Role of Color in Telemedicine Applications

Elizabeth A. Krupinski, PhD

Arizona Telemedicine Program

The University of Arizona Medical Center
Medical Imaging

Southwest TRC Telehealth Resource Center

Background

- Color displays common clinical practice
- Radiology growing acceptance & use
- Other "ologies" & telemed routinely used
- No *validated* color display calibration protocol for *medical* color image applications
- *No universally accepted image quality program* for color displays in medicine
- Significant lack data showing that poor/inappropriate calibration *affects diagnostic performance &/or workflow*

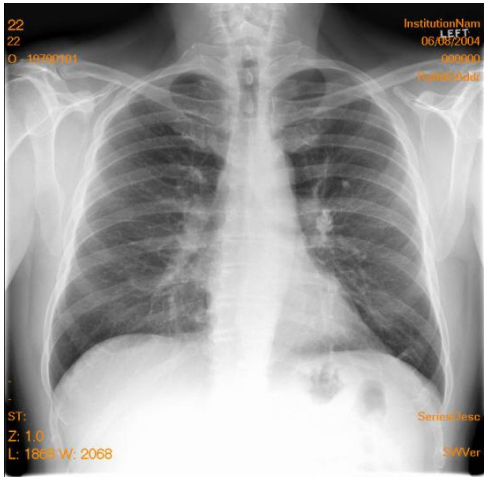
Telerad & Color

- **Color routine CT, MRI, US but what about bone, chest, mammo?**
- **Diagnostic accuracy & visual search efficiency high luminance (400 cd/m^2) MG color equivalent to high luminance (400 cd/m^2) MG monochrome display?**
- **Yes = MG color displays can be used as *all-purpose* diagnostic displays**

Krupinski et al. (2011). High luminance monochrome vs color displays: impact on performance & search. Proc SPIE Med Imag 7966 79661R1-6

Images

- 50 DR chest images
- Half solitary pulmonary nodule half without
- Locations confirmed DR & CT images
- Nodules size = 9 – 26 mm
- Shape = round or round with spiculations
- Subtlety = subtle to moderately subtle



Displays

- Eizo MG 3-Mpixel monochrome
- Eizo MG 3-Mpixel color
- Portrait mode
- Calibrated to DICOM GSDF
- Pixel pitch:
  color = 0.21075 mm
  (323.7 x 431.6 mm Diagonal: 540 mm)
  monochrome = 0.2115 mm
  (324.8 x 433.1 mm Diagonal: 541 mm)
- Max luminance = 400 cd/m^2
- Min luminance = 0.75 cd/m^2

Study Details

- 6 radiologists view images each display
- Counterbalanced min 3 wks between sessions
- Ambient room lights = 25 lux
- Viewing distance ~ 45 cm
- Window & level allowed
- Report nodule present or absent & rate confidence definite, probable, possible
- MRMC ROC used to analyze
- Total viewing time recorded

Eye-Position Study

- **Separate session viewed sub-set 15 images**
- **Eye-position recorded using 4000SU Eye-Tracker with head tracker**

Diagnostic Accuracy

No significant difference color vs mono

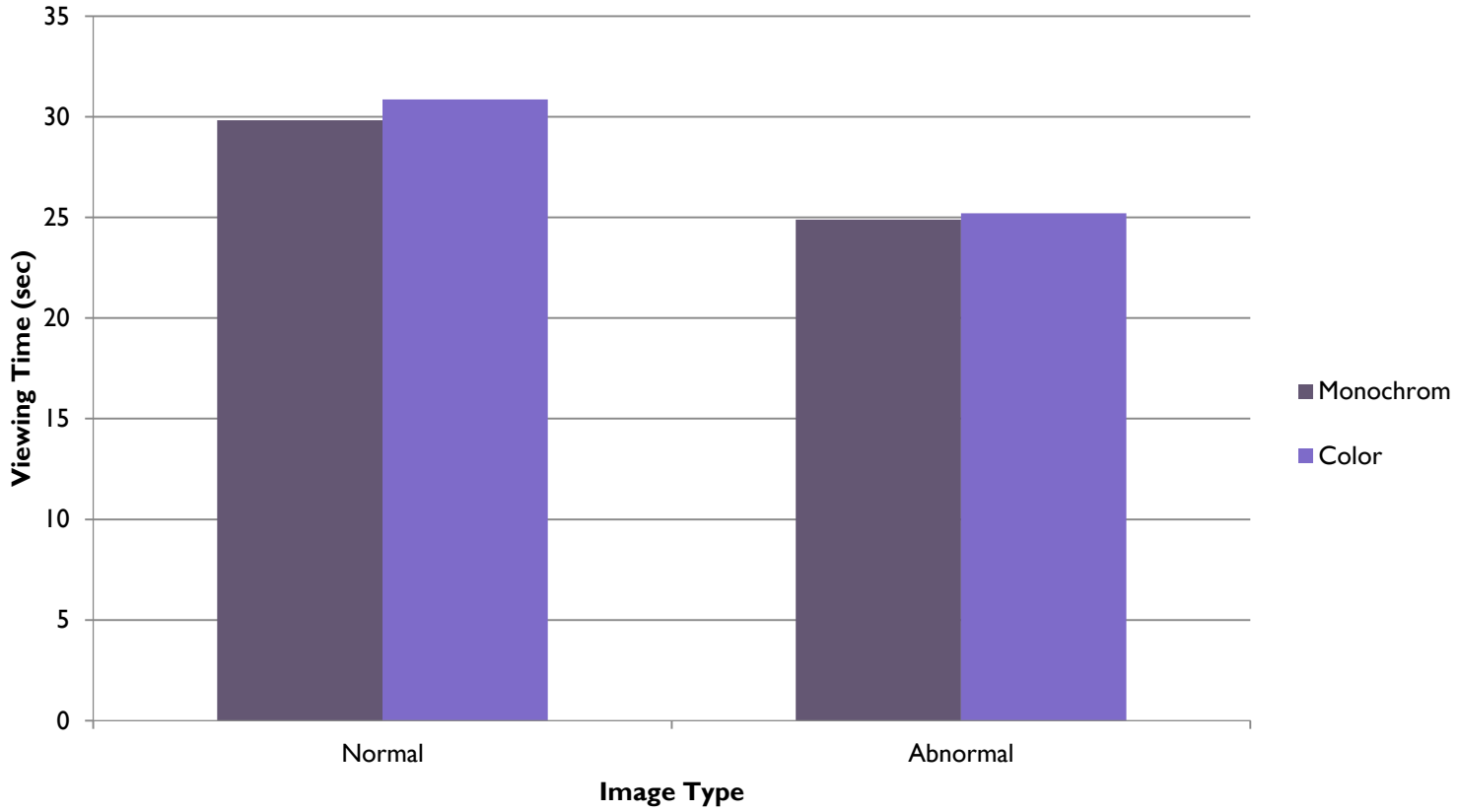
- **F = 0.0136, p = 0.9078**

Reader	Color Az	Monochrome Az
1	0.9209	0.9031
2	0.8913	0.8964
3	0.9002	0.8294
4	0.9455	0.9097
5	0.8597	0.9097
6	0.8710	0.9187
Mean	0.8981	0.8945

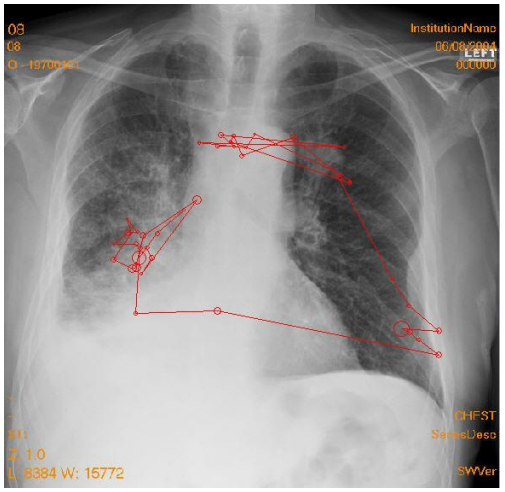
Total Viewing Time

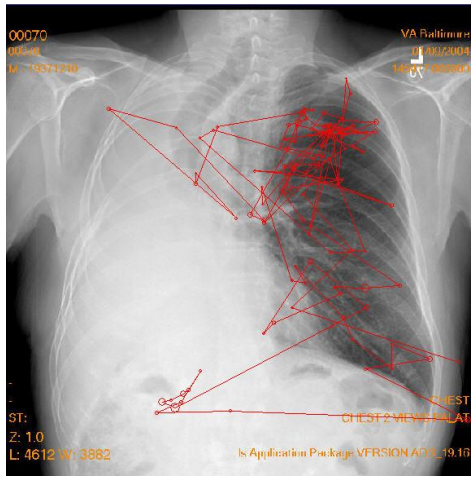
No significant differences color vs mono

