

# Evaluating Perceptual Color Reproducibility of Whole-Slide Imaging Devices

Firdous Saleheen<sup>1</sup>, Wei-Chung Cheng<sup>2</sup>

<sup>1</sup>Temple University, Philadelphia, PA; <sup>2</sup>DIDSR/OSEL/CDRH, U.S. Food and Drug Administration, Silver Spring, MD

Clinical Area: Digital Pathology

## REGULATORY RELEVANCE

Demonstrating quantitative methods for comparing color performance of a 510(K) whole-slide imaging submission against its predicate device

## ABSTRACT

The color reproducibility of two whole-slide imaging (WSI) devices was evaluated with real tissue slides. Three tissue slides – human colon, skin, and kidney – were used to test a modern and a legacy WSI devices. The color truth of the tissue slides was obtained using a multispectral imaging system. The output WSI images were compared with the color truth to calculate the color difference for each pixel. A psychophysical experiment was also conducted to measure the perceptual color reproducibility (PCR) of the same slides with 4 subjects. The results show that the mean color differences of the modern, legacy, and monochrome devices are 10.94, 22.35, and 42.74  $\Delta E_{00}$ , while their mean PCRs are 70.35%, 23.06%, and 0.91%, respectively.

## INTRODUCTION

### Motivation

- Color reproducibility is an essential factor when evaluating WSI devices for determining substantial equivalence. Color truth is required to assess color reproducibility.

### Challenge

- Color truth of biological tissues is difficult to measure due to their microscopic structures.

### Existing Methods

- Artificial color test targets were used. However, they differ from biological tissues in spectral and structural characteristics and therefore might confound evaluation results.

### Our Approaches

- Developed a multispectral imaging system to measure the color truth for each pixel
- Compare WSI devices with an optical microscope to determine perceptual color reproducibility

This study was supported by **Critical Path Initiative** and **ORISE**. The authors thank our study participants Yu-Han Cheng, Padraic Flynn, Qi Gong, and Jonathan Rosenberg as well as Katherine Shea and Rodney Rouse, CDER for providing the Aperio WSI scans.

