22 ARRIVE — animal research

R. Heller · Date: 2026-05-08 · Last changed: 2026-10-30

ARRIVE — *Animal Research: Reporting of In Vivo Experiments* — covers preclinical animal research. The 21-item checklist is at https://arriveguidelines.org/. The templates below cover the statistical analyses most commonly required.

This page contains **50 method-specific templates**. Each is a skeleton with placeholders in italics; the prose is written so you can paste it into a manuscript and replace only the bracketed values.

22.0.1 Alignment with BWA and Bowtie

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Alignment with BWA and Bowtie** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Alignment with BWA and Bowtie yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `bwa-bowtie-alignment` — For Reviewers

22.0.2 ATAC-Seq Analysis

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **ATAC-Seq Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> ATAC-Seq Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `atac-seq-analysis` — For Reviewers

22.0.3 Batch Correction with ComBat

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Batch Correction with ComBat** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Batch Correction with ComBat yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `batch-correction-combat` — For Reviewers

22.0.4 Bulk RNA-seq Differential Expression with DESeq2

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Bulk RNA-seq Differential Expression with DESeq2** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Bulk RNA-seq Differential Expression with DESeq2 yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `bulk-rnaseq-differential-expression` — For Reviewers

22.0.5 Cell-Type Annotation

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Cell-Type Annotation** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Cell-Type Annotation yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-cell-type-annotation` — For Reviewers

22.0.6 ChIP-Seq Analysis

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **ChIP-Seq Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> ChIP-Seq Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `chip-seq-analysis` — For Reviewers

22.0.7 Copy Number Variation Analysis

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Copy Number Variation Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Copy Number Variation Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `cnv-analysis` — For Reviewers

22.0.8 Counting Reads with featureCounts

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Counting Reads with featureCounts** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Counting Reads with featureCounts yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `counting-reads-feature-counts` — For Reviewers

22.0.9 Differential Expression with edgeR

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Differential Expression with edgeR** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Differential Expression with edgeR yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `edger-differential-expression` — For Reviewers

22.0.10 Differential Expression with limma-voom

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Differential Expression with limma-voom** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Differential Expression with limma-voom yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `limma-voom` — For Reviewers

22.0.11 DNA Methylation Analysis

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **DNA Methylation Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> DNA Methylation Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `methylation-analysis` — For Reviewers

22.0.12 Drug-Target Interaction Mining

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Drug-Target Interaction Mining** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Drug-Target Interaction Mining yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `drug-target-interaction` — For Reviewers

22.0.13 FASTQ Quality Control

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **FASTQ Quality Control** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> FASTQ Quality Control yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `fastq-quality-control` — For Reviewers

22.0.14 Finding Marker Genes

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Finding Marker Genes** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Finding Marker Genes yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-marker-genes` — For Reviewers

22.0.15 Gene Annotation with biomaRt

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Gene Annotation with biomaRt** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Gene Annotation with biomaRt yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `biomart-annotation` — For Reviewers

22.0.16 Gene Ontology Enrichment

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Gene Ontology Enrichment** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Gene Ontology Enrichment yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `gene-ontology-enrichment` — For Reviewers

22.0.17 GSEA Preranked Analysis

Methods paragraph (copy-paste, fill in italics):

[N=] participants/observations were analysed using **GSEA Preranked Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

GSEA Preranked Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `gsea-preranked` — For Reviewers

22.0.18 GSVA Single-Sample Enrichment

Methods paragraph (copy-paste, fill in italics):

[N=] participants/observations were analysed using **GSVA Single-Sample Enrichment** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

GSVA Single-Sample Enrichment yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `gsva-single-sample` — For Reviewers

22.0.19 Heatmaps for RNA-seq

Methods paragraph (copy-paste, fill in italics):

[N=] participants/observations were analysed using **Heatmaps for RNA-seq** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

Heatmaps for RNA-seq yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `heatmaps-rnaseq` — For Reviewers

22.0.20 Integration with Harmony

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Integration with Harmony** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Integration with Harmony yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-integration-harmony` — For Reviewers

22.0.21 KEGG Pathway Enrichment

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **KEGG Pathway Enrichment** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> KEGG Pathway Enrichment yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `pathway-enrichment-kegg` — For Reviewers

22.0.22 MA Plots

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **MA Plots** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> MA Plots yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `ma-plots` — For Reviewers

22.0.23 Metagenomic Profiling

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Metagenomic Profiling** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Metagenomic Profiling yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `metagenomic-profiling` — For Reviewers

22.0.24 Microbiome Analysis with DADA2

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Microbiome Analysis with DADA2** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Microbiome Analysis with DADA2 yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `microbiome-dada2` — For Reviewers