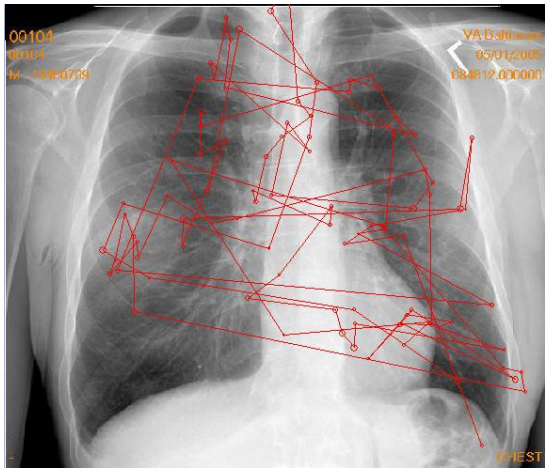
- F = 0.392, p = 0.5315



Typical Search Patterns







Red dots = fixation or location eye lands foveal vision

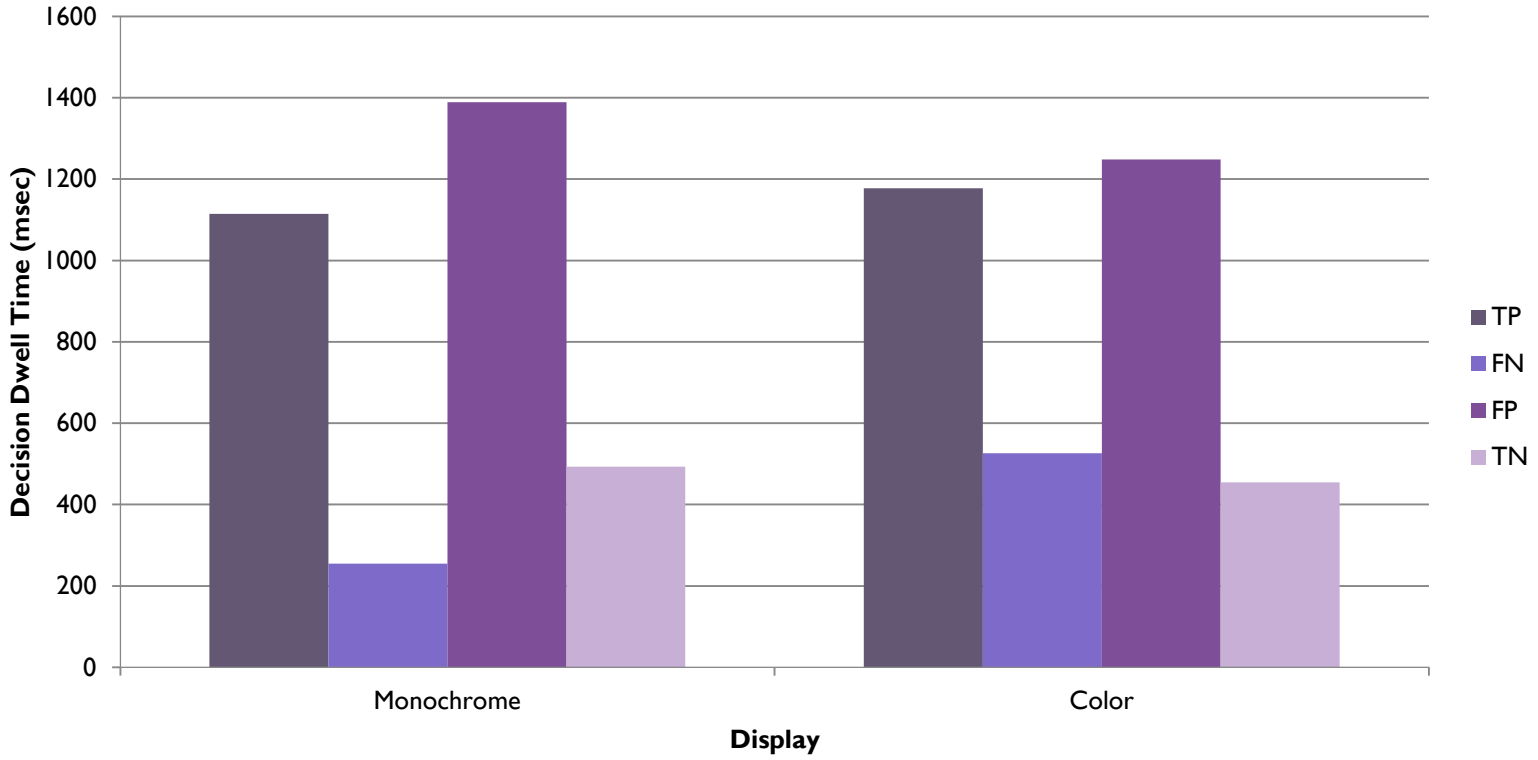
Size circle = dwell time: larger = longer dwell

Lines = order fixations generated during search

Decision Dwells

No significant differences color vs mono

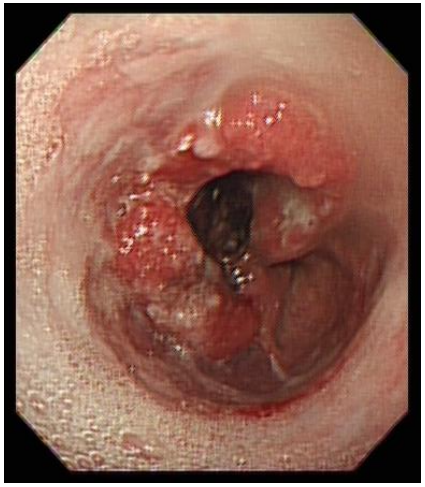
- **F = 0.133, p = 0.7154**

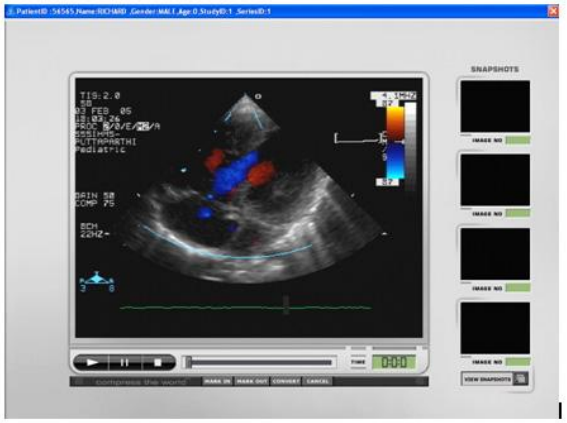


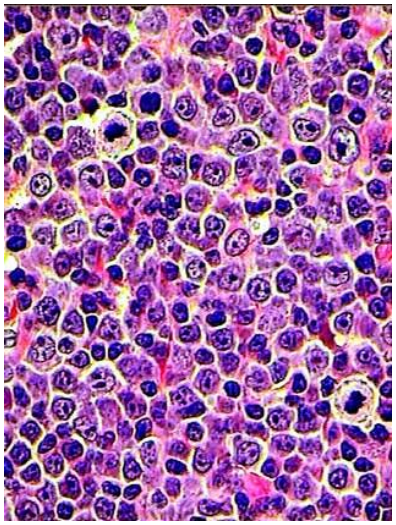
Summary

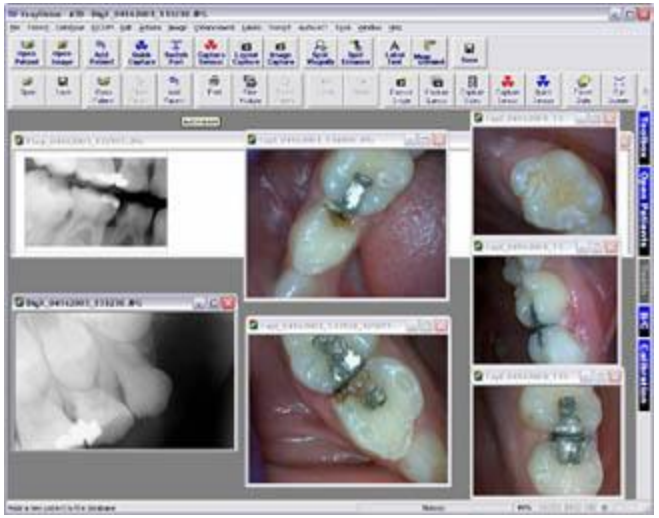
At least for task detecting chest nodules diagnostic accuracy & visual search efficiency *equivalent* using 3-Mpixel MG color & monochrome LCD displays

