stable over time. Acceptance testing and quality control testing of medical display devices are essential requirements for high-quality medical practice. The guidelines established by the AAPM Task Group 18 delineate specific testing procedures and acceptance criteria for that purpose that can be readily implemented in a clinical setting.

[a]Electronic mail: samei@duke.edu

[1]S. J. Ackerman, J. N. Gitlin, R. W. Gayler, C. D. Flagle, and R. N. Bryan, "Receiver operating characteristic analysis of fracture and pneumonia detection: comparison of laser-digitized workstation images and conventional analog radiographs," Radiology **186**, 263–268 (1993).

[2]W. W. Scott, J. E. Rosenbaum, S. J. Ackerman, R. L. Reichle, D. Magid, J. C. Weller, and J. N. Gitlin, "Subtle orthopedic fractures: Teleradiology workstation versus film interpretation," Radiology **187**, 811–815 (1993).

[3]W. W. Scott, D. A. Bluemke, W. K. Mysko, G. E. R. Weller, G. D. Kelen, R. L. Reichle, J. C. Weller, and J. N. Gitlin, "Interpretation of emergency department radiographs by radiologists and emergency medicine physicians: teleradiology workstation versus radiograph readings," Radiology **195**, 223–229 (1995).

[4]"AAPM, The role of the clinical medical physicist in diagnostic radiology," Report of the AAPM Task Group No. 2 (AIP, New York, 1994).

[5]E. Muka, H. Blume, and S. Daly, "Display of medical images on CRT soft-copy displays: a tutorial," Proc. SPIE **2431**, 341–359 (1995).

[6]E. Senol and E. Muka, "Spatial frequency characteristics of CRT soft-copy displays," Proc. SPIE **2431**, 302–315 (1995).

[7]E. F. Kelley, G. R. Jones, P. A. Boynton, M. D. Grote, and D. J. Bechis, "A survey of the components of display measurement standards," *SID Digest* (Society for Information Display Headquarters, San Jose, CA, 1995), pp. 637–640.

[8]E. Muka, T. Mertelmeier, R. M. Slone, and E. Senol, "Impact of phosphor luminescence noise on the specification of high-resolution CRT displays for medical imaging," Proc. SPIE, **3031**, 210–221 (1997).

[9]H. Roehrig, H. Blume, T. L. Ji, and M. Browne, "Performance tests and quality control of cathode ray tube displays," J. Digit Imaging **3**, 134–145 (1990).

[10]J. E. Gray, "Use of SMPTE test pattern in picture archiving and communication systems," J. Digit Imaging **5**, 54–58 (1992).

[11]R. D. Nawfel, K. H. Chan, D. J. Wagenaar, and P. F. Judy, "Evaluation of video gray-scale display," Med. Phys. **19**, 561–567 (1992).

[12]D. A. Reimann, M. J. Flynn, and J. J. Ciarelli, "A system to maintain perceptually linear networked display devices," Proc. SPIE **2431**, 316–326 (1995).

[13]M. P. Eckert and D. P. Chakraborty, "Video display quality control measurements for PACS," Proc. SPIE **2431**, 328–340 (1995).

[14]H. Kato, "Hard- and soft-copy image quality," *RSNA Categorical Course in Physics* (Radiological Society of North America (RSNA) Publication, Categorical Course Syllabus, Oak Brook, IL, 1995), pp. 243–258.

[15]NEMA PS 3.14-2000, "Digital imaging and communications in medicine (DICOM) Part 14: Grayscale display standard function," National Electrical Manufacturers Association, NEMA, 1300 North 17th Street, Suite 1847, Rosslyn, VA 22209, http://www.nema.org/ Draft viewable at http://medical.nema.org/dicom/2000/draft/00_14DR.PDF, Available for purchase in whole or in parts from Global Sales, Global Engineering Documents, 15 Inverness Way, East Englewood, Colorado, 80112; http://global.ihs.com/

[16]E. Samei, A. Badano, D. Chakraborty, K. Compton, C. Cornelius, K. Corrigan, M. J. Flynn, B. Hemminger, N. Hangiandreou, J. Johnson, M. Moxley, W. Pavlicek, H. Roehrig, L. Rutz, J. Shepard, R. Uzenoff, J. Wang, and C. Willis, "Assessment of display performance for medical imaging systems," Report of the American Association of Physicists in Medicine (AAPM) Task Group 18 (Medical Physics Publishing, Madison, WI, 2005), AAPM On-line Report No. 03.

[17]American College of Radiology (ACR), ACR Standard for Teleradiology, 1999. American College of Radiology, 1891 Preston White Drive, Reston, VA 20191-4397; http://www.acr.org/

[18]P. G. J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality* (SPIE Press, Bellingham, WA, 1999).

[19]M. J. Flynn, J. Kanicki, A. Badano, and W. R. Eyler, "High fidelity electronic display of digital radiographs," Radiographics **19**, 1653–1669 (1999).

[20]E. Samei, in "Digital mammography displays," *Advances in Breast Imaging: Physics, Technology*, and *Clinical Applications*, edited by A. Karellas and M. L. Giger (Radiological Society of North America (RSNA) Publication, Categorical Course Syllabus, Oak Brook, IL, 2004), pp. 135–143.

[21]H. Blume, A. M. K. Ho, F. Stevens, and P. M. Steven, "Practical aspects of gray-scale calibration of display systems," Proc. SPIE, **4323**, 28–41 (2001).

[22]E. Samei and S. L. Wright, "The effect of viewing angle response on DICOM compliance of LCD displays," Proc. SPIE Medical Imaging **5371**, 170–177 (2004).

[23]E. Samei and M. J. Flynn, "A method for in-field evaluation of the modulation transfer function of electronic display devices," Proc. SPIE **4319**, 599–607 (2001).

[24]H. Roehrig, T. L. Ji, M. Browne, W. J. Dallas, and H. Blume, "Signal-to-noise ratio and maximum information content of images displayed by a CRT," Proc. SPIE **1232**, 115–133 (1990).

[25]H. Roehrig, H. Blume, T. L. Ji, and M. K. Sundareshan, "Noise of CRT display systems," Proc. SPIE **1897**, 232–245 (1993).

[26]A. Rose, *Vision, Human and Electronic* (Plenum Press, New York, 1974).

[27]M. J. Flynn and A. Badano, "Image quality degradation by light scattering in display devices," J. Digit Imaging **12**, 50–59 (1999).

[28]A. Badano, M. J. Flynn, and J. Kanicki, "Accurate small-spot luminance measurements in display devices," Displays **23**, 177–182 (2002).

[29]K. A. Fetterly, S. N. Bernatz, D. S. Groth, and N. J. Hangiandreou, "Quantitative color measurement for the characterization of grayscale PACS CRTs," Radiology **209(P)**, 320 (1998).