APPENDIX -1

Instructions for Reporting Statistical Methods and Results

Statistical analyses are an essential feature of medical studies and the appropriate method must be used to properly answer research questions and hypotheses. To ensure that an article will not have any statistical deficiencies, the following steps should be adhered to carefully: (i) the study should be designed appropriately, (ii) the statistical methods employed should be selected and performed carefully and accurately, and (iii) the results should be reported correctly. Statistical analyses should be conducted in accordance with international statistical reporting standards, e.g., the SAMPL guidelines (*T. A. Lang and D.G. Altman. Basic statistical reporting for articles published in biomedical journals: The Statistical Analyses and Methods in the Published Literature or SAMPL Guidelines, Section 5.7 in Smart P, Maisonneuve H, Polderman A (eds). Science Editors' Handbook, European Association of Science Editors, 2013.*).

We highly recommend the use of study-specific guidelines, such as those created by the Consolidated Standards of Reporting Trials group (CONSORT) and the STrengthening the Reporting of OBservational studies in Epidemiology initiative (STROBE).

The study design (e.g., retrospective case-control, cross-sectional, cohort, etc.) should be described in detail in the Materials and Methods section. Please refer to the SAMPL guidelines for more details.

The description of the statistical analyses performed should be provided under a separate subheading of Statistical Analysis in the Materials and Methods section. This section should detail the following:

- **Data collection**: Information about how the data were collected. In this step, sampling strategy and details of power and sample size calculation should be provided. Any inclusion/exclusion criteria should also be provided;

- **Data processing:** Information about how the raw data were processed before data analysis (e.g., identifying outliers, normalization, data transformation, etc.);

- **Summarizing data:** Information about how data values were expressed (e.g., values are expressed as mean ± SD or median [1st-3rd quartiles]) as results of the analysis. A brief mention of selected summary statistics may also be useful;

- **Assumptions:** Information about how the statistical assumptions were tested (e.g., histogram and Q-Q plots were examined, the Shapiro-Wilk's test was used to assess data normality, the Levene test was used to test variance homogeneity, etc.);

- **Hypothesis tests and modeling:** Information about the statistical methods used and the purpose (e.g., to compare the miRNA levels of patients with and without cancer, a two-sided independent samples *t*-test was applied, a linear regression model was fitted to adjust for multiple covariates, etc.);

- **Software:** Information about any statistical software used in the data analysis (e.g., analyses were conducted using TURCOSA [Turcosa Analytics Ltd. Co., Kayseri, Turkey] statistical software.).

Statistical modeling (e.g., regression analysis, classification, clustering, linear models, etc.) requires significant effort and the steps used should be described in detail. What were the univariate analysis results? What criteria were used to select variables included for multivariate analysis? How did the final model fit with the data? What was the goodness-of-fit of the selected model? Finally, the model diagnostics to determine the validity of the built model should be explained.

Reference links to guidelines:

SAMPL: https://www.equator-network.org/2013/02/11/sampl-guidelines-for-statistical-reporting/

CONSORT: http://www.consort-statement.org

STROBE:

www.strobe-statement.org