Common TM Applications

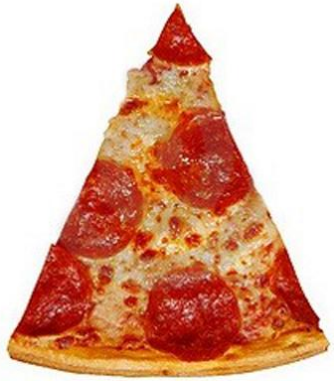






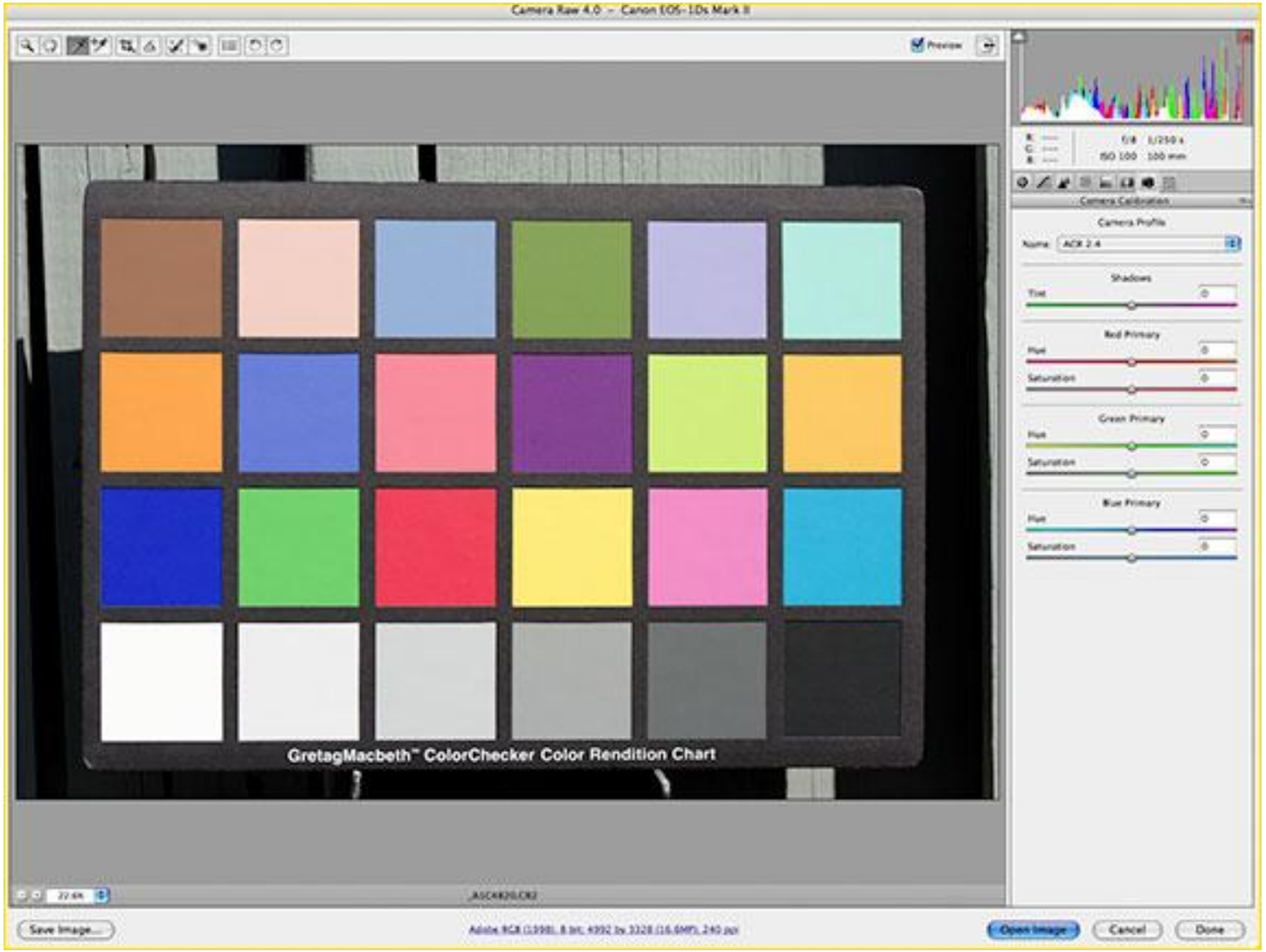


Common Calibration Tools

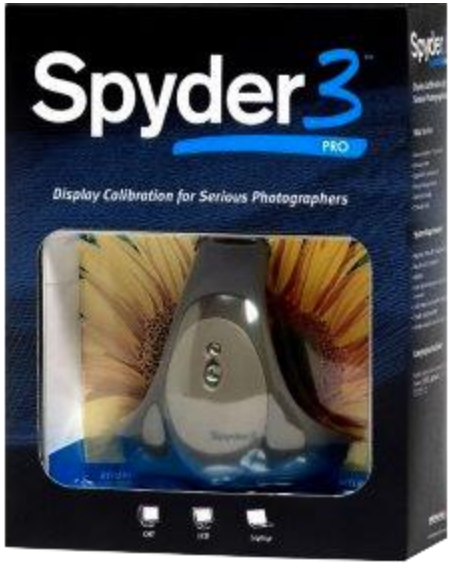


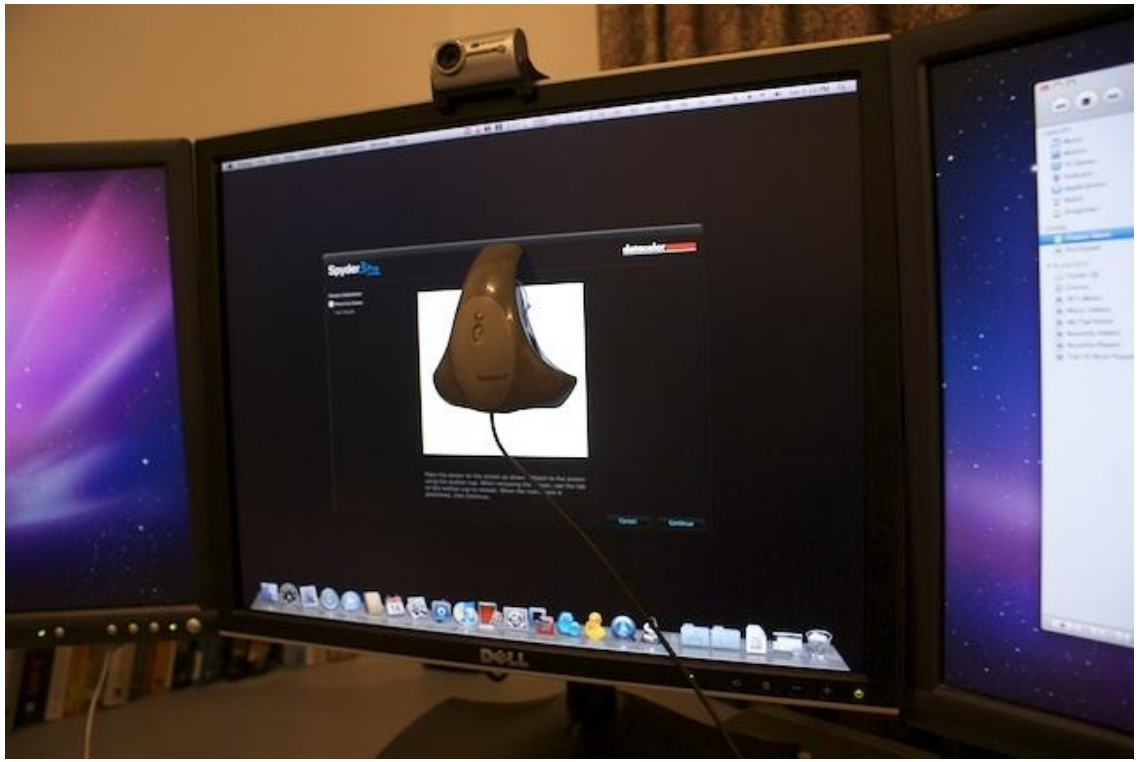
CLOSE ENOUGH

Common Calibration Tools

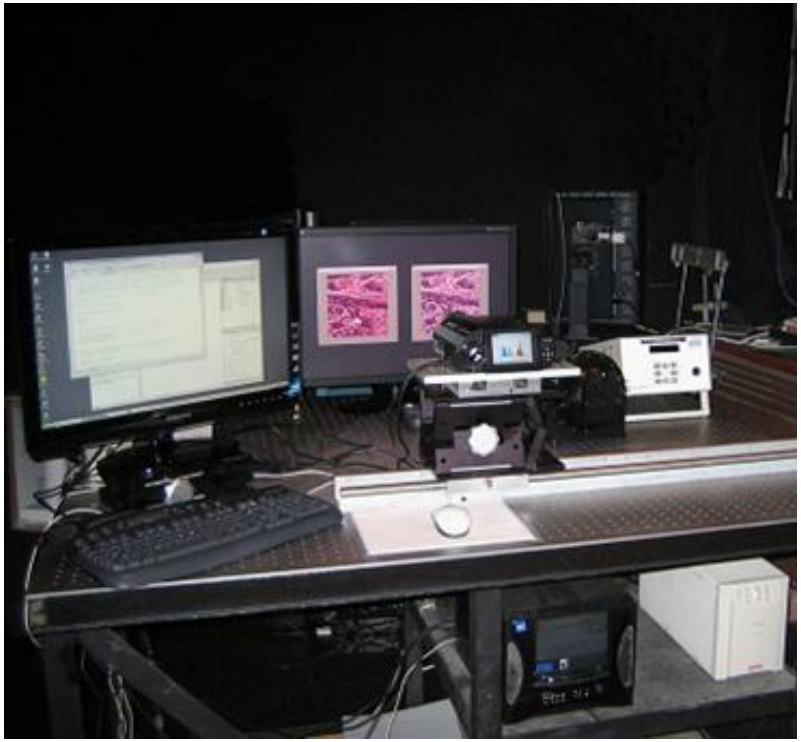


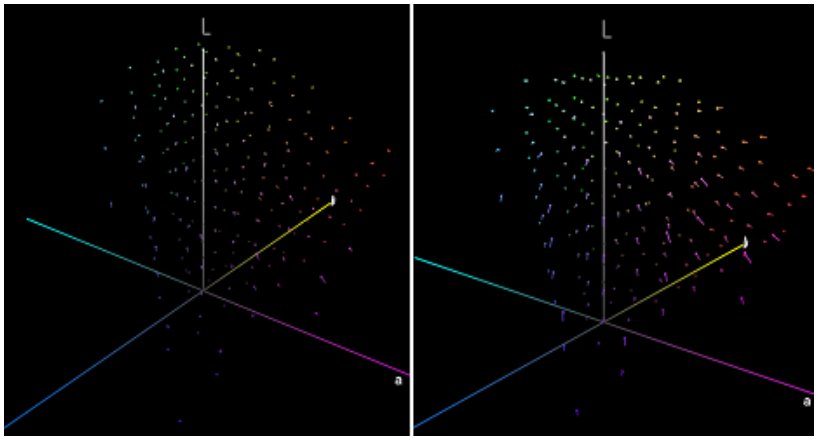
Less Common Tools

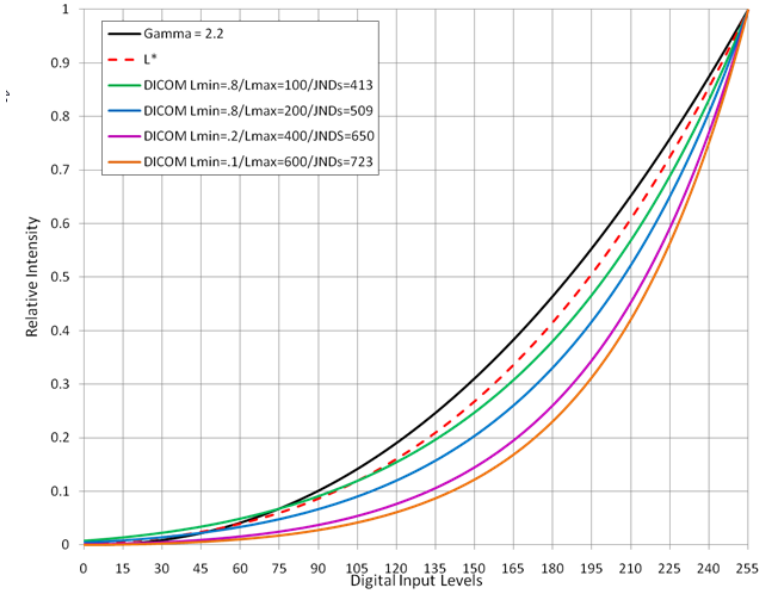




Rarely Used Methods





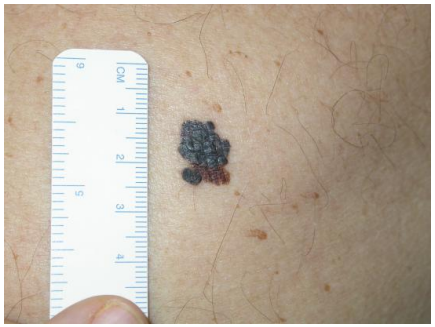


What About the Users?

- **Improved understanding human observers**
  - **Reduce errors**
  - **Develop better training tools**
  - **Improve workflow**
- **Improve image presentation**
  - **Develop better displays, calibrations, tools etc.**
  - **Intelligent compression**
