Topic	*No.*	*Item*
Subjects under study	1	Identify the target population, state the method of selection of the sample, total sample size, stratification if any, and the groups under study.
Sample size	2a	State the sample size for each group and justify the size for the stated precision, alpha error, and/or power. For power, specify the smallest effect size considered medically important with reasons.
	2b	State the number of missing values, outliers and other exclusions with reasons, comment on the representativeness of the sample finally available for analysis, and describe possible biases with measures taken to control them.
Hypothesis	3a	State all the hypotheses keeping the study objectives in mind.
	3b	State the minimum effect size to be considered as medically important, if applicable, with its rationale (see Item 1b). For equivalence and non-inferiority studies, give the largest medically unimportant margin with reasons.
Variables under study	4a	State all the variables on which the data were collected and identify the ones on which the present analysis was done along with the rationale of the choice of variables. State the unit of measurement of each, and describe the validity of the methods of measurement for each variable.
	4b	Categorize continuous data for presentation of distribution if needed. If helpful, give histogram and comment on the distribution pattern, particularly of the outcome variables.
	4c	If dichotomous or polytomous categories have been used in analysis of continuous variables, explain the rationale of these categories in terms of clinical implication.
Antecedents and outcomes	5a	In the case of analytical studies, identify the antecedent factors under study, the outcomes of interest, and the covariates included.
	5b	Define the effect of interest in terms of the variables included in the study (the effect size can be difference between means or between proportions, odds ratio, correlation coefficient, phi coefficient, or any other measure).
Descriptive summaries	6	Summarize the data –Provide mean (SD) (and not mean ± SD) or median (IQR) of each continuous variable depending upon the Gaussian or (highly) skewed distribution, respectively (do not use SE here). For IQR, give the values of the first and third quartile. Do not give such summaries for groups with $n \leq 4$; give the original values instead. For categorical data, state actual frequency in different categories and the percentage if $n \geq 20$. All summaries should be with the appropriate degree of decimal accuracy as specified at the end of these guidelines*.
Modification of raw data	7	Describe transformation such as log and square-root, if any, with reasons and the method of calculation of scores, and rates and ratios, and fully specify the numerator, denominator and multiplier (per cent, per million, *etc*.) for each where applicable. For rates, specify the time period (per day, per year, *etc*.).
Baseline information	8	Summarize all important demographic and clinical features of the subjects in each group, particularly those that can affect the outcome (see Item 6).
Comparability of two or more groups	9	Before comparing two or more groups with respect to outcomes in terms of summaries such as means, proportions in different categories, and rates, confirm that the groups are comparable with regard to the baseline composition of the subjects for factors (such as the age distribution) that can affect the outcome. If not comparable, report the re-computed summaries after proper standardization. If standardization required but not done, state reasons and explain how the outcomes in various groups can still be compared.
Main method of analysis	10a	Describe the method for each analysis, confirm the validity of the underlying assumptions, and justify the parametric and non-parametric methods used for different variables. Provide reference or explain the methods not in common use. State the software used for analysis with version.
	10b	Identify post-hoc analysis if any, including sub-groups analysis, and interpret this as exploratory and not confirmatory.
Estimation	11	For descriptive part of the study, provide estimate of the mean, proportion, difference, *etc*. with 95% confidence interval (CI). Justify the Gaussian approximation in case this is used for computing the CI. In case any other confidence level is used, provide the rationale.

Contd...

Topic	*No.*	*Item*
Tests of statistical hypothesis	12a	State the statistical hypothesis for each test. Give the name of each test and its exact *P*-value with df where relevant. For $P<0.001$, state with less than sign and for $P>0.999$ with more than sign. Indicate whether the test is one-tailed or two-tailed with the reasons thereof. Avoid the use of the term statistical significance and do not mention significance level (such as $\alpha = 0.05$) for your results. Mention about any adjustment made for multiple comparisons and for using multiple tests for any conclusion. Distinguish between family-wise error rate and experiment-wise error rate. Also mention the CI for the effect size such as mean difference between the groups.
	12b	Report all the results and not just those that have low *P*-value. Interpret larger *P*-value as inconclusive and not as negative result unless the power is high to detect a specified medically important effect. Distinguish between results with low *P*-value (conventional statistical significance) and medical significance of the results.
Robustness of results	13	Comment about the statistical limitations of the study in addition to the other limita-tions. Statistical limitations could be due to imprecision of the measurements, restricted analysis because of the nature of the data or size of sample in different groups, not fulfilling the underlying assumptions, lack of representativeness of the sample, compromised design, lack of internal or external validations, and such other deficiencies.
The following are needed if these methods have been used in your paper		
Correlation and cause-effect	14a	Report the value of the relevant correlation coefficient. If described as low, moderate or high, give the categories with their biological implications. Interpret conventional Pearson correlation coefficient for assessing linear relationship and not for any general relationship between continuous variables. For association between categorical variables, include the full contingency table and explain if any categories were merged for analysis purpose.
	14b	Distinguish between association/correlation and cause-effect. If cause-effect is implied, rule out all possible alternative explanations such as the role of confounders and biases.
	14c	Distinguish correlation/association from agreement.
Regression analysis	15a	Describe the purpose of the regression analysis (explanatory or predictive), identify the response (outcome) and regressor (antecedent) variables with the selection process if any, assess colinearity between independent variables, and provide medical and statistical rationale of the chosen model (linear/nonlinear, simple/multivariable). State the size of sample available for running each regression and comment on its adequacy. In case the model is being used for prediction of individual values, give prediction interval and not the CI for mean. Do not predict for values much beyond the values actually studied.
	15b	Report the regression equation with comments on its adequacy based on indicators such as coefficient of determination (η^2, whose linear component is R^2) for quantitative and generalized R^2 for logistic regression, and report exact *P*-value for each regression coefficient with the associated CI. For quantitative dependent in simple linear or curvilinear regression, plot the regression line or curve with scatter where helpful and comment on the randomness of the residuals. For logistic regression, specify the reference category for categorical regressors, give odds ratio (OR) and the CI for each variable – adjusted as well as unadjusted. For cohort studies, state the number of subjects with positive and negative outcomes, and the relative risk with their CI – again adjusted as well as unadjusted. In the case of multivariable regression, interpret regression coefficient as adjusted only for the other variables in the model and give plausible biological explanation of the model obtained.
	15c	Specify whether and how the model was validated, or why it could not be validated.
Survival analysis	16a	Describe the purpose of the survival analysis, identify the beginning- and the end-point for the duration under study, specify censoring, name the survival analysis method with the confirmation of the assumptions, plot the survival curve and report the median survival time with the CI, and discuss the points of inflexion in the survival curve, if relevant.
	16b	Where helpful, give the table with the estimated survival probability at each follow-up with the CI.

Contd...

Topic	*No.*	*Item*
	16c	Specify the method used for comparing two or more survival curves if applicable and give exact *P*-value. Interpret it for overall survival pattern and not for specific time-points.

Decimal accuracy (rounded) as follows

Percentages - One decimal place if n <100 and two decimal places for n ≥100;
Mean and SD (Median and IQR) - One decimal place more than the original values;
Correlation coefficient - Generally two decimal places;
Odds ratio, relative risk and hazard ratio - Generally two decimal places;
P-values - Exact P-values to three decimal places and not as P <0.05 or P ≥0.05 (For extremely small values, write P <0.001, and for extremely high values, write P >0.999).