All authors who are presenting data and data analyses in their manuscripts submitted to the Journal are now required to attest via Editorial Manager that they have reviewed sections 4A and 4B below and have implemented all of the relevant items.

This should be done preferably before implementing their study data collection but certainly as they undertook their statistical analyses and prepared their manuscript for initial submission and any requested revision(s).

While *Anesthesia & Analgesia* has elected not to implement a required formal statistical checklist to be completed and submitted by authors, adhering to the guidelines below will avoid delays in the review process and generally improve the likelihood of publication.

**A. Statistical Analyses and Methods as Promulgated by the Statistical Analyses and Methods in the Published Literature (SAMPL) Guidelines**


**B. For All Studies That Include Data Analysis and/or Estimation**

**BASIC STATISTICAL METHODS AND REPORTING THAT SHOULD BE INCLUDED IN ALL QUANTITATIVE MANUSCRIPTS.**

The items outlined below are commonly missing or deficient in submitted manuscripts, leading to a lengthier and less favorable statistical review.

Authors are strongly encouraged to proactively address all of these issues.

At the time of their initial online manuscript submission, the corresponding author will be asked to attest to reviewing and the online supplement that provides details for the following outline.

**PLEASE NOTE: EACH MANUSCRIPT WILL BE EXPLICITLY EVALUATED ON EACH OF THESE ITEMS DURING ITS STATISTICAL REVIEW.**

1. **Abstract clearly and accurately states the study objectives/hypotheses and describes data analysis and study findings:** As a standalone summary, the Abstract should convey this key information.

2. **Study objectives and/or hypotheses clearly stated:** Study objectives/hypotheses are clearly stated in the Introduction, as the applied statistical methods directly follow from them.

3. **Study design is appropriate for the stated aims:** For example, if the authors assessed the association between an exposure and outcome in an observational study, they should collect or obtain sufficient data on potential confounding variables.

4. **Primary and secondary outcomes:** Primary and secondary outcomes must be clearly identified and distinguished in the Abstract, Methods, Results (and/or Discussion). The designation as primary or secondary outcome should have been decided *a priori*. If true, this should be stated; if not, the reasons why should be explained. While it is acceptable to present findings not anticipated in the study design, these should be clearly identified as *post hoc* observations.

5. **Statistical methods subsection:** Statistical methods are appropriate for the study design and data, directly follow from the initially stated objectives and/or hypotheses, and are clearly and comprehensively described. This subsection should not be a general list of tests, but instead sufficiently detailed, including all conducted analyses, specifying the independent (predictor) and dependent (outcome) variables for each analysis (as appropriate). The most appropriate statistical method depends on several factors, including:

   - Aim of analysis: e.g., estimation of a population parameter versus testing of a hypothesis
   - Research design: e.g., repeatedly measured data in a longitudinal study or survival data
• Type and distribution of the outcome variable: e.g., nominal, ordinal or continuous

**Note:** Authors are referred to the series of basic statistical tutorials in *Anesthesia & Analgesia* for an overview of statistical techniques that are appropriate for different study designs and variables.

6. **Baseline comparisons:** In a randomized trial, authors should not include P-values or related tests comparing randomized groups on baseline characteristics. Rather, simply discuss whether clinically important differences in the observed values are apparent or not. Since there is no hypothesis being tested at baseline, such P-values are not appropriate. Instead consider assessing balance using standardized difference. Guidance is that absolute standardized difference greater than 0.10 is evidence of imbalance. In Statistical Methods, state what you had planned to do, if anything, if clinical imbalances were found at baseline (e.g., include those variables in a multivariable model when assessing association between exposure/treatment and outcome).

However, for non-randomized studies, comparing groups on baseline characteristics using statistical tests is important and highly recommended.

7. **Assumptions:** The majority of statistical tests require specific assumptions be met. Report how the key assumptions of the conducted analyses were assessed and confirmed -- for example, assumptions of normality, independence, relationship between the treatment/exposure and outcome variables (e.g., linearity), collinearity, missing data mechanism (see detail in item #9), proportional hazards assumption, model goodness of fit.

8. **Type I error/multiple testing:** If there are multiple primary outcomes or multiple testing, explain how a Type I error is protected at the given significance level (e.g., 0.05) (e.g., Bonferroni correction or other method). Differentiate the overall significance level for a hypothesis from the significance criterion (P-value cut-point) that is applied to individual tests.

**Note:** Authors are discouraged from using the argument that adjusting for multiple comparisons or multiple testing should not be done because it increases the risk of a Type II error (decreases power). While a more stringent significance criterion decreases power, that should be seen as the trade-off of multiple testing, and the sample size needs to be increased accordingly. Neglecting to adjust for Type I error can lead to much higher chance of some or many of a statistically significant results being false positives. The goal should be to focus on most important exposures and outcomes in the study design phase.

9. **Missing data:** The amount of missing data, as well as reasons for missing data, are clearly stated throughout the manuscript. Missing data can (a) lead to a loss in statistical power to detect effects and (b) bias the results, depending on the distribution of missing values across the observations. It is essential that authors report the total number of study subjects, as well as the number of missing data per key variable and per study group. If a relevant amount of data is missing (> 5% as a rough guide), authors should attempt to explore missing data mechanisms (missing completely at random, missing at random, missing not at random) and describe how missing data were handled in the analysis (e.g., listwise or pairwise deletion of the observations, imputation of missing values).

10. **Justify the sample size:** Sample size justification is a key feature of the design (not the results) of a research study. At end of statistical methods subsection, authors should describe how the sample size was chosen, including either a statement of power to detect clinically important differences at the chosen alpha level or available precision to estimate the effect or parameter of interest. Sample size justification should include appropriate adjustments for multiple comparisons, multiple testing or interim analyses for the primary aim. Justification is important whether the findings are positive or negative.