2016 OSEL Annual Student Science Poster Exhibit, August 3, 2016

## METHOD

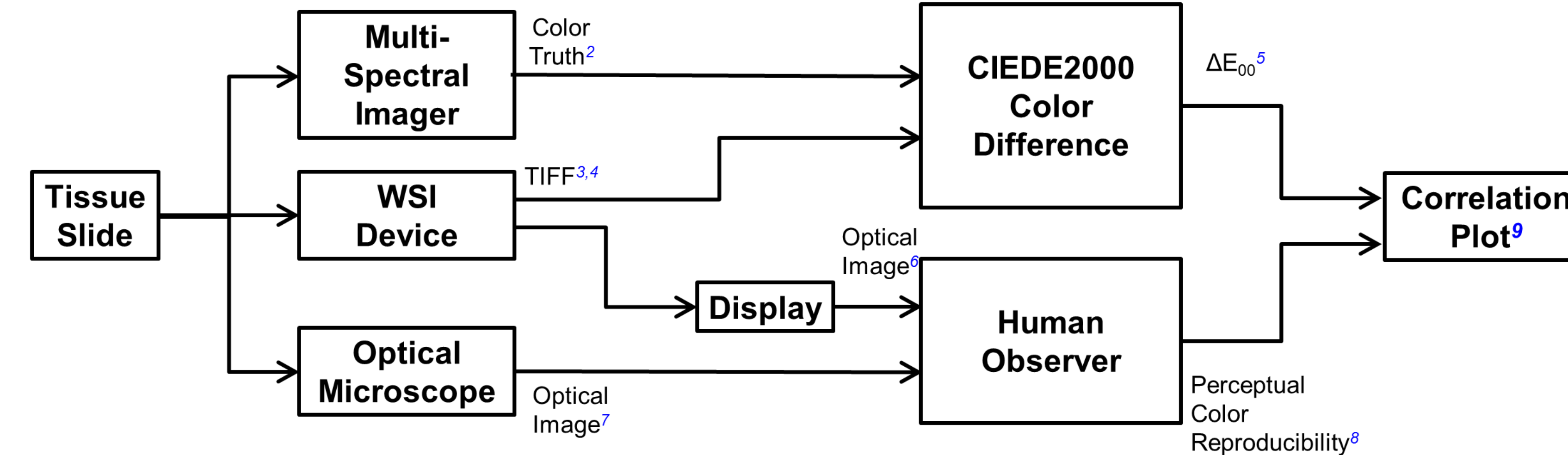


Fig. 1: Methodology workflow. The superscript numbers indicate the corresponding figures for each block.

