Study Protocol
with
Statistical Analysis Plan

Identification of Physiological Biomarkers of Gastro-intestinal Discomforts Induced by Milk Consumption

MIDI

NCT04205045

Version: 1
Date: 26/07/2022

Author: Paola Vitaglione, University of Naples, Federico II

Principal Investigator: Paola Vitaglione, University of Naples, Federico II

Background and rationale

Gastro-intestinal discomforts associated with the consumption of milk and dairy products are frequent in the general population, even in absence of lactose intolerance and casein allergy. Evidence from the literature suggests that these disorders (in the absence of other diagnoses) may be due to some peptides produced during the digestion of milk proteins. Moreover, some scientific evidence suggests that other factors, such as an altered intestinal permeability and/or a specific composition of the intestinal microbiota, may be responsible for gastrointestinal (GI) disorders resulting from milk consumption. It is also possible to hypothesize a role of the endocannabinoid system in the evolution of GI symptoms after milk ingestion given its implications in homeostatic and non-homeostatic regulation of feeding behavior, as well as in pain regulation processes.

Objectives

This study aims to compare two groups of healthy (non-lactose intolerant) subjects who are irregular milk consumers, due to milk induced gastrointestinal disorders, with another group of healthy subjects who consume milk regularly, without symptoms, in relation to the primary and secondary objectives here described.

By performing a milk test in which volunteers ingest 250 ml of UHT semi-skimmed milk, primary objectives include the evaluation of:

- plasmatic response of milk-derived peptides;
- plasmatic profile of GI hormones;
- plasmatic profile of endocannabinoids and N-acyl-ethanolamines.
- urinary profile of lactose over 24 h;

Secondary objectives include the assessment of:

- baseline and postprandial plasma profile of aminoacids;
- baseline and postprandial blood glucose;
- baseline and postprandial GI symptoms;
- baseline serum concentration of dipeptidyl peptidase IV (DPPIV);
- baseline and postprandial serum activity of dipeptidyl peptidase IV (DPPIV);
- urinary excretion of indican as marker of protein digestibility;
- correlations of GI symptoms with appetite ratings and blood glucose, plasma GI hormones, plasma peptides and amino acids as well as with urinary excretion of lactose and galactose + glucose;
- the composition of the intestinal microbiota (metagenomic analysis);
- the psychological profile through questionnaires on the perception of quality of life, on levels of depression, anxiety and stress and on eating behavior.

Further secondary objectives include the assessment of:

- the prevalence of lactose malabsorption in the two groups, by using results from Lactose H_2 breath test performed during the recruitment procedure;
- urinary excretion of lactulose, mannitol and sucralose as markers of intestinal permeability after an oral challenge of these 3 sugar probes.

Trial design

Recruitment

The recruitment process is based on:

- **Pre-recruitment questionnaires** to collect information on dietary habits and milk-related symptoms;
- **Lactose breath test** for the assessment of final eligibility. After lactose ingestion, H_2 measurements in breath will be performed over 4 hours; a cut-off value of 20 ppm over the baseline will be used to identify lactose malabsorbers. In parallel, and up to 24 hours, subjects will also record GI symptoms, through Visual Analogue Scales (VAS).

The eligibility of volunteers to participate in the study is defined on the basis of the following inclusion and exclusion criteria:

Inclusion Criteria:

- Drinking milk (maximum 150 ml/week for non-habitual milk consumers and minimum of 700 ml/week for habitual milk consumers).
- BMI between 18.5 - 30 kg/m^2.
- Availability to participate into three visits (with one week in between).
- Willing to drink 250 ml of milk in fasting condition within 10 min on one occasion.
- Signature of a written informed consent.
- Negative lactose breath test result (increase in H_2 concentrations < 20 ppm vs basal value) and without symptoms or with any symptoms except "vomiting" and "Loose, mushy or watery stools".
- Lactose malabsorbers (increase in H_2 concentrations > 20 ppm vs basal value but without symptoms).

Exclusion Criteria:

- Presence of any relevant organic, systemic or metabolic disease or abnormal laboratory values.
- Ascertained intestinal organic diseases, including celiac disease or inflammatory bowel diseases.
- Previous major abdominal surgeries.
- Active malignancy of any type, or history of a malignancy (patients with a history of other malignancies that have been surgically removed and who have no evidence of recurrence for at least five years before study enrolment are also acceptable).
- Untreated food intolerance.
- Lactose intolerant
- Assumption of probiotics or topic and/or systemic antibiotic therapy during the last month.
- Systematic/frequent assumption of contact laxatives.
- Pregnant and lactating women.
- Inability to conform with protocol.
- Treatment with any investigational drug within the previous 30 days.

