Statistical Analysis Plan (SAP)

Trial ID: VOMIT

Title of Trial

Validating the Effect of **O**ndansetron and **M**irtazapine **I**n **T**reating Hyperemesis Gravidarum. A Double-blind Randomized Placebo-Controlled Multicentre Trial.

Investigational Medicinal Products:	Mirtazapine and Ondansetron
EudraCT No.:	2018-002285-39
Sponsor:	Ellen Christine Leth Løkkegaard
GCP-Monitor:	Stine Kjær Hovgaard
Ethical Committee number:	H-18047191
Clinicaltrials.gov number:	NCT03785691
Data protection approval number:	VOMIT: P-2019-75 VOMIT biobank: P-2019-193
Date:	01.04.2022
Version:	2.2
Investigational Medicinal Product:	Mirtazapine and Ondansetron

SAP revisions:

January 2020:

Changes were made to the primary endpoint, which prior to these changes included difference between mirtazapine and placebo on day 2 and 14, exclusively. By introduction of hierarchical testing the difference between ondansetron and placebo, and mirtazapine and ondansetron were included in the co-primary endpoints.

March 2022:

Due to slow recruiting rates predominantly due to COVID-19, reaching the planned 180 subjects would substantially delay publication of the results. Thus, in January 2022 it was decided to discontinue the trial in July 2022, when the produced trial medication would expire.
Considering the lower-than-planned number of included subjects and the consequently compromised power of the trial, it was therefore necessary to reassess the pre-planned primary analyses.
Subsequently changes have been made to the planned analyses of the primary endpoint. It was initially planned to analyse two of the co-primary endpoints at 0.025 alpha each, but considering the compromised power of the trial we have decided to initially analyse just one of the co-primary endpoints at alpha 0.05 and if significant, subsequent primary endpoints will hereafter be tested hierarchically.

All changes in to the statistical analyses plan were based on consensus among the advisory board and the protocol author after consulting the statistician. Advisory board members, protocol author and statistician are listed on page 4.

Signatures

Sponsor:	**Name:** **Title:** **Department:** **Address:** **Telephone:** **E-mail:**	Ellen Christine Leth Løkkegaard MD, Head of Research, Professor, Consultant, PhD Department of Obstetrics and Gynaecology Nordsjællands Hospital Dyrehavevej 29, 3400 Hillerød +45 48296249 Eloe0002@regionh.dk
	Signature	Date
Statistician:	**Name:** **Title:** **Department:** **Address:** **Telephone:** **E-mail:**	Andreas Kryger Jensen Assistant Professor, PhD Department of Research Nordsjællands Hospital and Department of Biostatistics, Institute of Public Health, University of Copenhagen Øster Farimagsgade 5, 1014 Copenhagen K andreas.emil.kryger.jensen@regionh.dk
	Signature	Date
Chief investigator:	**Name:** **Title:** **Department:** **Address:** **Telephone:** **E-mail:**	Anne Ostenfeld MD, PhD student Department of Obstetrics and Gynaecology Nordsjællands Hospital Dyrehavevej 29, 3400 Hillerød +45 20779758 anne.ostenfeld.02@regionh.dk
	Signature	Date
SAP author:	**Name:** **Title:** **Department:** **Address:** **Telephone:** **E-mail:**	Anne Ostenfeld MD, PhD student Department of Obstetrics and Gynaecology Nordsjællands Hospital Dyrehavevej 29, 3400 Hillerød +45 20779758 anne.ostenfeld.02@regionh.dk
	Signature	Date

Roles and responsibilities:

