Author's response to reviews

Title: The HAPPY study (Holistic Approach to Pregnancy and the first Postpartum Year): design of a large prospective cohort study.

Authors:

Sophie EM Truijens (s.truijens@mmc.nl)
Margreet Meems (m.meems@uvt.nl)
Simone MI Kuppens (simone.kuppens@cze.nl)
Maarten AC Broeren (m.broeren@mmc.nl)
Karin CAM Nabbe (karin.nabbe@diagnostiekvoor.nl)
Hennie A Wijnen (h.wijnen@av-m.nl)
S Guid Oei (g.oei@mmc.nl)
Maarten JM van Son (m.vanson@uu.nl)
Victor JM Pop (v.j.m.pop@uvt.nl)

Version: 3 Date: 11 August 2014

Author's response to reviews:

Reviewer 1

(1) Abstract, Methods, Discussion: This manuscript details a study that is already ongoing. The language of the text throughout the manuscript should reflect that work on this study already has begun and that more work will continue.

We agree with the reviewer and adapted the text accordingly. To reflect that this study has already begun, the language of the text about the recruitment and inclusion phase has been rewritten from future tense into present tense throughout the text (most frequently the words ‘will be’ have been replaced by ‘is’ or ‘are’). Furthermore, we provided the recruitment timeframe to page 14.

(2) Introduction, Maternal and foetal thyroid function during pregnancy: The authors stated, “However, up until now, most studies looking at foetal outcome and maternal thyroid function did not use intra-uterine foetal parameters to evaluate the CNS development, such as ultrasound data.” Please provide references to support this statement.

“However, as shown in a review, most studies looking at foetal outcome and maternal thyroid function did not use intra-uterine foetal parameters to evaluate the CNS development, such as ultrasound data [Cassar et al., 2013].”


(3) Introduction, Breastfeeding: The authors also indicated that, “Determinants that predict start and continuation of breastfeeding are of multifactorial origin including education level, ... Also, absence of facilities at work to breastfeed or to express milk might result in quitting breastfeeding.”
Please provide references to support this statement.

To support these statements we included a systematic review of De Jager et al. (2013), a Cochrane review of Abdulwadud & Snow (2012), and a review of Meedya et al. (2010) to the following sentences (page 11, line 1-5):

“Determinants that predict start and continuation of breastfeeding are of multifactorial origin including educational level, socioeconomic status, marital status, partner involvement, and maternal distress during gestation and postpartum [De Jager et al., 2013]. Also, facilities at work to breastfeed [Abdulwadud & Snow, 2012] and the perceived milk supply [Meedya et al., 2010] influence the duration of breastfeeding.”

(4) Introduction, Breastfeeding: “It is important to detect women at risk for early cessation of breastfeeding because simple antenatal intervention techniques (including involvement of the partner in decision making) can increase the likelihood to breastfeed.”

While I agree with the authors that antenatal intervention can increase the likelihood of breastfeeding initiation and duration of breastfeeding, I believe this statement oversimplifies the challenges some mothers face with breastfeeding (e.g. low milk production, infant weight loss while being exclusively breastfed, etc.) despite their willingness and continuous efforts to try to breastfeed their infants.

We agree with the reviewer that the eventual success of breastfeeding depends also on other factors (like low milk production or infant weight loss). To state this more clearly we changed the text at page 11 accordingly: “It is important to detect women who experience difficulties with the challenges of breastfeeding (e.g. low milk production, infant weight loss while being exclusively breastfed, etc.) despite their willingness and continuous efforts to try to breastfeed their infants. Antenatal intervention techniques (including involvement of the partner in decision making) can increase the likelihood to breastfeed (Kummeling et al., 2005).

(5) Introduction, Breastfeeding: “A recent study from Spain showed an odds ratio…” Please include confidence intervals to help determine the precision of estimates on the odds of withdrawing breastfeeding in this research.

We agree with the reviewer and added the 95% confidence intervals to page 11, line 24-25.

(6) Methods/design, Design and setting: “Furthermore, the two obstetric departments of two large hospitals in the area will collaborate.” Please elaborate the role(s) of the OB departments (presumably to collect infant and OB data).

We agree and adapted the text accordingly: “The care providers of the obstetric departments collect umbilical cord blood after delivery and provide the fully completed obstetric record forms for data analysis after written informed consent of the participants” (page 14, line 6-8).

(7) Methods/design, Recruitment: Please provide recruitment timeframe.
We agree and provided the recruitment timeframe to page 14: “Recruitment takes place by the midwives of 17 primary care community midwife practices in the area of South-East Brabant, from January 2013 until July 2014.” Furthermore, we started the Recruitment section with: “Since January 2013, …” to indicate the timeframe.

(8) Methods/design, Recruitment: This study excludes women with gemelli pregnancy. Please verify and indicate on the manuscript that women with other higher ordered pregnancies (triplets, quadruplets, etc.) are also excluded.

Following the reviewer’s suggestion, we added to page 14 (line 17-18): “Exclusion criteria in this study are: gemelli pregnancy (or higher order pregnancies), …”

(9) Methods/design, Data collection: Please provide details on how missed clinic visits, non-response for questionnaire assessments (entire questionnaire, not selected items within a questionnaire), and participant lost-to-follow-up are handled in this study.

We especially choose for linear mixed model analysis (Mallinckrodt et al., 2003) as explained at page 22, to deal with missing data (page 22, line 1 to 5).

(10) Methods/design, Materials/Measures: The authors indicated in the Introduction section that a subset of maternal-fetal pairs will be chosen for the analysis on the relationship between maternal and fetal thyroid data. Please describe selection methodology (n and how. E.g. Simple random sample?) on how the pairs will be chosen.

