WORKING WITH AI · REFERENCE CARD

Session 1 — The Foundation

The Language of Seeing · keep this; you'll build on it for four sessions

The one idea: the quality of the science you get out is proportional to the richness of the biological context you put in. A vague prompt doesn't fail — it returns a confident, wrong answer that looks fine.

How every session works — talk to the model → build the analysis (code to measure or test a hypothesis) → validate it. Same arc every session; only the modality changes.

Describe every image in four dimensions

Dimension	What it covers	Thin → rich
01 Sample	Organism, tissue, cell type, model, treatment, timepoint.	"cardiomyocytes" → "adult mouse ventricular myocytes, TAC model, 16 wk"
02 Imaging	Modality, magnification, pixel size, channels, dyes.	"green channel" → "0.097 µm/px; single channel, cytosolic EGFP"
03 Structure	Shapes, sizes, spacing, expected vs. unexpected. Use analogies.	"bright spots" → "rod cells, regular ~1.8 µm banding, like rungs"
04 Question	The biological question, and what it's compared against.	"count cells" → "binucleation per cell, TAC vs. sham"

The imaging vocabulary (it travels into every session)

Pixel = a measurement — a number / intensity, not just a color

Bit depth & dynamic range — 8-bit = 0–255; 16-bit = 0–65,535 — the step size

Saturation — a pixel pinned at max is clipped; true value is gone. Check the histogram first.

Spatial calibration — µm/pixel (nm in super-res) turns pixels into real distance

Sampling — Nyquist — ≥2 pixels across the smallest detail (~2.3×). Undersample → aliasing; oversample → wasted dose & file size.

File & compression — pixel size & channels usually live in the file's metadata (.czi, .nd2, Olympus .oif) — have Claude read them, don't retype. Lossless (PNG, TIFF-LZW) is safe; never quantify on lossy JPEG.

Channels — which dye marks which structure — the basis of colocalization

Two modes of working

EXECUTION
You know what to measure.
DESCRIBE → SPECIFY → VERIFY
Rich description, exact steps + "test on one first," and state how you'll check it.

DISCOVERY
You want to know what's possible.
SHOW → WONDER → PROBE
Rich description, don't direct the analysis; ask what could be measured; chase the biology.

Discovery starter: upload with the description only — then ask "What could be quantified here that I might not have considered?" and chase what surprises you.

Setting your sampling — the calculation

SPATIAL — pixel size
Resolution $d = \lambda / 2 \cdot NA$. Pixel ≤ $d / 2$ (aim ~2.3×).
e.g. confocal GFP, NA 1.4: $d \approx 186$ nm → pixel ≤ ~93 nm.

TEMPORAL — sampling rate
Interval ≤ (feature time) ÷ (points you want).
e.g. spark rise 10 ms, ≥5 points → ≤2 ms → ≥500 lines/s.

Sampling finer than Nyquist doesn't add resolution — the optics or the kinetics set the limit; Nyquist just makes sure you capture what they deliver. **Homework before Session 2: run this on your own work — pick a structure or event you image, compute the pixel size or frame rate it needs, and have Claude check you.**

Trusting a result

The overlay is the interface. You don't need to read the code. If the overlay marks what a biologist would mark, the analysis is right; if not, it's wrong — however elegant the code looks.

A threshold is a decision, not a default. Turning intensities into objects or events means choosing a cut-off you own. Same idea recurs as a $\Delta F/F_0$ event cut, a segmentation cut, and a localization cut. Set it from the background — cutoff = mean + k×SD (k ≈ 3), not by eye — and never let it auto-adjust per image in a batch.

Every cut trades off. Too high → you miss real things (false negatives). Too low → you count noise (false positives). Signal-to-noise sets the ceiling on how well you can choose.

Known ground truth is the proof. Run your method on data where you already know the answer (synthetic, features planted in advance). Recovered vs. truth — the gap is the lesson, with no biological ambiguity.

Know the failure modes. It fails fluently — confident, plausible, sometimes inventing method names or citations. Random errors (scattered) usually sit within biological variability; systematic errors (the same mistake every time) are the ones to catch, and they're fixable with one instruction.

Before any result goes in a figure

- ☐ Tested on one image with a known expected result
- ☐ Overlay examined and biologically sensible
- ☐ 5–10 samples cross-checked against manual or an established tool
- ☐ Agreement quantified (Pearson r, Bland–Altman, or equivalent)
- ☐ Complete prompt archived with the raw data; model version recorded
- ☐ Batch run used identical parameters — no per-image auto-adjustment
- ☐ Known limitations written down

TODAY'S PORTFOLIO PIECE

Keep the four-dimension description you wrote and the first execution prompt you ran. They are the literal input to Session 2. The portfolio accumulates across all four sessions — it doesn't reset.

WHAT THE NEXT SESSIONS ASSUME YOU HAVE

Describe in four dimensions · read calibration & channels · judge sampling (Nyquist) · keep data quantitative (lossless, metadata, never JPEG) · treat a threshold as a decision · name the detection trade-off & SNR · validate against known ground truth.