Protocol author:	**Name:**	Anne Ostenfeld
	Title:	MD, PhD student
	Department:	Department of Obstetrics and Gynaecology Nordsjællands Hospital
	Address:	Dyrehavevej 29, 3400 Hillerød
	Telephone:	+45 20779758
	E-mail:	Anne.ostenfeld.02@regionh.dk
Head-advisor:	**Name:**	Ellen Christine Leth Løkkegaard
	Title:	MD, Head of Research, Professor, Consultant, PhD
	Department:	Department of Obstetrics and Gynaecology Nordsjællands Hospital
	Address:	Dyrehavevej 29, 3400 Hillerød
	Telephone:	+45 48296249
	E-mail:	Eloe0002@regionh.dk
Co-advisor:	**Name:**	Tonny Studsgaard Petersen
	Title:	MD, Associate Professor, Consultant, PhD
	Department:	Department of Clinical Pharmacology Bispebjerg Hospital
	Address:	Bispebjerg Bakke 23, 2400 København NV
	Telephone:	+45 38635855
	E-mail:	Tonny.studsgaard.petersen@regionh.dk
Co-advisor:	**Name:**	Hanne Brix Westergaard
	Title:	MD, Associate Professor, Consultant, PhD
	Department:	Department of Obstetrics and Gynaecology Nordsjællands Hospital
	Address:	Dyrehavevej 29, 3400 Hillerød
	Telephone:	+45 48296364
	E-mail:	Hanne.brix.westergaard@regionh.dk
Co-advisor:	**Name:**	Lars Henning Petersen
	Title:	MD, Professor, Consultant, PhD
	Department:	Department of Obstetrics and Gynaecology Aarhus Universitetshospital
	Address:	Palle Juul-Jensens Blvd. 99, 8200 Aarhus N
	Telephone:	+45 50526512
	E-mail:	lhp@clin.au.dk
Co-advisor:	**Name:**	Jon Thor Trærup Andersen
	Title:	MD, Associate Professor, Consultant, PhD
	Department:	Department of Clinical Pharmacology Bispebjerg Hospital
	Address:	Bispebjerg Bakke 23, 2400 København NV
	Telephone:	
	E-mail:	jon.thor.traerup.andersen@regionh.dk
Statistician:	**Name:**	Andreas Kryger Jensen
	Title:	Assistant Professor, PhD
	Department:	Department of Research Nordsjællands Hospital and Department of Biostatistics, Institute of Public Health, University of Copenhagen
	Address:	Øster Farimagsgade 5, 1014 Copenhagen K
	Telephone:	
	E-mail:	andreas.emil.kryger.jensen@regionh.dk
Monitor:	**Name:**	Stine Kjær Hovgaard
	Title:	GCP-coordinator, nurse
	Department:	Copenhagen University Hospital GCP-unit
	Address:	Bispebjerg Bakke 23, Byg. 51, 3. sal, 2400 Copenhagen NV
	Telephone:	+45 38635626
	E-mail:	Stine.hovgaard@regionh.dk

Contents

List of Abbreviations

AE	Adverse Event
BP	Blood Pressure
CRF	Case Report From
ECG	Electrocardiogram
e-CRF	Electronic Case Report Form
GCP	Good Clinical Practice
HG	Hyperemesis gravidarum
ICH	International Conference of Harmonisation
IMP	Investigational Medicinal Product
NVP	Nausea and Vomiting of Pregnancy
MD	Medical Doctor
OD	Once Daily
PUQE	Pregnancy-Unique Quantification of Emesis and Nausea
SAE	Serious Adverse Event
SAR	Serious Adverse Reaction
SmPC	Summary of Product Characteristics
SUSAR	Suspected Unexpected Serious Adverse Reaction

1. Protocol Synopsis / Summary

Title of trial:

Validating the Effect of **O**ndansetron and **M**irtazapine **I**n **T**reating Hyperemeis Gravidarum (**VOMIT**). A double-blind randomized placebo-controlled multicentre trial.

Objectives:

The primary objective is to compare the efficacy of mirtazapine, ondansetron and placebo on nausea and vomiting measured with PUQE-24 score in patients with HG.
The secondary objective is to compare the efficacy of mirtazapine, ondansetron and placebo in patients with HG when measured by outcomes other than PUQE-24 score as well as PUQE-24 score outcomes not included in the primary objective.

Trial design:

A double-blind randomized placebo-controlled multi-centre trial. The subjects will be randomised 1:1:1 to receive treatment with either mirtazapine, ondansetron or placebo. The intervention will last for 14 days.

Trial population:

We will include approximately 180 pregnant women with hyperemesis gravidarum, 60 in each group.

Main inclusion criteria:

- Pregnant with gestational age between 5+0 and 19+6
- Nausea and vomiting without other obvious reason
- PUQE-24[1] score ≥13 OR
  PUQE-24 score ≥7 AND
  1. weight loss >5% OR
  2. hospitalization due to nausea and vomiting of pregnancy
- Normal singleton pregnancy

Main exclusion criteria:

- Mola-, multiple- or non-vital pregnancy
- Congenital long QT-syndrome
- Ongoing treatment with antidepressant medication
- Not able to take medicine orally
- Not able to understand spoken and written Danish

Power calculation:

The trial was originally dimensioned to find a difference in PUQE-24 score of 2 with a power of 80% based on a standard deviation of 3 in change in PUQE-24 score, a type 1 error rate of 2,5% for each of the original co-primary endpoints and 35% drop out. This resulted in a required 58 participants in each of the 3 groups, which was round up to 60 participants. We thus aimed to include 180 participants in total.
However, the slow recruiting rates resulted in the decision to discontinue the trial before the aimed 180 participants included, and the co-primary endpoints have been narrowed down due to the reduced statistical power.

Trial endpoints:

Final primary endpoints tested hierarchically:

- change in PUQE-24 score from baseline to day 2 (short term) in the mirtazapine group versus the placebo group.

[1] PUQE-24 (Pregnancy-Unique Quantification of Emesis and Nausea) is a validated scoring system used to grade the severity of nausea and vomiting of pregnancy. It ranges 3-15.