  - **Tailored computer processing & decision tools**
  - **Better navigation tools**
  - **Streamlined retrieval, display & reporting tools**

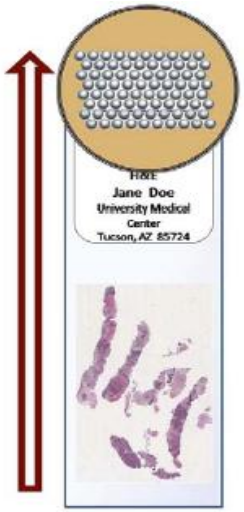
Early Studies

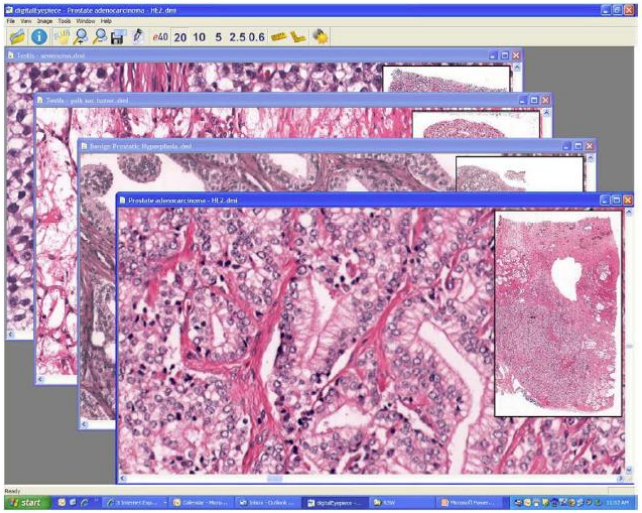
- **Teledermatology SF = In-person?**
  - **308 dermatology cases from IP**
  - **Digital photos 832 x 608, 24-bit**
  - **3 dermatologists (intra & inter)**
- **83% diagnostic concordance**
- **62% very definite or definite**
- **Image sharp 83% good/excellent**
- ***Image color* 93% good/excellent**

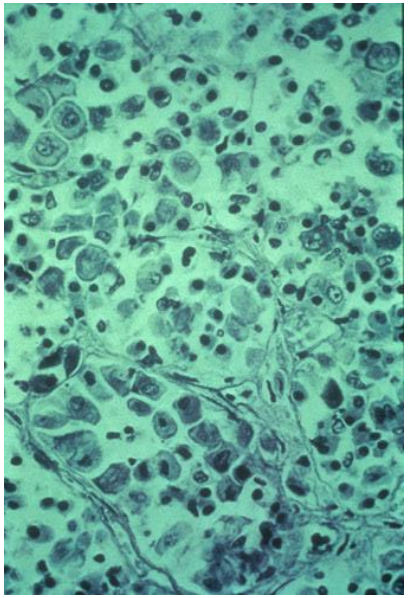


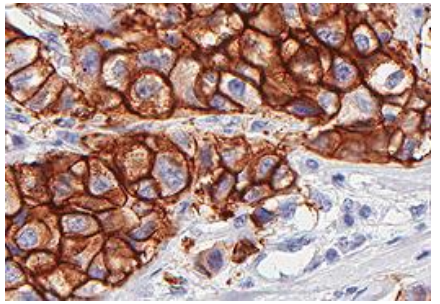
Krupinski et al. (1999). Diagnostic accuracy & image quality using digital camera for teledermatology. TMEH 5;257-263.