22.0.25 Microbiome Diversity Metrics

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Microbiome Diversity Metrics** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Microbiome Diversity Metrics yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `microbiome-diversity` — For Reviewers

22.0.26 Multi-Omics Integration

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Multi-Omics Integration** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Multi-Omics Integration yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `multiomics-integration` — For Reviewers

22.0.27 Multiple Sequence Alignment

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Multiple Sequence Alignment** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Multiple Sequence Alignment yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `msa-biostrings` — For Reviewers

22.0.28 PCA of RNA-seq Samples

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **PCA of RNA-seq Samples** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> PCA of RNA-seq Samples yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `pca-of-rnaseq` — For Reviewers

22.0.29 Phylogenetic Trees with ape

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Phylogenetic Trees with ape** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Phylogenetic Trees with ape yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `phylogenetic-trees-ape` — For Reviewers

22.0.30 Population Genetics Basics

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Population Genetics Basics** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Population Genetics Basics yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `population-genetics-basics` — For Reviewers

22.0.31 Protein Structure Prediction

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Protein Structure Prediction** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Protein Structure Prediction yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `protein-structure-prediction` — For Reviewers

22.0.32 Proteomics with MSstats

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Proteomics with MSstats** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Proteomics with MSstats yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `proteomics-msstats` — For Reviewers

22.0.33 Pseudoalignment with Salmon and Kallisto

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Pseudoalignment with Salmon and Kallisto** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Pseudoalignment with Salmon and Kallisto yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `salmon-kallisto-pseudoalignment` — For Reviewers

22.0.34 Read Trimming and Adapter Removal

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Read Trimming and Adapter Removal** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Read Trimming and Adapter Removal yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `read-trimming-adapters` — For Reviewers

22.0.35 RNA-seq Normalisation

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **RNA-seq Normalisation** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> RNA-seq Normalisation yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `rnaseq-normalization` — For Reviewers

22.0.36 scRNA Clustering

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **scRNA Clustering** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> scRNA Clustering yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-clustering` — For Reviewers

22.0.37 scRNA Normalisation

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **scRNA Normalisation** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> scRNA Normalisation yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-normalization` — For Reviewers

22.0.38 scRNA QC and Filtering

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **scRNA QC and Filtering** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> scRNA QC and Filtering yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-qc-filtering` — For Reviewers

22.0.39 Sequence Alignment: Overview

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Sequence Alignment: Overview** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Sequence Alignment: Overview yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `sequence-alignment-overview` — For Reviewers

22.0.40 Single-Cell with Seurat

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Single-Cell with Seurat** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Single-Cell with Seurat yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `single-cell-seurat-intro` — For Reviewers

22.0.41 Spatial Transcriptomics

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Spatial Transcriptomics** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Spatial Transcriptomics yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `spatial-transcriptomics` — For Reviewers

22.0.42 STAR: Spliced Alignment

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **STAR: Spliced Alignment** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> STAR: Spliced Alignment yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `star-spliced-alignment` — For Reviewers

22.0.43 Structural Variant Detection

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Structural Variant Detection** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Structural Variant Detection yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `structural-variant-detection` — For Reviewers

22.0.44 Surrogate Variable Analysis (SVA)

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Surrogate Variable Analysis (SVA)** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Surrogate Variable Analysis (SVA) yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `surrogate-variable-analysis-sva` — For Reviewers

22.0.45 Trajectory Analysis

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Trajectory Analysis** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Trajectory Analysis yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `scrna-trajectory-analysis` — For Reviewers

22.0.46 Transcript-to-Gene Summarisation

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Transcript-to-Gene Summarisation** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Transcript-to-Gene Summarisation yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `tx-to-gene-summarization` — For Reviewers

22.0.47 Variant Annotation with VEP

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Variant Annotation with VEP** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Variant Annotation with VEP yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `annotation-variant-vep` — For Reviewers

22.0.48 Variant Calling with GATK

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Variant Calling with GATK** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Variant Calling with GATK yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `variant-calling-gatk` — For Reviewers

22.0.49 VCF Manipulation

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **VCF Manipulation** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> VCF Manipulation yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `vcf-manipulation` — For Reviewers

22.0.50 Volcano Plots

Methods paragraph (copy-paste, fill in italics):

> *[N=]* participants/observations were analysed using **Volcano Plots** as implemented in *[package vX.Y]* in R *[vX.Y]*. Assumptions were checked: *[list checks performed]*. *[Effect size with 95% confidence interval reported as the primary outcome]*. Two-sided p-values use *α = 0.05*; *[multiple-comparison adjustment if any]*.

Results paragraph (copy-paste):

> Volcano Plots yielded *[estimate]* (95% CI *[lower, upper]*, p = *[value]*). *[Diagnostic / assumption summary]*. *[Sensitivity analysis result]*.