- change in PUQE-24 score from baseline to day 2 (short term) in the ondansetron group versus the placebo group (this will be tested only if mirtazapine versus placebo on day 2 is significant).
- change in PUQE-24 score from baseline to day 2 (short term) in the mirtazapine group versus the ondansetron group (this will be tested only if ondansetron versus placebo on day 2 is significant).

Secondary endpoints include:

- change in PUQE-24 score from baseline to day 14(+/-) (long term) in the mirtazapine group versus the placebo group (this will be tested only if mirtazapine versus placebo on day 2 is significant).
- change in PUQE-24 score from baseline to day 14(+/-) (long term) in the ondansetron group versus the placebo group (this will be tested only if ondansetron versus placebo on day 2 is significant).
- Change in PUQE-24 score on long term in the mirtazapine group versus the ondansetron group.
- Area under the curve for PUQE-24 score during the intervention.
- PUQE well-being score.
- Number of daily vomiting episodes.
- Nausea VAS.
- Side effects.
- NVPQOL (health-related quality of life for nausea and vomiting during pregnancy).
- HELP (hyperemesis level prediction).
- EQ-5D (health-related quality of life).
- Modified PSQI (Pittsburg Sleep Quality Index)
- Patient satisfaction with treatment VAS.
- Patient consideration of termination of pregnancy.
- Request for dosage increase.
- Request for continuation of trial medication after end of intervention.
- Use of rescue medication.
- Days on sick leave.
- Amount of treatments with i.v.-fluids.
- Days of hospitalisation.
- Weight change.
- Pregnancy outcome (birth weight, gestational age, APGAR, umbilical cord pH, placenta weight, sex, mode of delivery, delivery complications, hospitalisations on neonatal ward during the first month, congenital malformations, loss/termination of pregnancy).
- Occurrence of treatment failure.

Trial medication:

All trial medication will be provided by Glostrup Apotek in accordance with GMP. The investigational medicinal products (IMPs) will be similar in appearance, smell and taste due to the IMPs being encapsulated in gelatine. Each subject will be given two containers one labelled ‘Morning’ and one labelled ‘Bedtime’ to be administered orally in the morning and before bedtime for 14 days.

In the mirtazapine group, the ‘Morning’ medication will contain placebo and the ‘Bedtime’ medication will contain mirtazapine 15 mg.
In the ondansetron group both the ‘Morning’ and the ‘Bedtime’ medication will contain ondansetron 8 mg.
In the placebo group both the ‘Morning’ and the ‘Bedtime’ medication will contain placebo.

All subjects will be allowed to take metoclopramide as rescue medication, and IMP dosage increase will be possible on Day 7(+/-1) in case of insufficient effect during the first week.

Trial schedule:

Planned first subject first visit:	March 2019
Planned last subject randomised:	July 2022
Planned last subject last visit:	July 2022
End of trial:	May 2023

2. Flowchart

	Screening Day -2-0	Visit 1 Randomisation Day 0	Online questionnaire Day 1-14	Visit 2 Day 7(+/-1)	Visit 3 Day 14(+/-1)	Day 15-20	Postpartum
Informed Consent	x						
In- and Exclusion criteria	x	x		x	x		
Demography	x						
Medical History	x						
Previous and concomitant medication	x						
Changes in concomitant medication		x		x	x		
Vital signs (Height[1], weight, BP and pulse)	x	x		x	x		
Physical examination	x						
Ultrasound	x						
PUQE 24-score	x	x	x	x	x		
Nausea VAS, daily vomiting episodes		x	x	x	x		
Rescue medication			x	x	x		
NVPQOL, HELP, EQ-5D, PSQI		x		x	x		
Patient consideration of termination of pregnancy		x		x	x		
Patient satisfaction with treatment VAS				x	x		
Sick leave				x	x		
Hospitalization				x	x		
I.v. fluids				x	x		
Dosage increase				x			
Continuation of trial medication					x		
Side effects, adverse events			x	x	x	x	
Blood samples	x	x		x	x		
Urine sample	x						
ECG	x			x	x		
Randomisation		x					
Dispense of trial medication		x		x			
Drug accountability				x	x		
Birth Outcome							x

1) Height only at screening

3. Background

Refer to the Clinical Trial Protocol **3. Background/Rationale**.

4. Hypothesis

The hypothesis is that both mirtazapine and ondansetron reduce nausea and vomiting more than placebo in patients with hyperemesis gravidarum.

5. Objectives

5.1 Primary objective

The primary objective is to compare the efficacy of mirtazapine, ondansetron and placebo on nausea and vomiting measured with PUQE-24 score in patients with HG.

5.2 Secondary objectives

The secondary objective is to compare the efficacy of mirtazapine, ondansetron and placebo in patients with HG when measured by outcomes other than PUQE-24 score as well as PUQE-24 score outcomes not included in the primary objective.

6. Study Methods

6.1 Trial Design

A double-blinded randomized placebo-controlled multi-centre trial.