## COLOR DIFFERENCES ( $\Delta E_{00}$ )

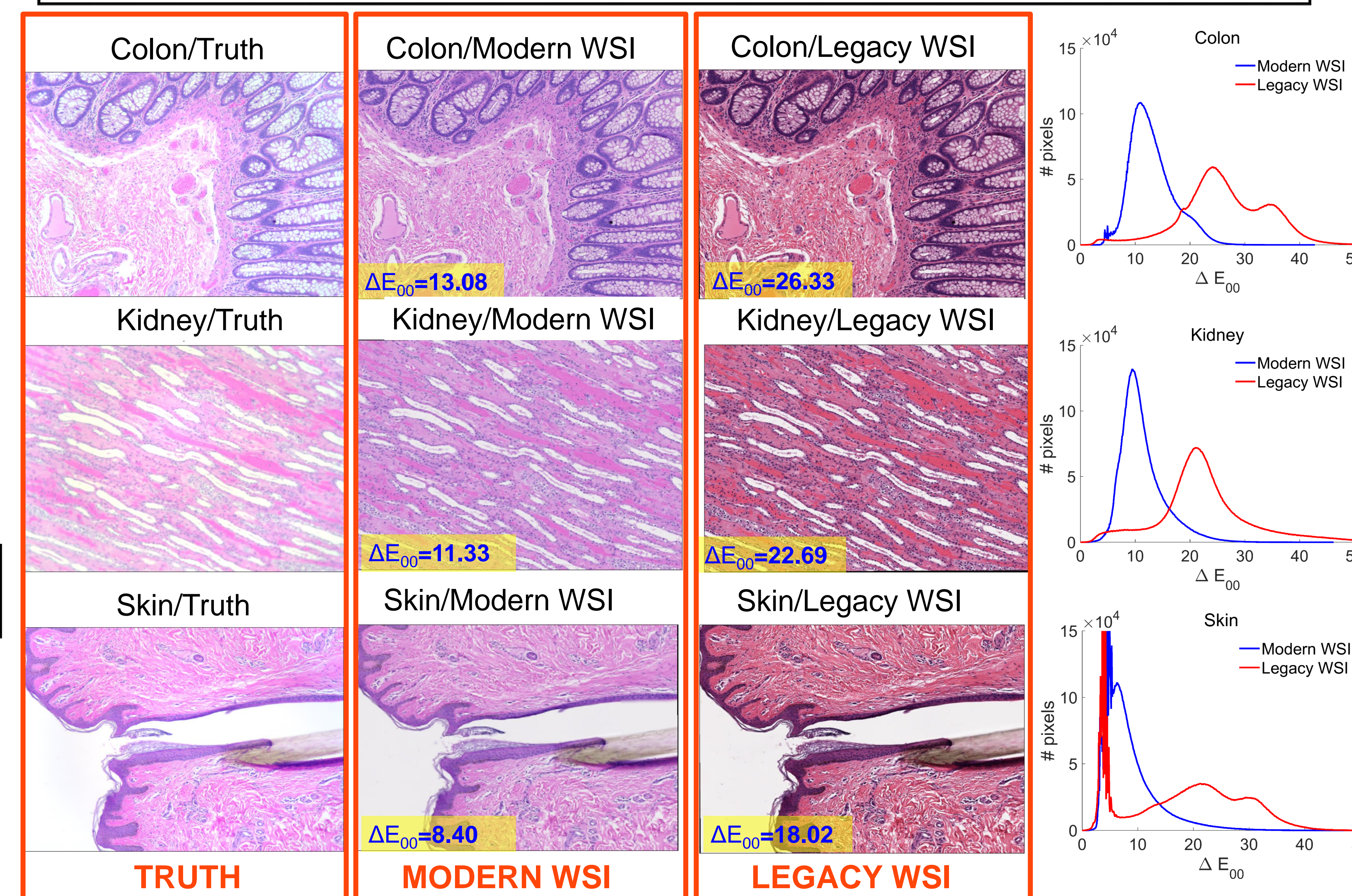


Fig. 2: Images from the color truth.

Fig. 3: Images from the modern WSI device.

Fig. 4: Images from the legacy WSI device.

Fig. 5: Histograms of  $\Delta E_{00}$  of each tissue type.

## PSYCHOPHYSICAL EXPERIMENT

- Objective: Determine perceptual color reproducibility quantitatively
- Study design: Scale PCR of 4 display images (truth, modern WSI, legacy WSI, and monochrome in Fig. 6) with respect to the microscope image (Fig. 7)
- Psychophysical method: Cross-modality ratio scaling - perceptual color reproducibility vs. perceptual length
- Subject task: For each image, place a mark on a line to represent its perceptual color reproducibility. The left end of the line represents the least PCR, while the right end the most (Fig. 8).
- 4 subjects, 6 trials per image

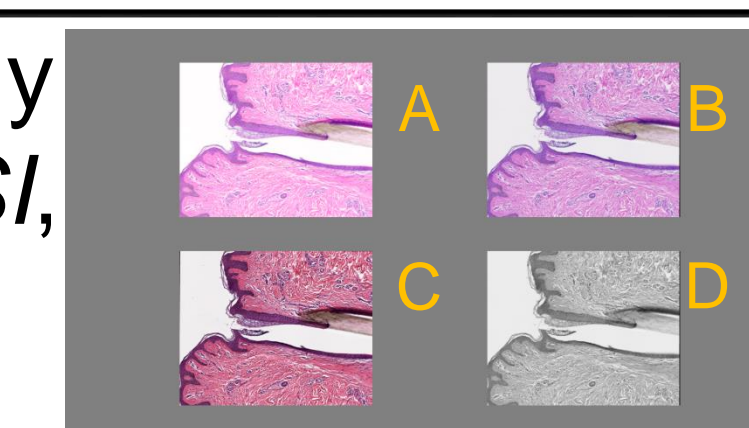


Fig. 6: Display images.

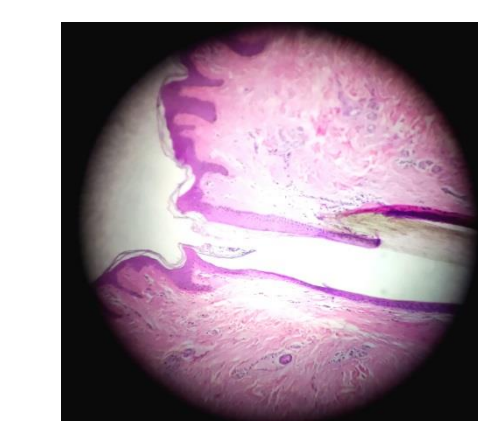


Fig. 7: Microscope image.