- Recent history or suspicion of alcohol abuse or drug addiction.
- Subjects having symptoms of "vomiting" and "Loose, mushy or watery stools" at any level of severity following milk consumption and the lactose breath test

Allocation

Subjects who will be assessed as eligible and will accept to enter the study protocol, will be allocated in the proper group, and the enrollment will be concluded when 2 groups of subjects will be constituted as follows:

- **Group 1**, habitual milk consumers (HMC): 20 subjects who are regular milk consumers (>700 mL/week), do not experience GI discomforts after milk intake and are lactose tolerant or lactose malabsorber.
- **Group 2**, non-habitual milk consumers (NHMC): 20 subjects who are not regular milk consumers (< 150 mL/week), because of GI discomforts after milk intake and are lactose tolerant or lactose malabsorber.

Intervention and other measures/samples

The study protocol includes a gut permeability test and a milk test to be performed for all subjects from both groups after one week from the lactose breath test and with one week in between.

Gut permeability test

Subjects will receive instruction to avoid medicine consumption and to follow a controlled diet for 2 days before the test (avoiding milk and dairy products, and food products containing polyols that were included in a list of products they received).

Fasting subjects will collect a baseline urine sample, before drinking a solution containing 5 g lactulose, 2 g mannitol and 2 g sucralose in 100 ml of water. Urines will be collected for time periods 0-5 hours and 5-24 hours in two containers containing 1 mL of a chlorhexidine solution at 1 ppm.

Urine samples collected over 24 hours and delivered the day after to the study centre will be aliquoted (1.5 ml) and frozen at −40°C until analyses that include lactulose, mannitol and sucralose concentration and baseline creatinine concentration.

The ratio between lactulose and mannitol (L / M) in the urine collected within the first 5 hours will represent permeability in the small intestine, while the total amount of sucralose excreted in urine in the 24 hours will be used to evaluate the colonic permeability.

Milk test

During the week before the milk test, subjects will fill a food diary. Moreover they will be instructed to not drink any milk, dairy product and any food product containing milk proteins during the 2 days before the milk test. On the day of the milk test, fasting subjects will fill out questionnaires about their actual appetite and GI symptoms, baseline urine and blood samples will be collected, and fasting glycaemia by finger-prick will be measured (time 0). Subsequently, subjects will drink 250 ml of milk within 10 min and thereafter, glycaemia, symptoms and appetite questionnaires, as well as blood samples and urine will be collected at specific time points (time 0.5, 1, 2, 4 and 6 hours after milk ingestion). After the last blood drawing, subjects will be offered lunch and, before leaving the study centre, they will be instructed to fill out GI symptoms and appetite questionnaires, to collect urine, to

fill out a 1-day food diary, and to consume a fixed dinner. Questionnaires, diaries and 6-24h urine samples will be delivered to the laboratory on the next day.

Blood samples will be centrifuged at 4°C and 4000 rpm; serum, plasma and plasma + aprotinin samples will be aliquoted and stored at −40°C.

Plasma samples will be analyzed for:

- casein digestion-derived peptides (beta-casomorphins and all known peptides derived from milk proteins);
- aminoacid profile;
- endocannabinoids and N-acyl-ethanolamines (ECs).

Plasma samples added with aprotinin will be analyzed for:

- gastro-intestinal hormones (in plasma samples pre-treated with aprotinin): insulin, GIP, GLP-1, glucagon, c-peptide, ghrelin, leptin.

Serum samples will be analyzed for:

- DPPIV activity;
- baseline DPPIV concentration.

In urine samples, aliquoted (1.5 ml) and frozen at −40°C, all the following analytes will be measured:

- lactose excretion levels;
- galactose + glucose excretion levels;
- indican excretion levels;
- baseline creatinine concentration.

Subjects will be also asked to collect a fecal sample under usual dietary habits and to fill out questionnaires as follows:

- physical activity questionnaire (short version of the IPAQ);
- questionnaire on quality of life;
- questionnaire on depression, anxiety and stress (DASS);
- semi-quantitative food frequency questionnaire (FFQ);
- the King's Stool Chart, to evaluate frequency and consistency of feces;
- Three Factor Eating Questionnaire (TFEQ) to evaluate individual eating behaviour.

Anthropometric data of weight and height will be measured as well.

Fecal samples, stored at -40°C, will be used for DNA extraction according to the standard operating procedures of the International Human Microbiome Consortium (http://www.human-microbiome.org/). Statistical analysis will be carried out by the most appropriate methods in order to infer the associations between specific microbiota signatures and clinical, dietary and metabolomic variables.

