

# Data Labeling Best Practices

## Folder Identification

Experiment name. Highest level of organization. Unique.

- A unique identifier
  - description of experiment (i.e., CTW, breast, double negative, 14MAY2021)

## Slide Identification

Once the original scan of the slide has been completed, you are asked to identify each slide.

- A unique identifier
  - barcode, unique slide number or name (i.e., tdT-TGF:23-4; MSI\_CRC67)
  - a means to refer to this specific slide
  - a description can be added for additional clarification

## ROI Selection

Highest level of tissue identification

- Identifies location of interest (biological target)
  - Stroma, germinal center, epithelium, infiltration, lumen
  - Use comment to provide description
  - Numbered sequentially (001)

## AOI Selection

Sub-region of the ROI (segments)

- Identifies specific target within your ROI
  - Immune cells, tumor cells, morphological structures
  - Each segment **MUST** be identified consistently to allow grouping between multiple slide scans. Once approved and collected, it cannot be changed or added
  - PanCK+, VEGF+, immune, CD3+, col2-, vimentin-, nuclei
  - Use comment to provide description

## Strategy

- Geometric ROIs only
  - Segmented ROIs only
  - Combination, segmented and geometric
1. Geometric ROIs only
    - a. Numbered sequentially (001)
    - b. Use comment to provide description
  2. Segmented ROIs only
    - a. Identified by segment name
    - b. Use comment to provide description
    - c. Segments generated first get first pass at claiming tissue area
    - d. Drag segments up and down to change order

- e. Consider establishing the collection order of the different segments within an ROI starting with the smaller segments (based on area or number of cells) and finishing with the largest segment within that ROI.
  - f. Strong staining is required for segmentation based on fluorescent markers
3. Combination of segmented and geometric ROIs  
 Establish segmented ROIs first, then place geometric ROIs  
 Naming conventions as described above

To Further describe your data, use the **Manage Annotation** feature to add tags. This feature allows filtering and grouping of your data.

Segment type: construct, muscle  
 ROI location: host muscle, construct, interface  
 Sample type: BMSC, EDL

Annotations								
Sample_ID	slide name	scan name	panel	roi	segment	aoi	area	tags
DSP-1001660004638-G No Template Control								
DSP-1001660004638-G		45 BMSC_17	(v1.0)	001	PanCK +	PanCK +-aoi-001	26193.75	host muscle
DSP-1001660004638-G		45 BMSC_17	(v1.0)	001	Syto 13 +	Syto 13 +-aoi-001	37964.54	host muscle
DSP-1001660004638-G		45 BMSC_17	(v1.0)	002	PanCK +	PanCK +-aoi-001	27631.01	construct
DSP-1001660004638-G		45 BMSC_17	(v1.0)	002	Syto 13 +	Syto 13 +-aoi-001	39445.83	construct
DSP-1001660004638-G		45 BMSC_17	(v1.0)	003	PanCK +	PanCK +-aoi-001	10368.50	interface
DSP-1001660004638-G		45 BMSC_17	(v1.0)	003	Syto 13 +	Syto 13 +-aoi-001	56359.47	interface
DSP-1001660004638-G		45 EDL_06	(v1.0)	001	PanCK +	PanCK +-aoi-001	10495.01	host muscle
DSP-1001660004638-G		45 EDL_06	(v1.0)	001	Syto 13 +	Syto 13 +-aoi-001	56502.83	host muscle
DSP-1001660004638-G		45 EDL_06	(v1.0)	002	PanCK +	PanCK +-aoi-001	14741.94	construct
DSP-1001660004638-G		45 EDL_06	(v1.0)	002	Syto 13 +	Syto 13 +-aoi-001	53386.89	construct
DSP-1001660004638-G		45 EDL_06	(v1.0)	003	PanCK +	PanCK +-aoi-001	22274.07	interface
DSP-1001660004638-G		45 EDL_06	(v1.0)	003	Syto 13 +	Syto 13 +-aoi-001	41501.19	interface
							396,865.03	