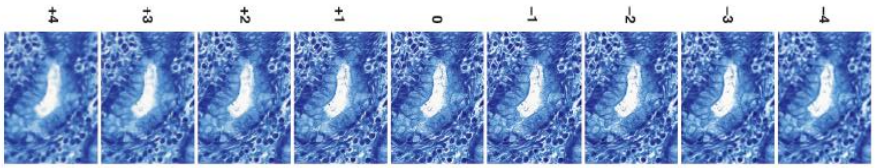
Recent Studies

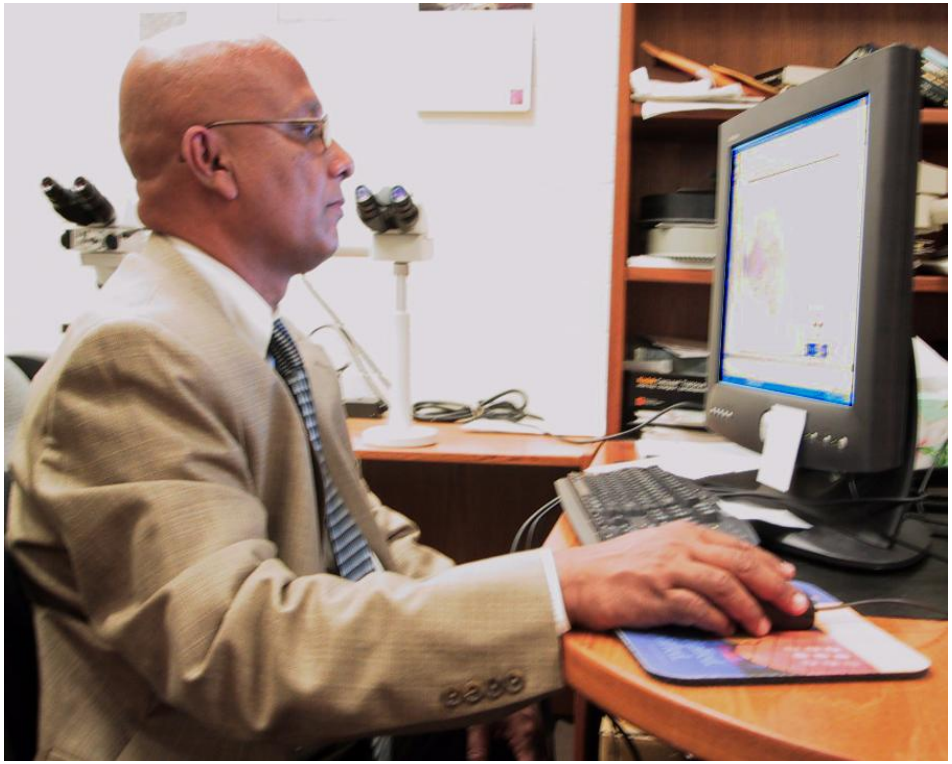










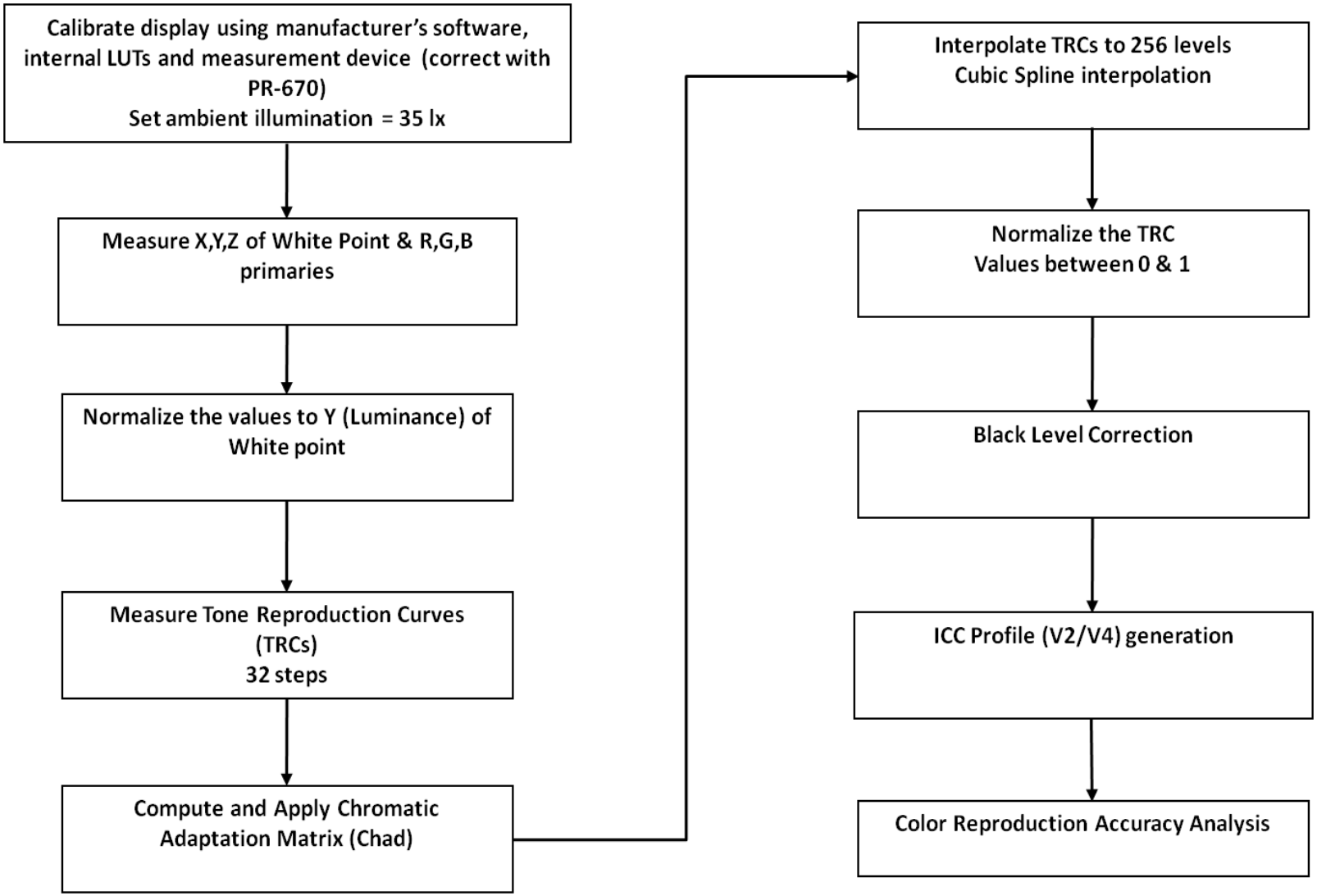


Calibration, Characterization & Profiling

Silverstein et al. Achieving High Color Reproduction Accuracy in LCDs for Color-Critical Applications. JSID 2012;20:53-62

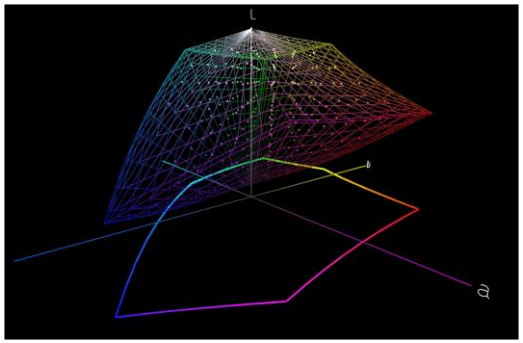
Experimental Setup

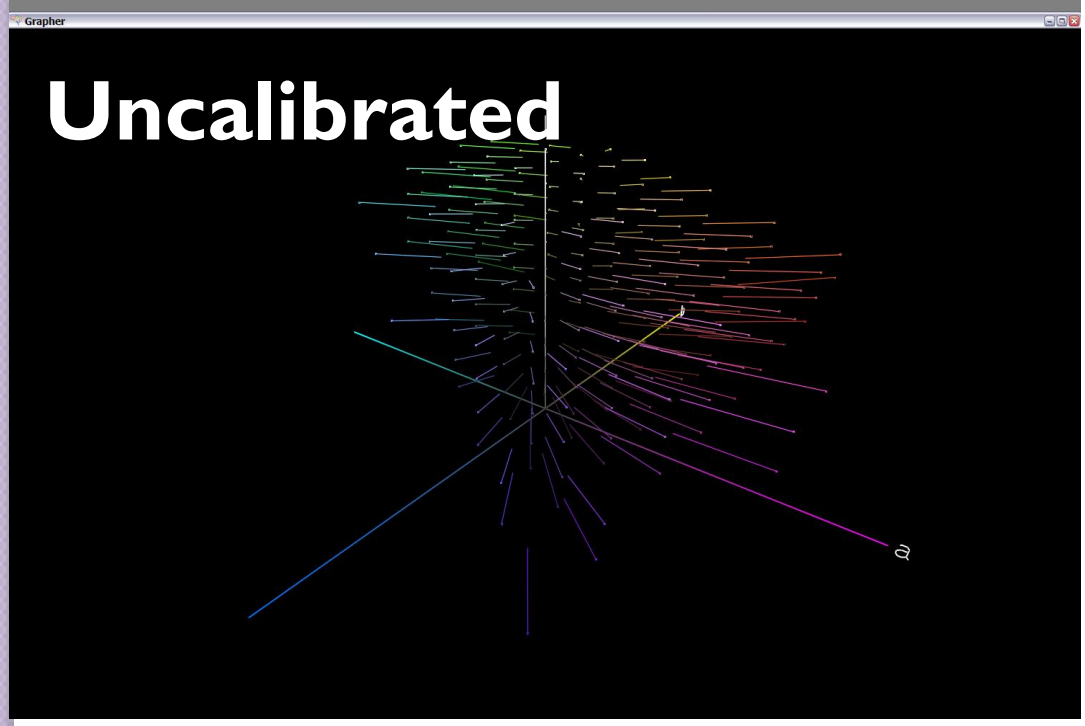
Calibration, Characterization & Profiling