180 subjects will be randomised 1:1:1 to receive treatment with either

a) Mirtazapine oral tablet 15 mg at bedtime and placebo oral tablet morning,
b) Ondansetron oral tablet 8 mg morning and ondansetron oral tablet 8 mg at bedtime, or
c) Placebo oral tablet morning and at bedtime.

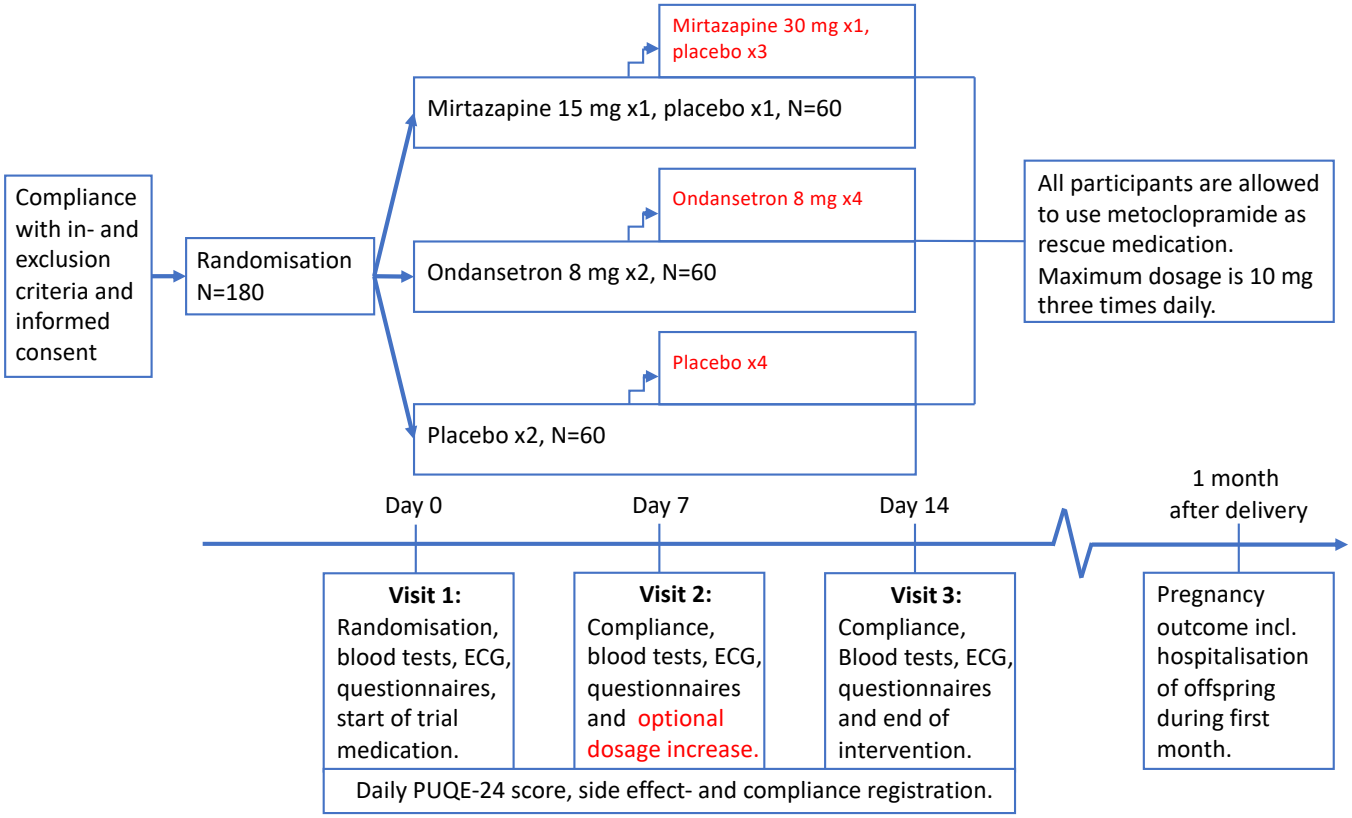
The trial subjects will receive treatment for 14 days.

A dosage increase will be optional on Day 7(+/-1). If desired the subject will then receive treatment with either

a) Mirtazapine oral tablet 30 mg at bedtime and placebo oral tablet morning, noon and late afternoon
b) Ondansetron oral tablet 8 mg morning, noon, late afternoon and at bedtime
c) Placebo oral tablet morning, noon, late afternoon and at bedtime

Data for the primary endpoint PUQE-24 score will be collected by the subjects completing online questionnaires related to their daily symptoms.

Figure 1 – Trial Design



6.2 Randomisation

Subjects will be randomised in a 1:1:1 ratio and stratified by site.
Glostrup Apotek will deliver the IMPs blinded to the study sites. Each treatment unit (containers sufficient for treating one subject during the intervention (two weeks)) will be individually marked with a randomisation number consisting of a combination of letters and numbers. Only Glostrup Apotek will know which randomisation number contains which type of trial medication.
After identification of an eligible subject, the subject will be assigned the lowest available randomisation number at that specific study site. The randomisation number will be entered in the database (REDCap) and the subject will subsequently receive the corresponding trial medication.