Reviewer checklist for this method. → `volcano-plots` — For Reviewers

22.1 Week 4, Session 5 — Explanation vs prediction; reporting

Workflow lab: Goal → Approach → Execution → Check → Report.

22.1.1 Learning objectives

- Distinguish explanatory from predictive modelling, and choose the correct evaluation metric for each.
- Map STROBE, TRIPOD, STARD, and CONSORT onto the study designs each is meant for.
- Produce a publication-ready regression table with `gtsummary` whose numbers trace directly to the fitted model.

22.1.2 Prerequisites

The regression tools covered earlier in the book.

22.1.3 Background

Breiman's 1998 essay *The Two Cultures* and Shmueli's 2010 essay *To Explain or to Predict?* draw the same line from opposite sides. In explanatory modelling the goal is inference — what is the estimated effect of X on Y, adjusted for confounders? — and the right evaluation metric is bias-unbiasedness, interval coverage, and interpretability. In predictive modelling the goal is accuracy on unseen data, and the right evaluation metric is out-of-sample loss, calibration, and decision utility. The statistical machinery overlaps but the habits around it do not; a regression that makes an excellent explanation often makes a lacklustre prediction, and vice versa.

Reporting guidelines are the mechanism by which the field enforces discipline around these two cultures. STROBE for observational studies, CONSORT for randomised trials, STARD for diagnostic accuracy, and TRIPOD (now TRIPOD-AI) for prediction models each specify the items a reader needs to evaluate the claim. A checklist filled in as you write is much easier than one filled in at submission.

22.1.4 Setup

```
library(tidyverse)
library(gtsummary)
library(broom)
library(palmerpenguins)
set.seed(42)
theme_set(theme_minimal(base_size = 12))
```

22.1.5 1. Goal

Fit a single linear model to the penguins data and present it two ways: once as an explanatory analysis (what drives body mass?) and once as a predictive model (can we predict body mass on held-out birds?).

22.1.6 2. Approach

```
peng ← penguins ▷ drop_na(body_mass_g, flipper_length_mm, sex, species)
fit_expl ← lm(body_mass_g ~ flipper_length_mm + sex + species, data = peng)
```

22.1.7 3. Execution

```
peng |>
  ggplot(aes(flipper_length_mm, body_mass_g, colour = species)) +
  geom_point(alpha = 0.6) +
  geom_smooth(method = "lm", se = FALSE, linewidth = 0.8) +
  labs(x = "Flipper length (mm)", y = "Body mass (g)")
```

22.1.8 4. Check

Explanatory evaluation: effect sizes with intervals, residual diagnostics, and adjusted R^2.

```
glance(fit_expl) |> select(r.squared, adj.r.squared, sigma, df, df.residual)
```

Predictive evaluation: a 5-fold split, hand-coded to keep the example transparent.

```
k <- 5
folds <- sample(rep(seq_len(k), length.out = nrow(peng)))
rmse_k <- sapply(seq_len(k), function(i) {
  tr <- peng[folds != i, ]
  te <- peng[folds == i, ]
  f  <- lm(body_mass_g ~ flipper_length_mm + sex + species, data = tr)
  sqrt(mean((predict(f, te) - te$body_mass_g)^2))
})
mean_rmse <- mean(rmse_k)
mean_rmse
```

22.1.9 5. Report

```
tbl_regression(fit_expl, intercept = TRUE) |>
  modify_caption("**Table 1. Linear-regression estimates for body mass (g).**")
```

> In the Palmer penguins dataset (n = `nrow(peng)`), body mass was associated with flipper length, sex, and species (adjusted R^2 = `round(summary(fit_expl)$adj.r.squared, 2)`). Out-of-sample performance, estimated by 5-fold cross-validation, was RMSE = `round(mean_rmse, 0)` g.

22.1.9.1 Reporting-guideline map

Design	Guideline	URL
Randomised trial	CONSORT	https://www.consort-statement.org/
Observational study	STROBE	https://www.strobe-statement.org/
Diagnostic-accuracy study	STARD	https://www.equator-network.org/reporting-guidelines/stard/
Prediction-model study	TRIPOD / TRIPOD-AI	https://www.tripod-statement.org/
Systematic review	PRISMA	http://prisma-statement.org/

22.1.10 Common pitfalls

- Choosing a predictor set by R^2 and reporting a predictive claim, or tuning a predictive model and reporting causal-sounding coefficients.
- Filling in a reporting checklist at submission rather than while drafting.
- Assuming in-sample R^2 generalises.

22.1.11 Further reading

- Breiman L. (2001). *Statistical Modeling: The Two Cultures.*
- Shmueli G. (2010). *To Explain or to Predict?* Statistical Science.

- Writing a report.

22.1.12 Session info

```
sessionInfo()
```