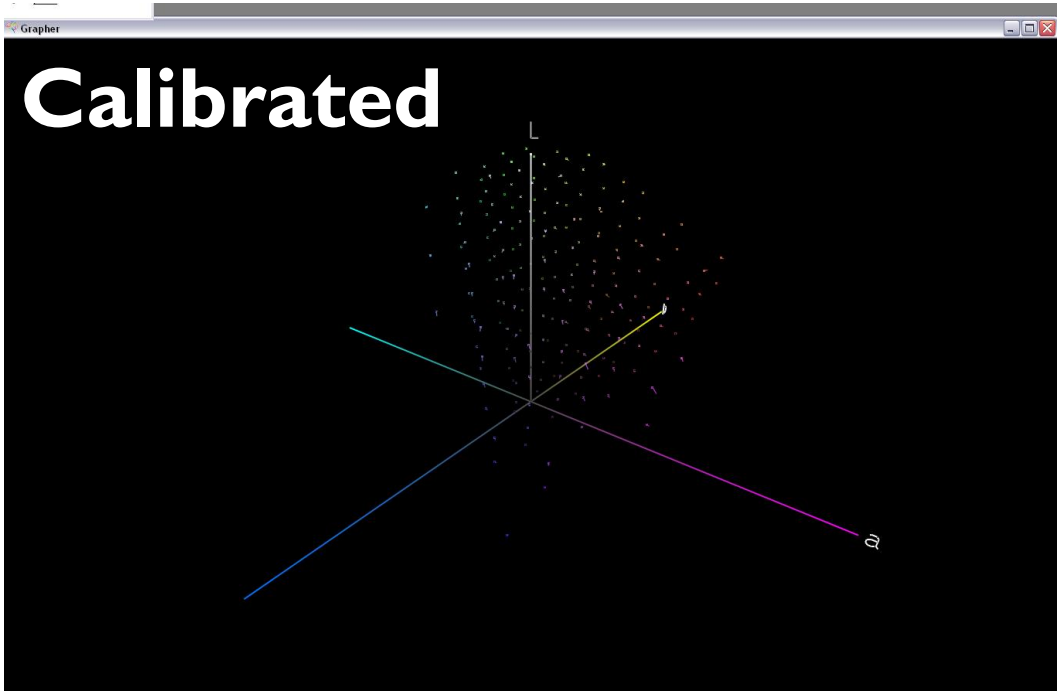


Methods

- MatLab module evaluates color reproduction accuracy using display characterization & ICC profile
- Sets test colors generated taking intersection uniform sampling CIE Lab color space axes with reference color space (sRGB)
- Set 210 test colors created

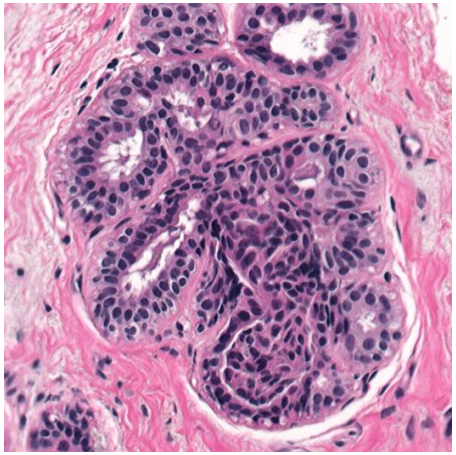


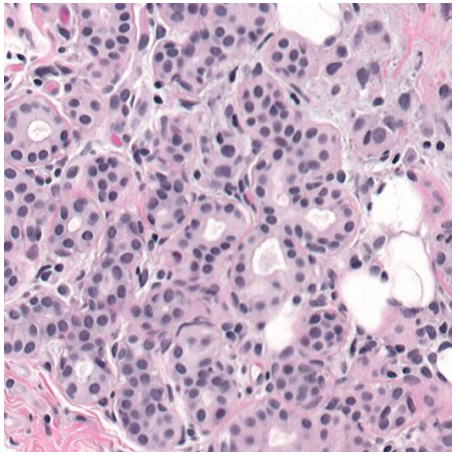

Grapher
Uncalibrated
L
Q


Grapher
Calibrated
L
Q

Does It Matter?

- **Whole slide images DMetrix scanner**
- **Breast biopsy specimens**
- **250 ROIs selected by expert pathologist**
  - **½ malignant & ½ benign**
- **Independently graded 2nd pathologist excellent or good quality**