6.3 Blinding

A double-blind set-up will be used. Mirtazapine, ondansetron and placebo will be similar in appearance, smell and taste and will be packaged identically to maintain the treatment blind.
Information on type of trial medication corresponding to each randomisation number will be kept by Glostrup Apotek and thus will not be accessible to trial personnel involved in the conduct of the trial during the intervention. Neither the sponsor, the subject, nor the investigational site staff will know which treatment the subject is receiving.
If a subject desires to continue the IMP after the intervention, the randomisation code for that particular subject can be broken. Thus, for unblinded subjects data collected after unblinding (side effects and AEs 0-5 days after the intervention and pregnancy outcome) will not be blinded for the subjects or the clinicians, but the statistician will remain blinded.
The randomisation code for a particular subject can be broken during the intervention in a medical emergency or if knowledge of the IMP is necessary for the optimal treatment of the subject.

To break the randomisation code the investigational staff must call Glostrup Apotek which is available at all times of day and night.
Whenever a randomisation code is broken, the person breaking the randomisation code must record the time, date and reason as well as his/her initials in the source documents. The sponsor and investigator must be notified immediately after randomisation code break if this happens because of an emergency during the intervention. No matter the reason for unblinding, the subject must be discontinued from the trial (side effects and AEs 0-5 days after the intervention and birth outcome will still be collected).

6.4 Sample size

Sample size calculation is based on originally planned co-primary endpoints, which have since been changed.
The trial was dimensioned to detect a difference of 2 or more in the co-primary outcomes; change in PUQE-24 score from baseline to day 2 in the mirtazapine versus the placebo group, or in the ondansetron versus the placebo group. Each hypothesis for co-primary outcomes will be tested at significance level 2.5% in order to obtain at family-wise type I error of 5% based on a Bonferroni correction. The power calculation is based on a standard deviation of change in PUQE-24 score of 3 which is a conservative estimate based on a placebo-controlled trial with doxylamine-pyridoxine(1). Expecting dropout of 35%, 58 participants are required in each of the 3 groups to obtain a power of 80%. Rounding that up to 60 participants we aim to include 180 participants in total.

6.5 Framework

Two-sided equality test with the null hypothesis that the expected difference of the effects of mirtazapine and placebo, and ondansetron and placebo, respectively, is zero.

6.6 Interim analysis

There are no planned interim analyses.

6.7 Early termination of the trial

The Sponsor reserves the right to terminate the trial under the following conditions:

- Safety concerns
- If an interim statistical analysis shows that the trial has no scientific value or too low power
- If recruiting is not feasible or so slow that it results in significant delay in publication of results

If the trial is prematurely terminated or suspended, the investigator will promptly inform enrolled subjects and ensure appropriate therapy and follow-up. Furthermore, the investigator and/or sponsor will promptly inform the pertinent ethics committee and regulatory authorities.

6.8 Timing of final analysis

Final analysis of all endpoints will be performed after the last patient has finished the intervention. All endpoints except birth outcome will be analysed shortly after the last subject has finished the intervention. Analysis of birth outcome will be performed at least one months after the last patient has given birth to allow for registration of possible hospitalization of the new-born.

6.9 Timing of outcome assessments

Baseline characteristics will be assessed on Day 0 (Visit 1).
PUQE-24 score, daily vomiting episodes, nausea VAS, administration of IMP and rescue medication and possible side effects will be registered daily by subjects via online questionnaires.
NVPQOL, HELP, EQ-5D, PSQI, patient satisfaction with treatment (only Visit 2 and 3), patient consideration of termination of pregnancy, number of days on sick leave, amount of treatments with i.v.-fluids. and number of days of hospitalisations will be registered via online questionnaires at Visit 1, 2 and 3.

Weight will be measured by trial personnel at Visit 1, 2 and 3, and request for dosage increase and continuation of trial medication as well as additional side effects will be registered by trial personnel at the visits.
The visits on Day 7(Visit 2) and Day 14(Visit 3) can be scheduled for day 6-8 and day 13-15 respectively.
Possible treatment failure will be registered when it occurs.
Birth outcome will be registered at least one month after delivery.

7. Statistical Principles

7.1 Level of statistical significance and adjustment for multiplicity

Statistical significance will be claimed if the null hypothesis is rejected with an alpha level of 0.05 (two-sided) for the primarily tested primary endpoint. This will ensure a family-wise type I error of 5% based on a Bonferroni correction. A hierarchical testing procedure will be followed for the remaining co-primary endpoints to control the type 1 error. If a test fails to confirm a given hypothesis all subsequent tests will be regarded as explorative.
All secondary outcomes are per definition explorative.

