Display Quality Assurance: Recommendations from AAPM TG270 for Tests, Tools, Patterns, and Performance Criteria

Nicholas B. Bevins, Ph.D.
TG270 Co-chair

Display Check

![Display Check Image]
TG270 Goals

- Provide an update to the TG18 report
  - Test methodology
  - Test criteria
  - Test frequency
  - Test patterns

Outline

- Display Classifications
  - Diagnostic
  - Non-diagnostic

- Display Test Patterns
  - Existing Patterns
  - New Patterns

- Display Performance Evaluation
Display Classification

- Four classifications based on use
  - Diagnostic Displays
  - Non-diagnostic Displays (TG18 “secondary displays”)
    - Modality Displays
    - Clinical Specialist Displays
    - Electronic Health Record (EHR) Displays

- Diagnostic Displays (TG18 “primary displays”)
  - Primary interpretation of medical images
  - Improved performance characteristics
    - Luminance stability (both in level and uniformity)
    - Smaller pixel pitch
    - Lower noise
    - Greater bit depth
    - Self-testing functionality
  - Stringent performance criteria
  - High cost
  - Does not include navigation displays
Display Classification

- **Modality Displays**
  - Displays used during acquisition and generation of medical images
    - May or may not be attached to modality
    - Only displays that show images (not for acquisition control)

- **Clinical Specialist Displays**
  - Review of images before or independently of primary radiology read
    - ER, surgical environments
    - Patient care decisions, often before primary read by radiologist

- **EHR Display**
  - Images used to review images following interpretation
    - Referring physicians offices
    - Exam room with patient
    - Pre-surgical planning

- The goal of display QA is consistent image presentation across all displays (image review chain)
  - Similar goals, but different tolerances, tests, frequencies
Outline

- Display Classifications
  - Diagnostic
  - Non-diagnostic
- Display Test Patterns
  - Existing Patterns
  - New Patterns
- Display Performance Evaluation

Existing Test Patterns

- TG18-QC
- TG18-LN
- TG18-UN
- TG18-AFC
Existing Test Patterns

- SMPTE

“As a result of the pattern’s grayscale insensitivity and CRT-specific features, this report considers the SMPTE test pattern deprecated for qualitative display evaluation in favor of either quantitative measurement or updated test patterns.”

New Patterns – TG270-sQC

- Simple QC test pattern for routine checks by users, technologists, physicists
New Patterns – TG270-sQC

- Low contrast test patterns at multiple gray levels
- Spatial resolution verification
- Luminance patches for uniformity and min/max measurements

Continuous Gradient Effects

- No issues
- Mis-calibrated gray level
- Bit-depth configuration error
New Patterns – TG270-pQC

- Detailed QC pattern for physicists and other advanced users
- Same gray levels as sQC, but with more contrasts and frequencies
- Use as follow up to quantitative failures for context

New Patterns – TG270-pQC

- Low contrast patterns at multiple gray levels
- Spatial resolution verification
- Luminance patches for 18-point measure
- Continuous ramp
New Patterns – TG270-ULN

- Replaces the TG18 LN and UN pattern series
- Generated for all 256 8-bit gray levels
- Grid for quantitative uniformity measures

New Patterns – TG270-TR

- Temporal resolution pattern for qualitative evaluation of short-term temporal resolution
- Use to help guide purchasing decisions, display usage, latency effects
- Used with digital camera to capture frames
All of the new TG270 test patterns were generated using ImageJ macros (.ijm)
- Included with TG270 report
- Available on the TG wiki on AAPM website
Outline

- Display Classifications
  - Diagnostic
  - Non-diagnostic
- Display Test Patterns
  - Existing Patterns
  - New Patterns
- Display Performance Evaluation

Display Performance Evaluation Tools

- Equipment
  - Photometers and colorimeters
    - Contact and telescopic
    - External and internal
  - Loupe
  - Digital camera
Display Luminance

- Assessment of display luminance includes measuring:
  - $L_{\text{amb}}$
  - $L'_{\text{min}} (L_{\text{min}} + L_{\text{amb}})$
  - $L'_{\text{max}} (L_{\text{min}} + L_{\text{amb}})$
  - Luminance ratio ($L'_{\text{max}} / L'_{\text{min}}$)
  - Luminance response function

- Each of these is related to the others. Understanding these relationships is critical to proper display QA.

$L_{\text{amb}}$

- Ambient luminance is due to reflected light from the display
  - Specular reflection
  - Diffuse reflection

- Setting and maintaining proper environmental lighting for consistent and predictable image presentation

- Setting ambient lighting in reading rooms to minimize visual strain
  - 25-50 lux
Display Luminance

- Avoid Lamb effects from obscuring darkest regions of image
  \[ L_{\text{min}} = 4 \cdot L_{\text{amb}} \]
- Approximately 80% of contrast seen with no ambient lighting is still visible with ambient lighting
The minimum and maximum luminances are combined with the ambient luminance:
- \( L'_\text{min}, L'_\text{max} \)

The ratio gives the luminance ratio \( LR \):
\[
LR = \frac{L'_\text{min}}{L'_\text{max}}
\]