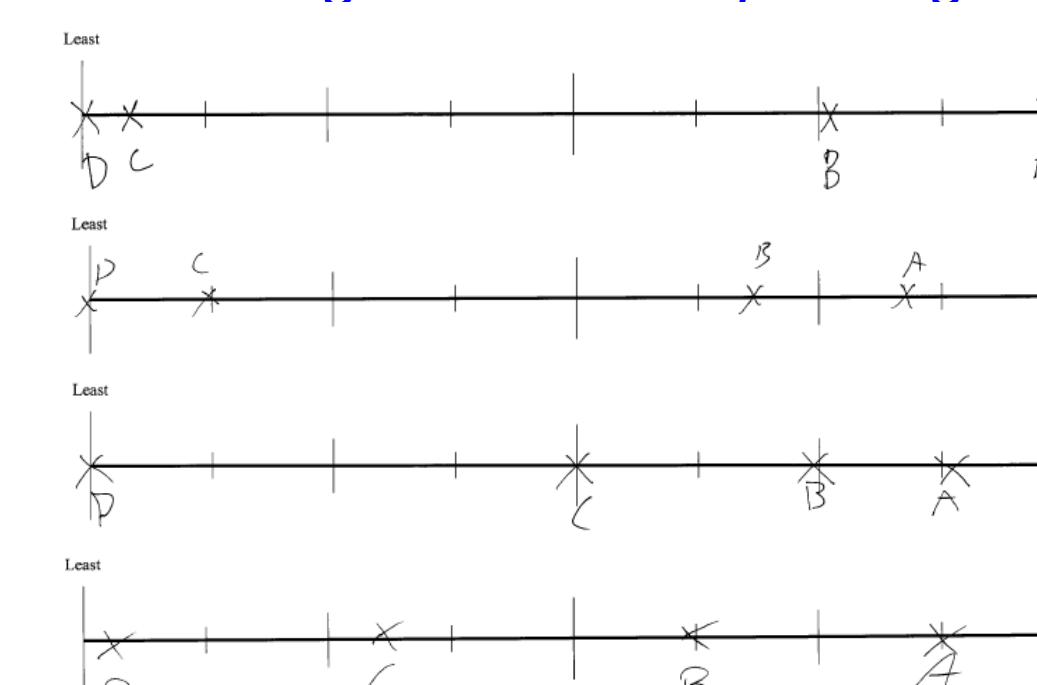


Fig. 8: Responses to the stimuli in Fig. 6 and 7 from 4 subjects.

Table 1: Perceptual Color Reproducibility

Tissue	Perceptual Color Reproducibility (%)			
	Truth	Modern	Legacy	Monochrome
Colon	89.71	69.08	20.63	0.75
Kidney	89.75	70.82	24.65	0.97
Skin	86.84	71.14	23.89	1.01
Mean	88.77	70.35	23.06	0.91