7.2 Confidence intervals

Results will be reported with 95% confidence intervals and will be regarded as statistically significant when not including 0 or 1 on a relative or absolute scale, respectively.

7.3 Adherence and protocol deviations

Adherence to intervention will be monitored at Visit 2 and 3 and via online questionnaires. Adherence will be presented as part of the results.
Deviations from the protocol including violations of inclusion/exclusion criteria will be assessed as “minor” or “major.” Major protocol deviations are defined as those deviations from the protocol likely to have an impact on the perceived efficacy and/or safety of study treatments, and therefore, would result in a patient being excluded from the per protocol population. Major deviations will be identified and finalized prior to database lock and documented. Examples of major protocol deviations and any action to be taken regarding the exclusion of patients or affected data from specific analyses are outlined below.

7.4 Major Protocol Deviation Population Exclusions

Randomized but did not take any study medication: Exclude from statistical analyses and per protocol analyses
Consumed the incorrect study medication per randomization scheme: Exclude from per protocol analyses
Major Protocol Violations: Exclude from per protocol analyses
Non-compliant (< 80% compliant with intended study medication): Exclude from per protocol analyses
Took prohibited medication that impacts efficacy: Exclude from per protocol analyses.

7.5 Analysis populations

The primary analyses will be performed on all subjects randomized in the trial (intention to treat). Sensitivity analyses will be performed on the per protocol population.

8. Trial Population

8.1 Screening data

Possible subjects are pregnant women suffering with HG.
We will collect data on all patients who were informed about the trial as to describe the representativeness of the trial population. If patients do not wish to be enrolled in the trial, the reason will be registered.

8.2 Eligibility

Inclusion criteria:

- Written informed consent obtained before any trial related procedures are performed
- Female age >18 years
- Pregnant woman with gestational age between 5+0 and 19+6
- Nausea and vomiting without other obvious reason
- PUQE-24 score ≥13 OR

  PUQE-24 score ≥7 AND
  1. weight loss >5% OR
  2. hospitalization due to nausea and vomiting of pregnancy

  Compliance with this criterion on any given day within 3 days prior to screening (possibly not compliant with this criterion on day of randomization)

- Singleton pregnancy
- The subject must be willing and able to comply with trial protocol

Exclusion criteria:

- Mola pregnancy, multiple gestation or non-vital pregnancy
- Nausea and vomiting of other aetiology than NVP
- Allergic to selective 5-HT3-receptor antagonists
- Ongoing treatment with antidepressant medication
- Pre-existing diagnosis of chronic kidney disease, diabetes type 1 or 2, significant cardiac disease (incl. long QT syndrome), epilepsy, HIV. In case of other pre-existing conditions subjects might be excluded based on individual assessment by an MD
- Elevated liver enzymes (ALAT >150)
- Elevated creatinine (>100)
- ECG showing long QT-syndrome (QTc >460msek)
- Alcoholism; subject drinks more than recommended by health authorities
- Not able to take medicine orally
- Not able to understand spoken and written Danish
- Participation in another investigational drug trial within current pregnancy

8.3 Recruitment

A flow chart of recruitment will include number of subjects who a) received oral information on the trial, b) consented to participation c) were included d) were randomised to each of the interventions e) discontinued the intervention early f) completed the intervention g) were included in the final analysis.

8.4 Withdrawal/follow-up

Number, time and reason for discontinuation (withdrawal from intervention), lost-to-follow-up and final number analysed will be reported. Discontinued subjects and lost-to-follow-up will not be replaced.

8.5 Baseline patient characteristics

Baseline characteristics to be summarized in Table 1:

- History of HG in previous pregnancy
- Maternal age
- Weight, kg
- Pre-pregnancy weight, kg
- BMI, kg/m2
- Gestational age at randomisation, w+d
- PUQE-24 score at randomisation
- NVPQOL at randomisation
- HELP score at randomisation
- EQ-5D at randomisation
- 7 days PSQI at randomisation

9. Analysis

9.1 Outcome definitions

Primary endpoint:
The co-primary endpoints will be tested hierarchically in the listed order. Only if a result is statistically significant, will the subsequent analyses be regarded as primary:

- Change in PUQE-24 score from baseline to day 2 (short term) in the mirtazapine group versus the placebo group.
- Change in PUQE-24 score from baseline to day 2 (short term) in the ondansetron group versus the placebo group (this will be tested only if mirtazapine versus placebo on day 2 is significant).
- Change in PUQE-24 score from baseline to day 2 (short term) in the mirtazapine group versus the ondansetron group, non-inferiority (this will be tested only if mirtazapine versus placebo and ondansetron versus placebo on day 2 is significant).
- Change in PUQE-24 score from baseline to day 2 (short term) in the mirtazapine group versus the ondansetron group, superiority (this will be tested only if mirtazapine versus placebo and ondansetron versus placebo on day 2 is significant).