Statistical analysis plan

Sample size calculation

The sample size for this study has been calculated on the basis of the primary endpoints here listed:

- milk-derived peptides plasma concentrations after milk consumption;
- lactose urinary excretion after milk consumption;
- post-prandial plasma profile of GI hormones and ECs.

Considering circulating mean values and standard deviations reported by Deth et al. (2016) in blood of healthy population, power analysis indicated that a sample size of 19 participants per group is sufficient to detect a minimum variation of 40% in circulating BCM-7 with a power of 80% and an $\alpha = 0.05$.
Considering excreted % of lactose in urine, hydrogen concentration in breath and relative standard deviations as reported by Bezerra et al. (1990) a sample size of 19 participants per group is sufficient to detect a 15% minimum variation of lactose and a 50% minimum variation of hydrogen in healthy subjects, with a power of 80% and an $\alpha = 0.05$.
Regarding post-prandial plasma profile of GI hormones and ECs, our previous studies showed that a sample size of 13 participants is adequate to find significant differences with an 80% power and an $\alpha = 0.05$. (Mennella et al., 2015; Mennella et al., 2016).

The participant number will be increased up to 40 (20 subjects per group) in view of possible dropouts.

Statistical analysis

Data will be expressed as means ± SEMs, unless otherwise specified.
Two-tailed P values lower than 0.05 were considered significantly different

Missing data will be handled as such, i.e., no imputation of missing data will be performed.

<u>Categorical variables</u>

For variables with 2 categories, Chi-square test will be used to evaluate the association between variables if the expected values in each cell are ≥ 5 in at least 25% of the cells. When the expected values are < 5 in more than 25 % of the cells, Fisher's exact test will be used.

To assess the association between two polytomous variables (>2 categories), the Pearson's Chi-square test will be used if the frequency expected is ≥ 5 in at least 75% of the cells. If not, we will regroup to increase the expected frequencies.

<u>Continuous variables</u>

Normality of continuous variables will be assessed using the Shapiro-Wilk test. For non-normal variables showing a significant positive skewness, logarithmic, square-root, exponential, or inverse transformation will be applied. For those variables which will not show a normal distribution after arithmetic transformation, non-parametric tests will be used.

Continuous variables without repeated measures

For those variables which will show a normal distribution at the Shapiro-Wilk test, an independent two-sample T test will be performed to check differences between groups, whereas ANCOVA analysis will be performed to include covariates in the analysis.

Differences between groups of non-normally distributed variables will be assessed by Mann-Whitney U test, while non-parametric analysis of covariance (Quade's test) will be used to explore the effects of any covariates.

Continuous variables with repeated measures

Normality of measures, for each time, will be assessed using the Shapiro-Wilk test.

Differences of variables between baseline and over time within and between the groups, for normally distributed variables, will be tested by ANOVA with repeated measures on one factor (time) and Bonferroni adjustment for multiple comparisons. For non-normally distributed variables, Friedman test and pairwise Wilcoxon as a *post hoc* test will be applied.

Analysis of quartiles will be performed for both baseline values and pairwise time points (Δ) upon milk ingestion.

For these variables with repeated measures, the total area under the curves (AUCs) will be also estimated using the linear trapezoidal rule, and differences in AUCs between the groups will be assessed by proper parametric or non-parametric analyses including as covariates the potential confounding variables measured at baseline.

Pearson's correlation coefficients will be calculated to assess bivariate associations among continuous variables at baseline or partial associations of estimated AUCs controlling for baseline values for variables with repeated measures.

Spearman's rank correlation will be used to assess association between non-parametric variables collected and microbiota related variables (e. g. species abundance and genes abundance).

In order to explore differences in microbiome profiles, a Partial Least Squares Discriminant Analysis was applied.

Furthermore, principal component analyses (PCA) on specific variables at baseline or unstandardised residuals correlation between AUCs and related baseline values will be also performed in order to extract features relevant to the cluster structure. Moreover, statistical significance of the distance between groups in the PCA analysis will be computed using the Hotelling T2 test.

Further analyses

Since the prevalence for lactose malabsorption over the enrollment will be assessed, all the statistical analyses reported above may be applied to subgroups of subjects classified as "lactose tolerants (LT)" and "lactose malabsorbers (LM)" in order to check differences between the two groups and/or correlations between variables assessed.

Chi-means will allow to assess differences between subgroups classified on the basis of lactose absorption capacity of subjects as detected by lactose H_2 breath test.

Statistical analyses and data visualization will be carried out in R environment (http://www.r-project.org).