If no *a priori* calculation was done, authors should still explain how the sample size was arrived at and also give either (a) the available power with the given sample size to detect what the authors believe would be a clinically important association and/or (b) what difference the authors had 90% power to detect with the given sample size and observed or expected variability, and/or (c) the width of the confidence interval that was attainable for the primary aim with the given sample size.

**Note:** *Anesthesia & Analgesia* is not asking for a “post-hoc” power analysis, in which authors report power for observed differences. We are instead asking for power to detect clinically important and preferably *a priori* specified differences, independent of what was actually observed.

11. **Results section follows clearly from the study objectives and statistical methods:** Primary and then secondary aims should be addressed in sequence, with clear differentiation. No new statistical methods should be introduced in the Results, when they have not been stated and described earlier in the Methods.
Reporting follows the SAMPL guidelines, which can be accessed at [http://www.equator-network.org/reporting-guidelines/sampl/](http://www.equator-network.org/reporting-guidelines/sampl/).

12. **Treatment effect estimates and their variability are reported:** Effect estimates and their confidence intervals should be reported in both the Abstract and Results sections, at least for the primary outcomes. Also report confidence intervals for estimates of incidence, prevalence, when the primary outcome. Confidence intervals for the primary outcomes are interpreted as the best evidence for where the treatment effect or association of interest lies.

   **Note:** Non-significant results should only be considered as conclusively negative when the confidence interval does not include what authors or others would consider to be clinically important effects.

13. **Confounding:** For non-randomized studies assessing the association between and exposure and outcome, address potential confounding of the relationship of interest as thoroughly as possible using multivariable regression, propensity score methods, or other methods. Since the goal is typically to adjust for as much confounding as possible, it is usually neither desirable nor ideal to use a so-called parsimonious model when considering which variables to adjust for. Adjustment should instead be more liberal. When limited adjustment is made, for whatever reason, list this as a study limitation in the Discussion.

   **Example:** In retrospective database studies, researchers may assess the association between an exposure of interest (such as receiving an intraoperative blood transfusion) and a major postoperative complication or event. Because the exposure groups are not randomized, they may differ on baseline variables (e.g., age, sex, race, BMI, comorbid diseases), variables that themselves may be strongly associated with the outcome variable. Researchers will typically either control for such variables in a multivariable model when assessing the association of interest; or alternatively, use propensity score (PS) methods either (a) to match exposed and non-exposed patients on the set of potentially confounding variables or (b) to inversely weight by the PS. With each method, the goal is to reduce confounding.

14. **Tables and Figures should be clear and self-explanatory:** Tables and figures, along with their legends and footnotes, clearly present the results and include enough information about what was done statistically to basically stand alone, independent of the statistical methods subsection of the manuscript. Tables should generally include the patient or unit denominator (sample size).

15. **Limitations:** All limitations of the study design (e.g., inherent limitations of observational studies) or the statistical analysis (e.g., potential residual confounding, limited sample size, missing data) are transparently acknowledged and clearly presented in the Discussion section.

16. **Conclusions and Interpretations:** Interpretations and conclusions are justified by the design and results. They do not go beyond what was tested or assessed in the study, and instead focus on primary endpoint(s).

   **Specifically:**

   - **Causation/association:** In observational studies—whether retrospective or prospective—authors generally avoid using language that would imply a cause and effect relationship (such studies can typically only identify an association between a variable and an outcome). For example, generally avoid saying that the exposure “reduced” the outcome, or “effect” of the exposure on outcome. Instead, state and discuss that an “association” was observed between exposure and outcome. When discussing observational results, please be as conservative as possible. Many observational studies essentially demonstrate that sicker patients do worse! Methodologic limitations, including the potential for unidentified confounding, should be transparently discussed.

     In observational studies in which the authors have (a) made a strong case that they have appropriately and thoroughly adjusted for confounding and (b) conducted rigorous sensitivity analyses to demonstrate the robustness of their findings, they may be given more latitude to suggest causality in the relationship of interest. However, even in such settings, the limitations inherent to any observational study should be clearly outlined and stated in Discussion.

   - **Non-significant instead of similar/equivalent:** Non-significant results from superiority tests should not be interpreted as evidence of no difference or no effect in the population. They also do not imply that groups are “similar” or “equivalent.” A specific design (equivalence study) and tailored analytic methods are required to make claims of equivalence or similarity. Rather, authors should simply state that no significant difference or effect was found or that the data do not provide evidence for a difference or effect, and report the confidence interval for the effect estimate.
• **Make inference on population not sample**: Avoid claiming that a certain outcome variable was “higher” or “more frequent” in one group than the other if (a) there is no statistically significant difference or (b) a formal hypothesis test was not performed. While the statement may be true for the sample, the inferences from the study apply to the population from which the sample was taken. Therefore, such are not supported by the data and must be avoided.

• **Trend**: Do not state that the nearly statistically significant result represents a trend in the data. Do not state “there was an effect of X on Y” and then state that it was non-significant—instead, simply state that it was non-significant or that no association was found.

17. **P-values**: Report all actual P-values, not a P < 0.05, P > 0.05, or “NS.” P-values are usually be rounded to 2 or 3 decimal places. Actual P-values < 0.001 should be reported as P < 0.001, while P-values > 0.99 should generally be reported as P > 0.99, not P = 1.0. However, reporting “P=1” is appropriate when the observed means or proportions comparing groups are exactly the same (e.g., in a 2-tailed test against null hypothesis of equal means or proportions).

18. State “multivariable” instead of “multivariate” when there are multiple independent/predictor variables and a single dependent/outcome variable.