Table 8 Generic Recipe for data analysis with the General Linear Model (normal errors).

1. Construct model. Begin with verbal and graphical model.
   Distinguish response from explanatory variables
   Assign symbols, state units and type of measurement scale for each.
   Write out statistical model.
2. Execute model Place data in model format, code model statement.
   Compute fitted values from parameter estimates.
   Compute residuals and plot against fitted values.
3. Evaluate the model, using residuals.
   If straight line inappropriate, revise the model (back to step 1).
   If errors not homogeneous, consider using generalized linear model (step 1)
   If n small, evaluate assumptions for using chisquare, t, or F distribution.
   residuals homogeneous ? (residual versus fit plot)
   residuals independent ? (plot residuals versus residuals at lag 1)
   residuals normal ? (histogram of residuals, quantile or normal score plot)
   If not met, empirical distribution (by randomization) may be necessary
4. State population and whether the sample is representative
5. Decide on mode of inference. Is hypothesis testing appropriate?
   If yes step 6, otherwise, skip to step 10.
6. State H_0/H_A pair (some analyses may require several pairs).
   State test statistic, its distribution (t or F), and tolerance of Type I error.
7. ANOVA: Partition df and SS according to model.
   Table Source, SS, df, MS, F-ratio.
   Type I error (p-value) from distribution(F or t).
8. Recompute p-value if necessary.
   If assumptions not met compute better p-value by randomization if:
   sample small ($n < 30$) and if p near α.
9. Declare decision about model terms: If $p < \alpha$ then reject H_o and accept H_A
   If $p \geq \alpha$ then accept H_o and reject H_A
   Report conclusion with evidence: Either the ANOVA table or
   F-ratio (df1,df2) or t-statistics (df) and p-value (not α) for terms of interest.
10. Report and interpret parameters of biological interest (means, slopes)
    along with one measure of uncertainty (st. error, st. dev., or conf. intervals).
    Use appropriate distribution (step 8) to compute confidence limits.

Notes on Generic Recipe

1. **Construct model**
   Define variables with symbols, type of measurement scale, and units if possible
   Response variable is... Explanatory variable is ...
   Graph response relative to explanatory variables.
   Write formal model (GLM) with graph sketched over individual terms.
2. **Execute model**. Place data in model format, one column per variable.
   Computer routines use model statement to estimate parameters and to partition df and SS according to model. Routines produce ANOVA table with p-values, fitted values, residuals, and plots of residuals.
3. **Evaluate model.**
   a. If regression used but straight line inappropriate (bowls or arches), revise the model (back to step 3).
   b. If errors expand (cones in residual plot) consider generalized linear model
   c. If n is small, evaluate assumptions for using chisquare (or t or F) distributions. Graphical analysis because tatistical tests of assumptions lack power at small sample sizes (violations matter) while being too sensitive at large sample sizes (when violations do not matter).
      Examine residual plots for heterogeneity, non-independence, non- normality. If present, randomized distributions may be needed (below).
4. **State sample, population**. Statistical inference is restricted to the population for which the sample is representative.
5. **Hypothesis testing ?** Hypothesis testing not mandatory.
   Appropriate to binary (yes/no) research questions.
   Required in some contexts (*e.g.* journal policy, high cost of Type I and II errors).
6. **State Ho/HA pair** about parameters (t-tests) or about variance (F-tests).
   State test statistic, its distribution (t or F), and tolerance of Type I error.
7. **ANOVA**: GLM routines produce ANOVA table and Type I error from appropriate distribution (t or F).
8. **Re-compute p-values** by randomization if assumptions are not met if sample small ($n < 30$), and p near α.
9. **Declare decision about model terms.**
   If H_o accepted, decide whether Type II error matters. If so report an appropriate measure such as minimum detectable difference, minimum sample size.
   If interaction term significant, then decisions about component terms are not possible.
   Report conclusion with evidence.
10. **Examine parameters of biological interest**.
    For simple models, report parameters with confidence intervals.
    These are more informative (which hypotheses excluded?) than hypothesis testing.
    One-way ANOVA: Planned (*a priori*) comparisons are more informative than unplanned (*a posteriori)* comparisons.
    If interaction term significant report contrasts (*e.g.* compare slopes).