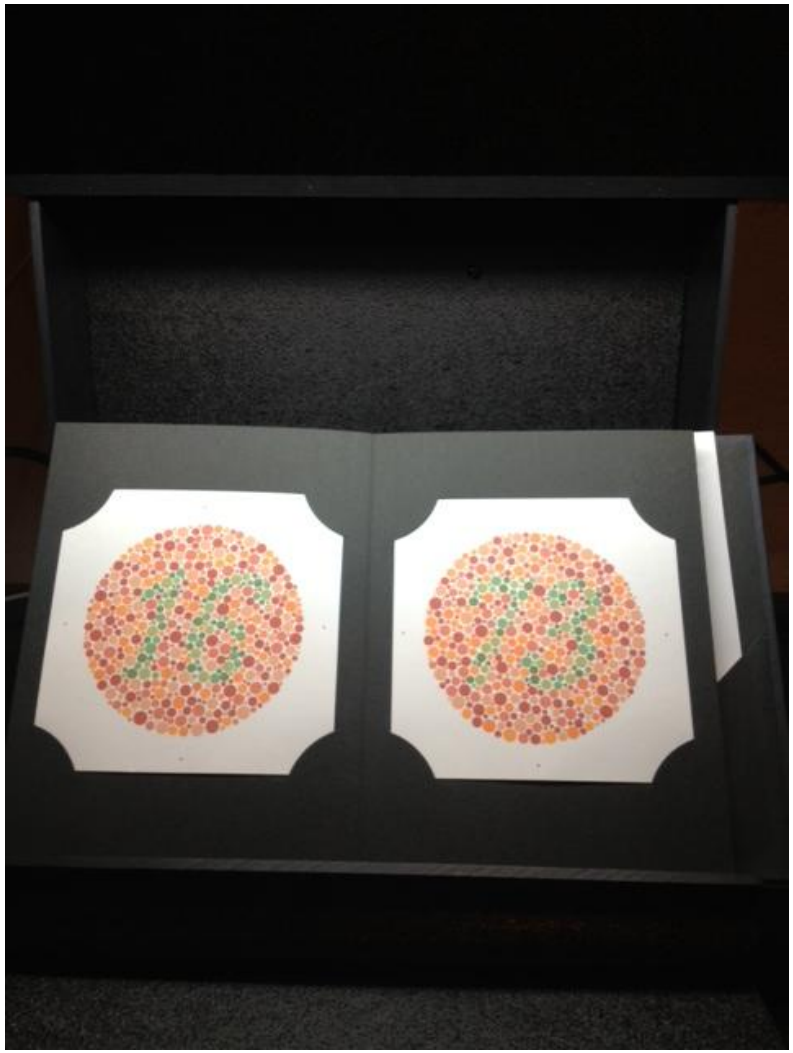




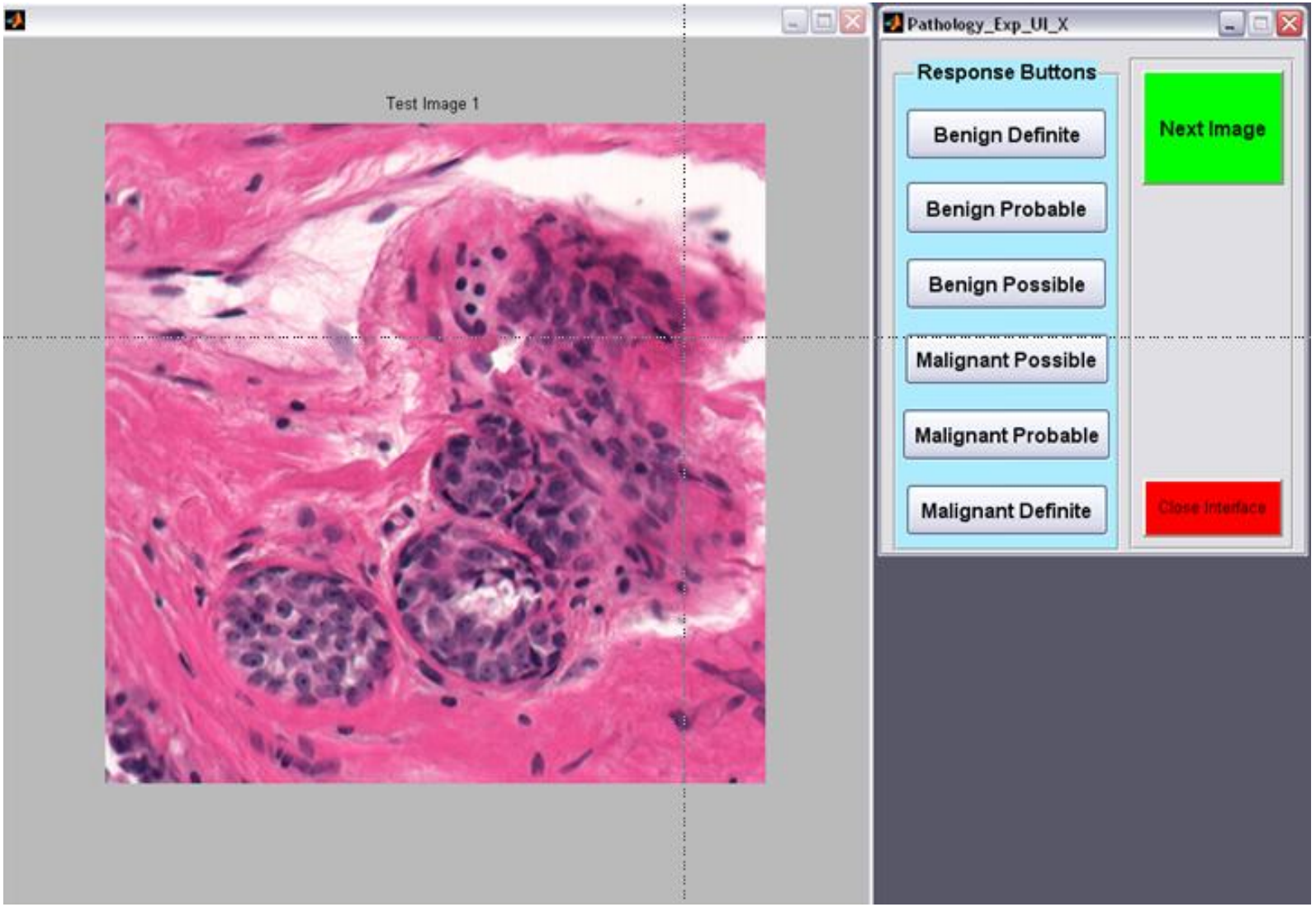
Study Methods

- 6 pathologists – 2 Board certified, 4 residents
- NEC 2690 Color LCD
  - 1920 x 1200
  - Lmax = 320 cd/m^2
  - Contrast ratio = 1000:1
  - Wide gamut
  - Calibrated/color managed & off-the-shelf
- Counterbalanced min 3 weeks between
- Rate benign vs malignant
- Trials timed automatically

Ishihara's Test Color Deficiency



Interface

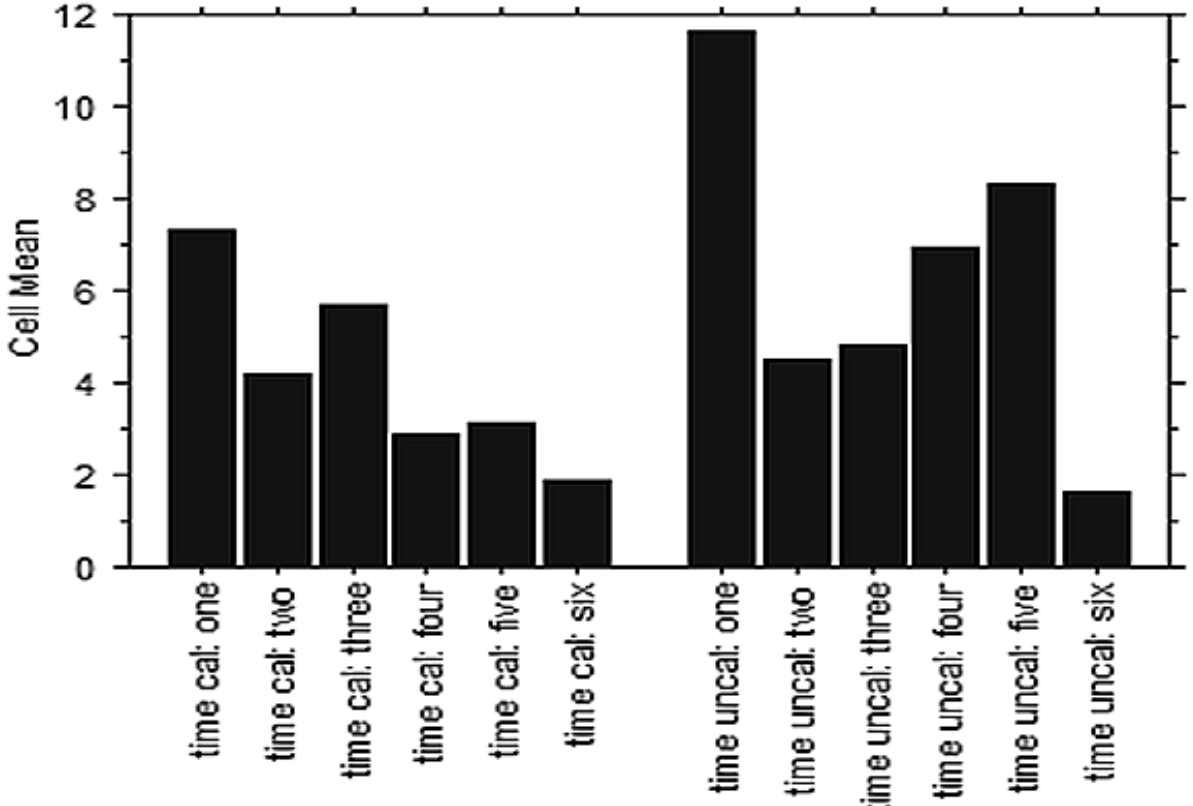

Test Image 1
Pathology_Exp_UI_X
Response Buttons
Benign Definite
Benign Probable
Benign Possible
Malignant Possible
Malignant Probable
Malignant Definite
Next Image
Close Interface

MRMC ROC Az

Reader	Uncalibrated Az	Calibrated Az
1	0.9003	0.9142
2	0.9747	0.9856
3	0.8235	0.8586
4	0.7827	0.7884
5	0.8098	0.7889
6	0.8015	0.8062
Mean	0.8488	0.8570

F = 0.71 p = 0.4112

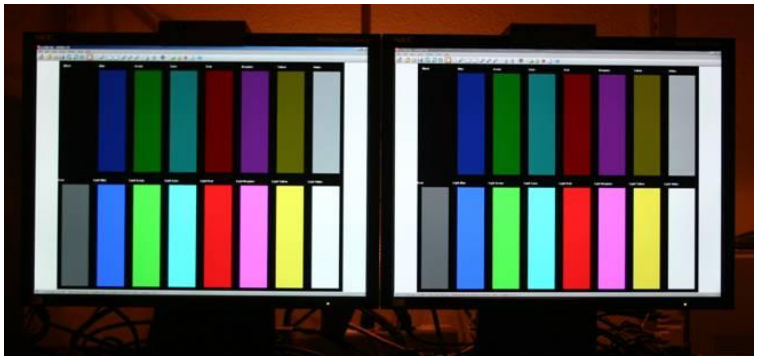
Timing Results



Average 4.895 sec vs 6.304 sec p = 0.0460

Conclusions

- 1st study examine impact calibrating/color managing displays on diagnostic accuracy & interpretation efficiency
- Diagnostic accuracy *equivalent*
- Interpretation speed significantly *slower*
- Unless viewed side-by-side do not notice differences



ATA Guidelines Derm

- *Image acquisition:* **device should allow ≥ 24 bits color; set white balance; attach calibration label (GreTag) at least 1 image (pref all)**
- *Image display:* **max/min luminance, gamma (2.2), white point, color temp 6500 K, bit-depth 24, ambient 20-40 lux; if GreTag doesn't match adjust using display settings function &/or calibration software tools**

ATA Guidelines DR

Monitors & settings should be validated for clinical diagnostic accuracy. Any validated monitor technology can be used (e.g., cathode ray tube, liquid crystal display, gas plasma panel). Retinal images used for diagnosis should be displayed on high-quality monitors of appropriate size and resolution. Displays should be calibrated regularly to ensure fidelity with original validation display conditions. Re-validation should be performed if settings are changed. Ambient light level, reflections and other artifacts should be controlled to ensure standardized viewing.