## PERCEPTUAL COLOR REPRODUCIBILITY VS. $\Delta E_{00}$

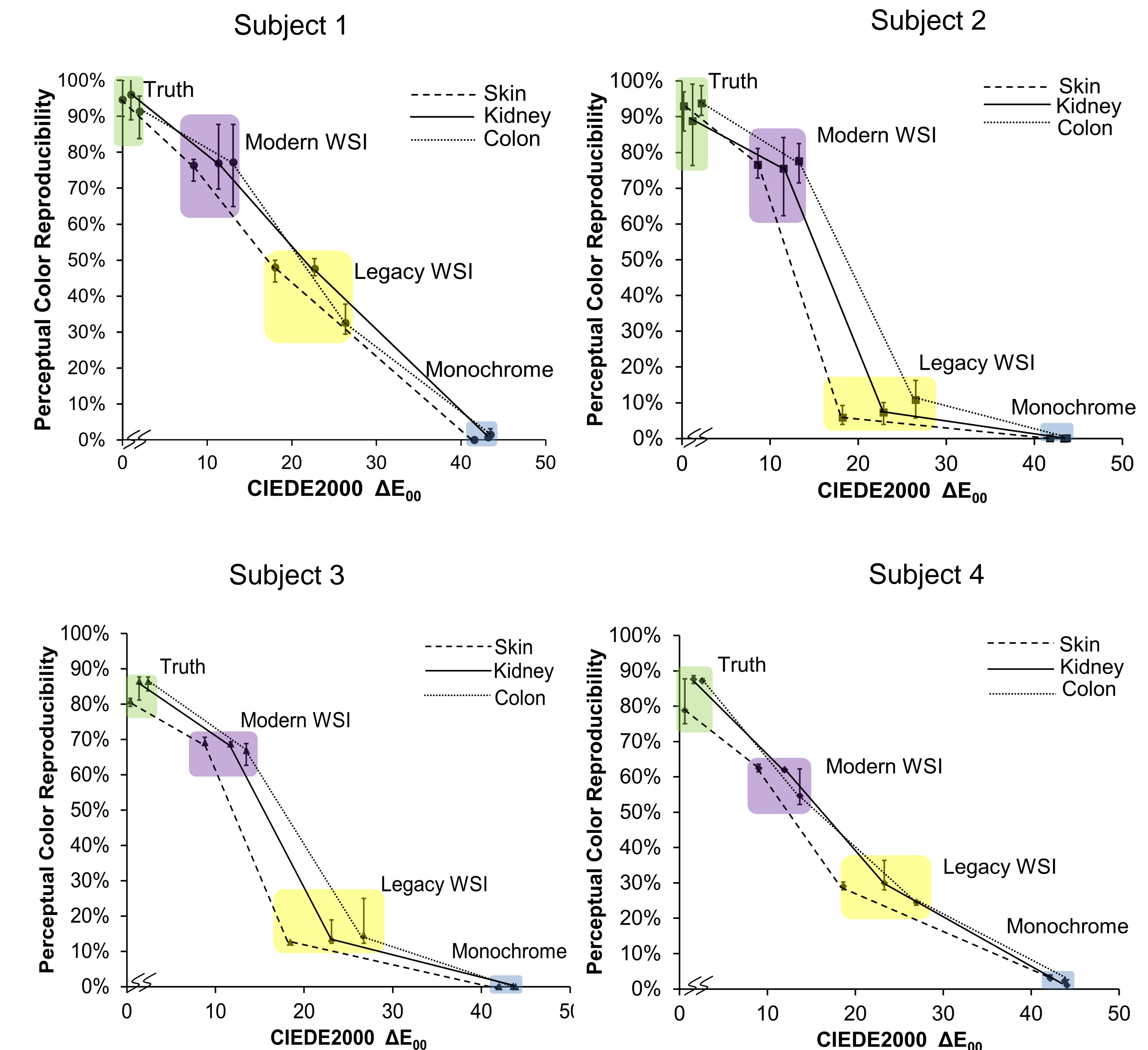


Fig. 9: PCR vs. color difference of 3 tissue types (colon, kidney, and skin) and 4 devices (truth, modern, legacy, and monochrome) from 4 subjects. Each data point represents the mean PCR of 6 trials, while the error bars represent the range. For clarity, the data points, especially the "truth" dataset, are slightly staggered in the X-axis direction.

## FINDINGS

- Monochrome images ( $\Delta E_{00}=42.74$ ) have near zero PCR (0.91%) in high consensus.
- Truth images ( $\Delta E_{00}=0$ ) were not perfectly reproduced on the display and perceived as predicted by CIEDE2000 (PCR=88.77%).
- Modern WSI has  $\Delta E_{00}=10.94$  and PCR=70.35%.
- Legacy WSI has  $\Delta E_{00}=22.35$  and PCR=23.06%.
- Both PCR and  $\Delta E_{00}$  can perfectly detect legacy vs. modern WSI devices (sensitivity=specificity=100%).
- Inter-reader variability in PCR for legacy WSI is greater than modern WSI (lower consensus).
- Subject 1 and 4 show near linear responses. Legacy WSI was lowly scaled by subject 2 and 3.

## CONCLUSION

The non-inferiority of color reproducibility in the modern WSI device was deduced using two quantitative methods..