The change in PUQE-24 score is calculated as baseline score minus Day 2 score. Baseline values will be recorded at randomisation.

PUQE-24 score (Pregnancy-Unique Quantification of Emesis and Nausea) is a validated scoring system used to grade the severity of NVP. PUQE-24 score consists of three elements:

1) Number of hours with nausea within the last 24 hours (from none at all to more than 6 hours),
2) Number of vomiting episodes in the last 24 hours (from 0 to >6 times), and
3) Number of retching episodes without vomiting within the last 24 hours (from 0 to >6 times).

Each element gives a sub-score from 1 to 5 points and the total score is calculated by adding the three sub-scores giving a total score ranging from 3 to 15.
The condition is graded as mild if the score is 3-6, moderate with a score of 7-12 and severe if the score is 13-15(3). See Appendix 1.

Secondary endpoints:

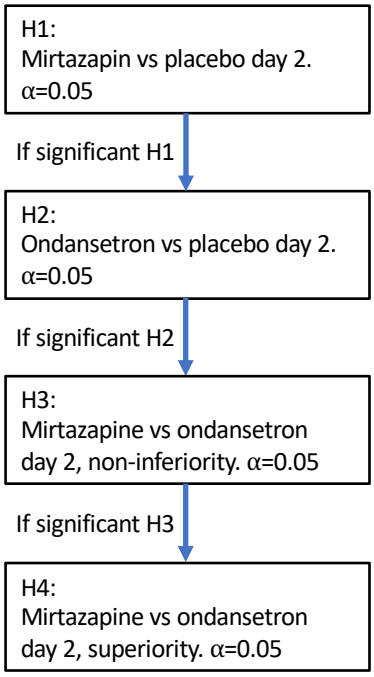
- Change in PUQE-24 score from baseline to day 14(+/-) (long term) in the mirtazapine group versus the placebo group.
- Change in PUQE-24 score from baseline to day 14(+/-) (long term) in the ondansetron group versus the placebo group.
- Change in PUQE-24 score from baseline to Day 14(+/-1) in the mirtazapine group versus the ondansetron group.
- Area under the curve for PUQE-24 score during the intervention in the three different groups.
- Change in PUQE well-being score during the intervention in the three different groups.
- Change in daily nausea VAS during the intervention in the three different groups.
- Change in number of daily vomiting episodes during the intervention in the three different groups.
- Occurrence of side effects in the three different groups.
- Change in NVPQOL (health-related quality of life for nausea and vomiting during pregnancy) from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Change in HELP score (hyperemesis level prediction) from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Change in EQ-5D-5L (health-related quality of life) from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Change in modified PSQI (Pittsburg Sleep Quality Index) from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Patient satisfaction with treatment VAS on Day 7(+/-1) and Day 14(+/-1) in the three different groups.
- Change in patient consideration of termination of pregnancy from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Frequency of request for dosage increase in the three different groups (categorical).
- Frequency of request for continuation of trial medication after end of intervention in the three different groups (categorical).
- Use of rescue medication during the intervention in the three different groups.
- Number of days of sick leave during the intervention in the three different groups.
- Amount of treatments with i.v.-fluids during the intervention in the three different groups.
- Number of days of hospitalisations during the intervention in the three different groups.
- Weight change in kg from baseline to Day 7(+/-1) and baseline to Day 14(+/-1) in the three different groups.
- Pregnancy outcome (birth weight, gestational age, APGAR, umbilical cord pH, placenta weight, sex, mode of delivery, delivery complications, hospitalisations on neonatal ward during the first month post-partum, congenital malformations, loss/termination of pregnancy). This endpoint will not be double blind on all patients since patients can be unblinded after end of the intervention. The statistician will remain blinded.
- Frequency of and time to treatment failure in the three different groups (categorical).

9.2 Analysis methods

Primary analyses
The primary analysis will consist of a linear mixed model with intervention group and baseline PUQE-24 score as covariates adjusted for site due to the stratified design. Intention to treat analysis will be performed where participants will be included in the analysis irrespectively of their adherence to the protocol.
A hierarchical testing procedure will be used to control the type 1 error for the primary endpoints (Figure 2). Statistical significance will be claimed if the null hypothesis is rejected at the alpha level of 0.05 (two-sided), and subsequent comparisons will be regarded as explorative if a test fails to reach a statistically significant difference. Thus initially, significance of difference in PUQE-24 score change from baseline to day 2 in the mirtazapine versus the placebo will be tested (H1). If significance is obtained a test will be performed on difference in PUQE-24 score change from baseline to day 2 in the ondansetron versus the placebo group (H2). If significance of H2 is obtained

difference in PUQE-24 score change from baseline to day 2 in the mirtazapine versus the ondansetron group will be tested as non-inferiority (H3), and if H3 reaches statistical significance difference in PUQE-24 score change from baseline to day 2 in the mirtazapine versus the ondansetron group will be tested as superiority (H4).
If one of the co-primary endpoints fails to reach statistical significance all subsequent comparisons will be regarded as exploratory and only point estimates and confidence intervals will be reported.

