

Study Design Cheat Sheet

BCB744 — Biostatistics Version 1 (4 April 2026)

Professor A.J. Smit

Design principles drawn across Chapters 1–26. Use this sheet before sampling or experimentation so that later inference rests on defensible data rather than statistical repair.

QUICK DESIGN FLOW

- 1. Question first.** Define the biological question, target population, and response variable before choosing a method.
- 2. True replicates?** Identify the sampling unit and check that planned n counts independent biological or experimental replicates.
- 3. Dependence expected?** If rows cluster in space, time, site, subject, block, or batch, design for that structure explicitly.
- 4. Measured well?** Standardise measurements, metadata, coding, and missing-value rules before collecting data.

Stage	Cue	Design Element	Decide Before Sampling	Why It Governs Later Inference	Good Practice / Implementation
QUESTION, POPULATION, AND VARIABLES					
What exactly is being studied?					
Question	Q	Biological objective	State whether the goal is description, comparison, association, explanation, or prediction.	The question determines whether you need a test of differences, a regression, a predictive workflow, or no inferential model at all.	Write one sentence in biological language first; only then translate it into statistical terms.
Question	y Resp	Response variable	Choose one primary response and define its scale, units, and admissible range.	Response type determines whether Gaussian, count, binomial, ordinal, or other models are appropriate. Poorly defined responses cause later method drift.	Record units, scale, zero meaning, transformation plans, and whether the response is continuous, binary, proportion, or count.
Question	Q	Target population and inference scope	Specify the population to which inference should generalise.	Parameters belong to populations, not to the convenience sample alone. Weak population definition weakens every later conclusion.	State where, when, and to what organisms, sites, or processes the inference is intended to apply.
What counts as one observation?					
Replication	n Unit	Sampling unit	Define the smallest independently sampled biological or experimental replicate.	This is the n in your analysis. If it is misidentified, standard errors, p -values, intervals, and effect estimates become misleading.	Name the unit explicitly: organism, plot, mesocosm, site, clutch, transect, cage, or patient.
Replication	n Unit	True replication vs technical repetition	Distinguish repeated measurements on the same unit from new independent units.	Repeated instrument reads, subsamples, and pseudo-subsamples do not create more independent evidence.	Average technical repeats when justified, or retain them only if a later grouped model is planned.
Replication	•• Grp	Grouped structure	Decide whether units are nested in site, block, tank, observer, family, or subject.	Grouped data often require blocking, pairing, random effects, or repeated-measures logic. Ignoring the grouping inflates apparent sample size.	Sketch the hierarchy before fieldwork: sample within quadrat within site within region, for example.
ALLOCATION, CONTROLS, AND COMPARABILITY					
How will units be compared fairly?					
Allocation	R Rand	Randomisation	Plan how treatments, order, or sampling locations are assigned.	Randomisation protects against systematic bias and hidden confounding. Without it, later causal language is much weaker.	Randomise treatment assignment, sampling order, or lab order where feasible; record the scheme used.
Allocation	C Ctrl	Control and comparator groups	Define the reference condition against which change or difference will be judged.	Tests of differences require a meaningful comparator, not merely multiple labels. Poor controls limit interpretation more than statistics can fix.	Specify untreated, baseline, sham, historical, or environmental controls and justify why they are informative.
Allocation	•• Grp	Blocking and pairing	Use blocks or matched pairs when known background heterogeneity would otherwise obscure the treatment effect.	Blocking improves precision by absorbing predictable nuisance variation; pairing changes the sampling unit to the within-pair difference.	Match units only on variables known before treatment; keep the block structure in the analysis plan.
Allocation	≠ Conf	Confounding avoidance	Check whether treatment, environment, batch, site, season, or observer will be entangled.	Collinearity and confounding later destabilise coefficients and make attribution impossible.	Vary or balance nuisance factors during design rather than hoping to "control for them" later.
When and where will sampling happen?					
Dependence	t Time	Temporal independence	Decide whether measurements taken through time are genuinely independent or repeated on the same units/process.	Temporal pseudoreplication and autocorrelation make rows look more informative than they are.	Spread sampling over appropriate time windows, or design explicitly for repeated measures / time series.
