Statistical Analysis Plan
Camel Milk Project

Topic: Statistical analysis plan

This supplementary material has been provided by the project researchers to give readers additional information about the work.

Version 1

Approval date: 12/08/2024

Total number of pages: 10

Statistical Analysis Plan

Version 1, Date: 12/08/2024

PI:

Name	**Function**	**Date of the signature**	**Signature**
Dr. Mo'ath Bataineh	Principal Investigator	12/08/2024	

Co-PI:

Name	**Function**	**Date of the signature**	**Signature**

1- Changes history

All changes and versions of this document are reported in the table below:

Version	Author	Description of modifications	Date
V1	Dr. Mo'ath Bataineh	First version	12/08/2024

---

2 - Definitions

CM Camel Milk

3 - Objectives

3.1 Primary Endpoint

- To assess the impact of camel milk on exercise recovery and subsequent physical performance

3.2 Secondary Endpoints

- Check impact on inflammatory markers.

4 - Project Design

The study will use randomized crossover counterbalanced design to collect information about impact of camel milk on physical performance.

4.1 Inclusion criteria

The inclusion criteria for this study are the following:

(1) Recreational athlete, active 5 days a week.

(2) Male, healthy, non-obese.

(3) Age 18 to 39 years old.

(4) No allergies to milk.

(4) provide consent to participate and able and ready to complete the study's interview.

(5) Speaking and understanding Arabic or English, and open to visit him at home.

(6) Subject consenting to participate in the study.

4.2 Exclusion criteria

(1) Sedentary

(2) Female

(3) Allergies to milk

(4) Presence of diseases.

4.3 Sampling

The sample for this study will be obtained from UAEU community populations.

4.4 Sample size calculation

Sample size was calculated with the use of G*Power 3. 1.9.7 that revealed the need for 12 participants to achieve a 0.95 power of the test with medium effect size and alpha set at a 0.05 value. Thus, a minimum of 12 participants will be recruited for the current study.

5 - Measured variables retained for the analysis

Table of variables

Variables	Format	Description
Code	String	Three letter code of participant
Date	Date	Date of data collection
Time	Numeric	Time of data collection
Sex	Numeric	Sex: male or female
Age	Numeric	Age in years
Marital	Numeric	Marital status
Education	Numeric	Education level
Self-Health	Numeric	Self-perceived health
Wt	Numeric	Self-reported weight in kg
Ht	Numeric	Self-reported height in cm
Order	Numeric	Random order of trial
Visit	Numeric	
Treatment_Drink	Numeric	
Treatment_Recognition	Numeric	
Session	Numeric	
BMI_Measured	Numeric	
HR_rest	Numeric	
HR_Peak	Numeric	
RPE_rest	Numeric	

RPE_Peak	Numeric	
VO2 rest	Numeric	
VO2 Peak	Numeric	
RQ rest	Numeric	
RQ Peak	Numeric	
Cytokine1	Numeric	
Cytokine2	Numeric	
Cytokine3	Numeric	
CK	Numeric	

6- Statistical Analysis

The statistical analysis will be done according to the intention-to-describe principle. We will report summary statistics according to the type and normality of data. Data will be reported as mean and standard deviation, median and interquartile range for continuous variables and for categorical variables as frequency and percentage. When appropriate data will be reported with a 95 % confidence interval.

6.2 Methods of Analysis

6.2.1 Principal Analysis

The Chi-square test will be used to assess the impact of demographics on dependent variables. The repeated measures ANOVA will be used to assess main effects. One way ANOVA and t-test will be used to compare effects of independent variables on measured variables. A p value <0.05 will be considered statistically significant. Effect size will be calculated and presented to provide an estimation of the significance magnitude. If appropriate the 95% confidence interval will be reported.

6.2.2 Secondary Analysis

The same methods will be used to assess the impact inflammation markers.

6.3 Specific methods for handling missing data

The researchers will perform a missing data analysis to report on the mode and the percentage of missing values. Variables with less than 5% of missing values will be imputed with the use of median for the continuous variables and/or the highest frequency for the categorical variables. In the case of missing values higher than 5%, the variable will be excluded from the final analysis.

7 - Information systems, software, and program outcome validation

7.1 Information system and used analysis software

All analyses will be performed using SPSS Version 29.0 (IBM Inc., USA) running Microsoft Windows11.

Graphic representations will be performed using PowerPoint.

7.2 Validation of software outcome

The statistician will be responsible for reviewing each program and the associated SPSS output. The SPSS "Journal" will be examined for logical errors, syntax errors, and serious errors. The results will be examined for typographical errors, misspellings, and irrational values. The consistency of the results will be examined.

reporting using descriptive tables

8.1 Description of demographic characteristics

Variable name	Visit			Missing (n)	p-value
	1	2	3		
Age, mean (SD)					
Self-Health, n (%) Poor/Fair Good Very Good/Excellent					
HR rest, median (IQR)					

8 - Suggested data