Figure 2 – Flowchart of primary analyses



The treatment effect for the primary outcomes will be presented as the mean difference between the groups in PUQE-24 score change from baseline to day 2. Results will be reported with 95% confidence intervals.

Secondary analyses
Intention to treat analysis will be performed on secondary outcomes where participants will be included in the analysis irrespectively of their adherence to the protocol.
Continuous data will be analysed using a linear mixed model and results will be reported as point estimates with 95% confidence intervals. Categorical data will be analysed using and chi-squared test.
Secondary analyses include per protocol analyses of the primary outcome.

Other analyses
Subgroup analyses will be performed by stratifying on the baseline variables PUQE-24 score, number of daily vomiting episodes, gestational age and concomitant medication with first line treatment.
Per protocol analysis including only participants who completed the intervention will be performed.
All data will be described including data-incompleteness as well as reasons for data-incompleteness. Data will be analysed blinded by the statistician. Any changes to the statistical analysis plan will be described in any future publications.

9.3 Missing data

Missing data will be imputed using multiple imputation where the imputation model will include all available baseline variables (listed below) and intervention group as well as outcome values (listed below) up till the time of drop out. Twenty five datasets will be imputed using the R-package mice (Multiple Imputation by Chained Equations)(3) and a seed of 3741.

Available baseline variables:

- Background data: Age, pre-pregnancy weight, height, BMI, occupation.
- Obstetric history: Parity, previous HG (resulting in sick leave, treatment, hospitalisation, >5% weight loss, termination of pregnancy)
- Current pregnancy: Gestational age, baseline weight loss.
- Symptom measures: PUQE-score, well-being-score, daily vomiting episodes, nausea VAS, NVPQOL, HELP, EQ-5D-5L, modified PSQI, consideration of termination of pregnancy
- Concomitant medication

Outcome values:

- PUQE-score, well-being-score, daily vomiting episodes, nausea VAS, NVPQOL, HELP, EQ-5D-5L, modified PSQI, consideration of termination of pregnancy, satisfaction with treatment, request for dosage increase, days of sick leave during the intervention, treatments with i.v.-fluids, days of hospitalisations during the intervention, weight change in kg from baseline to Day 7, use of rescue medication.
- Concomitant medication

9.4 Harms

Refer to the Clinical Trial Protocol **11. Assessment of Safety** for information on definitions of adverse events (AEs) and how they will be reported.

9.5 Statistical software

R will be used for the statistical analysis.

10. References

1. Koren G, Clark S, Hankins GD V, Caritis SN, Miodovnik M, Umans JG, et al. Effectiveness of delayed-release doxylamine and pyridoxine for nausea and vomiting of pregnancy: A randomized placebo controlled trial. Am J Obstet Gynecol [Internet]. 2010;203(6):571.e1-571.e7. Available from: http://dx.doi.org/10.1016/j.ajog.2010.07.030
2. EQ-5D. EQ-5D [Internet]. Available from: https://euroqol.org/eq-5d-instruments/
3. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3):1–67.
4. DSOG. Hyperemesis Gravidarum - Guideline - DSOG [Internet]. 2013. Available from: http://gynobsguideline.dk/sandbjerg/Hyperemesisgravidarum.pdf

Appendix 1

PUQE-score (Pregnancy Unique Quantification of Emesis and Nausea- score)

Marker det svar, der bedst beskriver dine symptomer de sidste 24 timer				
1. Hvor lang tid har du følt dig forkvalmet, i løbet af de sidste 24 timer?				
Slet ikke	≤ 1 time	2-3 timer	4-6 timer	> 6 timer
(1)	(2)	(3)	(4)	(5)
2. Hvor mange gange har du kastet op, i løbet af de sidste 24 timer?				
Ingen opkastninger	1-2 gange	3-4 gange	5-6 gange	≥ 7 gange
(1)	(2)	(3)	(4)	(5)
3. Har du haft opkastningsbevægelser (uden at der kommer noget med op), i løbet af de sidste 24 timer?				
Nej	1-2 gange	3-4 gange	5-6 gange	≥ 7 gange
(1)	(2)	(3)	(4)	(5)

PUQE-24 score: Mild ≤ 6, Moderat = 7-12, Svær = 13-15.

På en skala fra 0 til 10 hvordan vil du vurdere dit generelle **velbefindende**? ____
0 (værst tænkelig), 10 (så godt, som du havde det før, du blev gravid)

Dansk oversættelse af: Ebrahimi et al. Nausea and Vomiting of Pregnancy: Using the 24-hour Pregnancy-Unique Quantification of Emesis (PUQE-24) Scale. J Obstet Gynaecol Can. September 2009;31(9):803-7 (4)