Dependence	\mapsto Space	Spatial independence	Consider whether nearby samples share environment, dispersal, or observer effects.	Spatial clustering violates independence and can mimic treatment or habitat effects.	Separate locations appropriately, stratify if needed, and record coordinates or site identifiers for later checking.
Dependence	•• Grp	Repeated measures / longitudinal design	If the same unit is measured before/after or across occasions, design it as such from the start.	Repeated measures are informative, but not as extra independent rows. They require subject IDs and clear visit structure.	Assign stable IDs, schedule visits consistently, and record missed visits explicitly rather than deleting them silently.
MEASUREMENT QUALITY AND DATA RECORDING					
Will the recorded data be analysable and credible?					
Measurement	M Meas	Measurement protocol	Specify instruments, calibration, observers, rounding, and units before data collection.	Unstable measurement inflates noise and can induce artefactual outliers, heteroscedasticity, or attenuated slopes.	Pilot the protocol, calibrate instruments, and train observers to use the same definitions and stopping rules.
Measurement	y Resp	Predictor quality	Decide how explanatory variables are measured and whether serious measurement error is expected.	Predictor error biases slopes and complicates causal interpretation, especially in regression.	Prefer direct measurement where possible; otherwise document proxies and their likely limitations.
Measurement	D Data	Coding and metadata	Predefine column names, factor levels, missing-value codes, date formats, and units.	Messy coding creates analysis errors, hidden row loss, and invalid model comparisons.	Use one data dictionary for the project; never mix blanks, zeros, and ad hoc symbols to mean missing.
Measurement	D Data	Missing-data rules	Decide what will count as missing, impossible, censored, or not applicable.	Ad hoc row deletion changes sample size across models and can invalidate comparisons such as AIC.	Record why values are missing and retain that reason in the dataset rather than deleting rows without trace.
ANALYSIS READINESS AND REPRODUCIBILITY					
Can someone else reconstruct the inferential path?					
Readiness	Δ Work	Planned analysis family	Identify the likely model family before data collection, even if details may change.	Thinking ahead exposes design gaps: missing IDs, absent controls, unbalanced cells, or unusable scales.	State the likely primary analysis, for example paired t -test, one-way ANOVA, mixed model, Poisson GLM, or GAM.
Readiness	Δ Work	Sample-size adequacy	Decide what constitutes biologically informative precision or detectable effect size.	Very small n gives unstable estimates; very large n can make trivial effects "significant".	Use pilot data, literature, or precision targets to justify the planned sampling effort.
Readiness	D Data	Data management and versioning	Plan where raw data, cleaned data, scripts, and outputs will live.	Reproducibility depends on traceable decisions. Lost provenance turns later inference into guesswork.	Keep raw data read-only, script cleaning steps, and version both code and derived datasets.
Readiness	Δ Work	Reporting commitments	Decide in advance what estimates, uncertainty measures, and diagnostics will be reported.	This reduces method shopping and keeps the focus on effect size, uncertainty, and design assumptions rather than p -values alone.	Plan to report estimates, confidence intervals, sample sizes, design structure, and diagnostic limitations.

SAMPLING - UNIT CHECK

- Ask: if I remove one row, what real-world replicate have I removed?
- If the answer is "nothing new" because the row is a repeat, subsample, or extra read from the same unit, it is not a new n .
- Mixed models and repeated-measures models partition dependence; they do not create new independent replicates.

COMMON DESIGN FAILURES

- Treating technical replicates as biological replicates.
- Sampling one site, tank, estuary, or individual repeatedly and calling it replication.
- Letting treatment coincide with site, season, observer, or batch.
- Choosing variables after data collection without a prior biological rationale.

MINIMUM DESIGN RECORD

- Biological question and target population.
- Sampling unit, planned n , grouping hierarchy, and dependence structure.
- Response variable definition, units, and measurement protocol.
- Allocation, controls, randomisation, blocking, and missing-data rules.

Use this sheet first to decide whether the planned study can support a defensible analysis. Use the inferential statistics cheat sheet only after the design has established a valid response, sampling unit, independence logic, and comparison structure.