ACR Electronic Practice

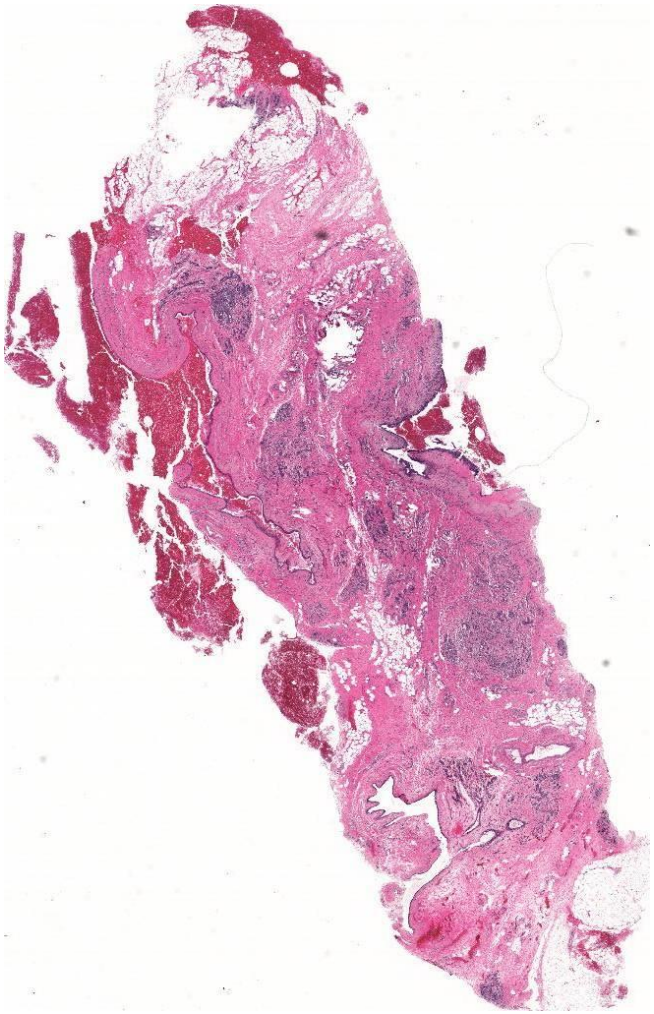
White point: The *color characteristics* of a display with respect to the presented color space are *not considered* in this technical standard. However, the white point associated with presentation of grayscale images is important for medical imaging systems. It is recommended that monitors be set to a white point corresponding to the CIE daylight standard D65 white point. This corresponds to a color temperature of about 6,500 °F.

When is Color Important?

- **Color displays (MG & COTS) quite good now**
  - **Luminance, noise, stability etc.**
  - **Cost reasonable even MG**
  - **Calibration tools exist – should be used!**
  - **Diagnostic accuracy & workflow affected little**
- **Humans vs computers!**
  - **Image processing & analysis tools require more color fidelity, calibration, color management**

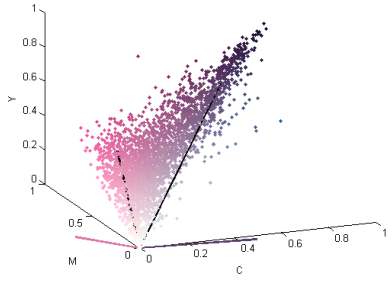
Automatic Analyses

- **Automatic tissue classification**
  - **Diagnostically Relevant (B or M) vs. Irrelevant**
    - **Application: Efficient region-of-interest compression for faster digital workflow**
  - **Benign vs. Malignant**
    - **Application: Computer-aided diagnosis**
- **Task-driven saliency**
  - **Predict regions most likely to attract *expert* attention**
    - **Application: Efficient region-of-interest compression for faster digital workflow**

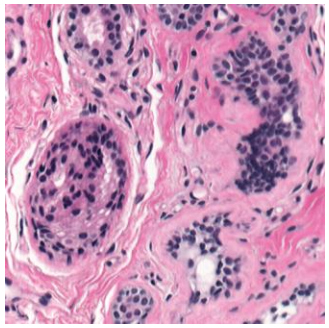


H&E Color Transformation

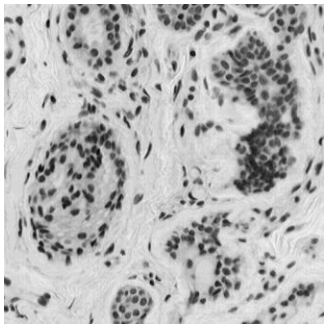
- **Isolate info on nuclei & cytoplasm**
  - **Pixelwise color transform CMY -> HE**
  - **H shows nuclei, E shows cytoplasm, etc.**



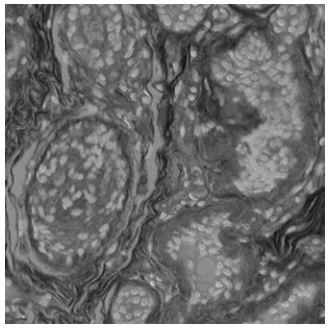
Pixel distribution CMY color space



CMY



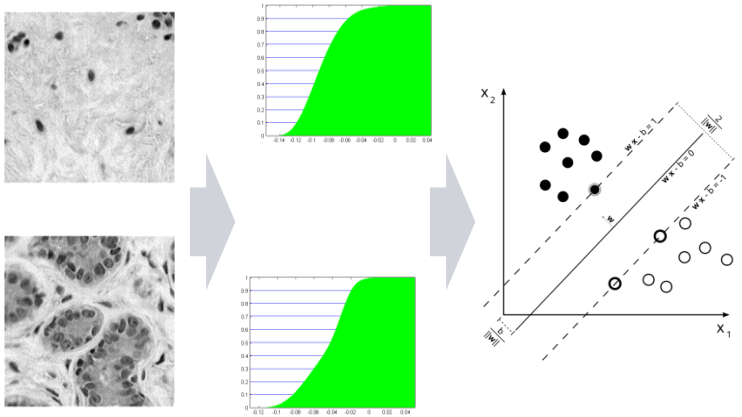
H (nuclei dark)



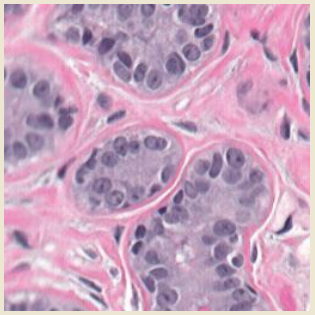
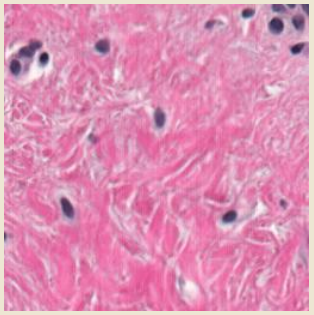
E (cytoplasm dark)

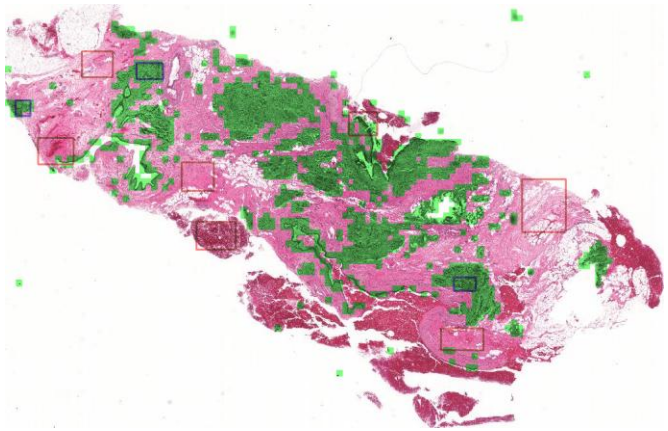
Relevant vs. Irrelevant

- **Requirements**
  - **Computationally efficient**
  - **100% detection Relevant (True Positive), certain false alarm rate permissible**
- **Methodology**



- **Performance**
  - **Accuracy: 98.8 % cross-validation EER on 400 cropped images**
  - **Speed: 10 min 4GB slide, can be sped up**

Relevant	Irrelevant
	
↑ Epithelial nuclei & Tubule formation	↓ Epithelial nuclei & Tubule formation
↓ Cytoplasmic tissue	↑ Cytoplasmic tissue



THANK YOU!
krupinski@radiology.arizona.edu