Recommended \( LR = 350 \)
- Set \( L'_\text{max} \) based on \( L'_\text{min} \) and LR, not maximum of display.
Luminance Response Function

- Measurement of luminance response function
  - 18-point (TG18 methodology)
  - 52-point
  - 256-point
  - 11-point (SMPTE pattern)

- Analysis of luminance should be of L’, which includes the effects of ambient luminance

Display Luminance

![Graph showing the relationship between luminance and gray level]
Luminance Response Function

- Confirm conformance with DICOM GSDF
  - Mean JND/GL
  - dL/L per JND
  - Both to within 10% for diagnostic displays, 20% for non-diagnostic

- More frequent qualitative verification
  - Test pattern based
    - TG270-sQC, TG270-pQC, TG18-QC
    - Verify contrast performance at multiples levels (especially in the darks)

Display Color (White Point)

- Color of the light output by the display throughout the grayscale
- Evaluate by measuring the color difference

\[ \Delta = \sqrt{(u'_1-u'_2)^2+(v'_1-v'_2)^2} \]

- Compared against
  - Other display
  - Standard illuminant (e.g., D65)
  - Full brightness (TG196 methodology)
Display Color (White Point)

- Standard illuminant (e.g., D65) should be used instead of correlated color temperature (CCT)
  - CCT is defined as multiple points in color space
  - The maximum difference between the points is large
Display Color (White Point)

- **Comparing two displays**
  
<table>
<thead>
<tr>
<th>Optimal Limit</th>
<th>Acceptable Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_{D65} (u', v') \leq 0.005$</td>
<td>$\Delta_{D65} (u', v') \leq 0.01$</td>
</tr>
<tr>
<td>$\Delta_{D65} (u', v') \leq 0.01$</td>
<td>$\Delta_{D65} (u', v') \leq 0.02$</td>
</tr>
</tbody>
</table>

- **Comparing display to standard illuminant**
  
<table>
<thead>
<tr>
<th>Optimal Limit</th>
<th>Acceptable Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_{D65} (u', v') \leq 0.005$</td>
<td>$\Delta_{D65} (u', v') \leq 0.01$</td>
</tr>
<tr>
<td>$\Delta_{D65} (u', v') \leq 0.01$</td>
<td>$\Delta_{D65} (u', v') \leq 0.02$</td>
</tr>
</tbody>
</table>

Display Uniformity

- Display uniformity evaluated both quantitatively and qualitatively
  
  - Quantitative assessment for global uniformity issues across display
  
  - Qualitative assessment for local non-uniformity

- Global uniformity is less important for clinical image review
  
  - Global non-uniformity is low frequency, likely not to be confused with anatomy
  
  - Local non-uniformities are common failures with flat panel displays, and are of similar size/contrast as image features
Display Uniformity

- New methodology for evaluating global uniformity

\[ \text{LUDM} = \max \left( 100 \times \frac{|L_n - L_{\text{med}}|}{L_{\text{med}}} \right) \]

- Evaluates all measured points against the median value
  - Measure 9 points (corners, edges, center)
  - Median less affected by outliers
  - LUDM < 30% for passing. At 15%, clinical impact should be evaluated visually

Display Uniformity

- Local non-uniformities
  - Mura
  - Bad pixels (stuck pixels)
  - Image burn-in

- Evaluated qualitatively
  - Must be done on site
  - Use multiple gray levels to evaluate
Display Uniformity

Display Noise

- Qualitative noise assessment for product evaluation
  - Test pattern (e.g., TG18-AFC) for pixel-by-pixel variation

- Use clinical images for evaluation of clinical impact

- Unnecessary for routine display quality assurance
Display Temporal Performance

- Several scales of temporal performance
  - Long term (luminance stability, uniformity)
  - Medium term (warm up time, image retention)
  - Short term (response time, input latency)

- Qualitative evaluation of short term performance
  - Evaluate impact of display performance on the viewing of dynamic images
    - Fluoroscopy, ultrasound, etc.
Modern flat-panel displays have discretized pixel structures, with little light dispersed into neighboring pixels.

Quantitative measures of spatial resolution unnecessary assuming:
- Advanced pixel structure (e.g., IPS, VA)
- Digital graphic interfaces (e.g., DVI-D, DisplayPort)

Visual verification of driver settings to native display resolution
- Magnifier, loupe is helpful
Display Spatial Resolution

- Pixel pitch selected depending on use and viewing distance
  - Minimize the appearance of pixel structure
  - Radiologist workstation recommended distance of 65 cm
    - Minimize eye strain
  - Other workstations often have larger viewing distance
    - Larger pixel pitch is acceptable

<table>
<thead>
<tr>
<th>Pixel Pitch</th>
<th>Radiologist Workstation</th>
<th>Modality, Other Clinical Workstation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 210 µm</td>
<td>&lt; 250 µm</td>
</tr>
</tbody>
</table>

Conclusion

- Display QA for flat-panel displays is an important part of general QA across all of medical imaging
- Awareness of current standards and guidelines is critical for appropriate QA
Status of Report

- Report draft circulating for comments
- Goal is final draft before RSNA 2017

- Intention is to incorporate report into other TG reports
  - No need to re-state display testing in every modality testing guideline
  - Replace references to TG18 in future reports

Thank you