We agree with the reviewer and adapted the text at page 9 (line 16-20) accordingly:

“Within the HAPPY study, pairs of maternal and neonatal thyroid data will be sampled to investigate whether poor maternal thyroid function has neonatal thyroid function consequences. From previous research in the same area it is known that because of logistic difficulties (home delivery, referral during labour from home to hospital, delivery in different hospitals) in about 70% of the participating women umbilical cord blood will be obtained (Kuppens et al., 2011)”

(11) Methods/design, Statistical analyses: Given CTS is one of the primary study outcomes, the authors should describe analytical techniques that will be used to answer research questions pertaining to CTS.

We agree with the reviewer and extended the information about how to analyse the CTS research questions. The questions about severity of CTS will be answered using the score on the 19-item Boston CTS Questionnaire (BCTQ), with its two subscales symptom severity (11 items) and functional status (8 items) (page 19). Furthermore at page 19 we added: “To examine CTS, the question: “Did you suffer any of the following symptoms during pregnancy: pain, tingling sensations, or numbness in hands or wrists?” is used including the three key symptoms of CTS.” (page 19, line 19-21).

Furthermore, we added to page 21: “Regarding CTS, a multiple logistic
regression analysis will be performed, with the presence of self-reported CTS (based on the question: “Did you suffer any of the following symptoms during pregnancy: pain, tingling sensations, or numbness in hands or wrists?”) as the dependent variable and fluid retention during pregnancy as independent variable, adjusted for age, parity, BMI, and depression. Furthermore, a multiple linear regression analysis will be performed, with BCTQ score (indicating CTS severity) as dependent variable and fluid retention during pregnancy as independent variable, adjusted for age, parity, BMI, and depression.” (page 21, line 20-26)

(12) Methods/design, Statistical analyses: How or what criteria will be used by the authors to determine which confounders will be included into each final multivariable models?

It is beyond the scope of the design paper to describe all of the models (including the relevant confounders) that will be used answering all the different research questions. However, we added to the page 22 (Method section – Statistical analyses): “For answering the various research questions, different models will be used containing different dependent and independent variables. Confounding variables will be entered into the model according to the literature. The variables age, parity and educational level will be used as confounders in all the models. For example: in examining determinants of start/continuation of breastfeeding during the early postpartum period, obstetric outcome will be entered into the regression as a confounder, while for assessing determinants of CTS, the variable BMI will be included as confounder.” (page 22, line 6-12).

Minor Essential Revisions - The author can be trusted to make these.

(12) Table 1: Define each abbreviation in footnote of table.

We added the abbreviations with definitions of the syndromes and questionnaires in Table 1.

Reviewer 2
- Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore).

We thank the reviewer for his exhaustive comments on our manuscript. We hope that we respectfully answered his comments. In the case that we changed the original text according to these comments, we specifically mentioned this in this letter.

(1) I personally would not recommend the Generalised Anxiety Disorder Scale as this is specifically used for identifying clinical cases of anxiety disorders (which will possibly be excluded as considered high risk by criteria given in the manuscript). I would suggest a more general measure of anxiety that has been routinely studied with pregnant cohorts, such as the STAI. This is more likely to pick up greater variability in responses.

We agree with the reviewer and have personal experiences using the STAI (Brouwers, Van Baar & Pop, 2001). We found little evidence for differences in state and trait anxiety scores in relation to independent variables during
pregnancy. We prefer the use of a scale of general anxiety based on the seven symptoms of anxiety according to the DSM-IV criteria, to compare with a pregnancy-specific distress scale (TPDS). Moreover, the GAD-7 is much more user-friendly than the STAI.

(2) You may want to incorporate a measure of additional activities pregnant women engage in such as antenatal yoga and pilates. It is likely that midwives may anecdotally recommend yoga to sub-clinical stressed/anxious/depressed mothers and those with carpal tunnel syndrome.

We agree, and we ask women whether they joined antenatal courses. In The Netherlands, these courses are standardized, including yoga exercising and are followed by more than 80% of the pregnant women.

- Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Introduction

(3) Due to the multi-dimensional nature of the project, a high number of concepts and mechanisms are described to the reader (e.g. mental health, physiological function, hormonal function etc). It is unlikely any reader will be entirely familiar with every aspect of this project, and also within the context of pregnancy. Therefore I think it is important that the authors are explicit on every term introduced. For myself, I am not overly familiar with thyroid function so concepts such as hyperthyroidism could benefit from a brief explanatory sentence. In contrast psychological health is more my speciality and so some of the concepts may be obvious to me but not other readers. Nothing too in depth but a simple sentence. Many proceeding comments touch upon this main criticism.

We agree with the reviewer and we added some words to the Introduction section to explain the terms that we introduce (page 5-8):

- “…hyperemesis gravidarum (HG), a potentially life threatening but rare condition (0.5-2%) which is characterized by nausea, vomiting, severe dehydration and weight loss resulting in hospitalisation” (page 5, line 12-13)
- “…high levels of the hormone human chorionic gonadotropin (HCG)” (page 5, line 10)
- “An underactive thyroid, called hypothyroidism (HT), has long time been reported…” (page 6, line 5-6)
- “…antibodies against thyroid peroxidase enzyme (TPO-Ab). The TPO-Ab are involved in the autoimmune thyroiditis process resulting in an inactive thyroid.” (page 8, line 4-5)

(4) ‘Maternal signs’ is an odd term that is frequently used throughout the manuscript.

I’m not entirely sure what the authors are implying and this needs to be changed.

We agree with the reviewer and changed the term ‘signs’ into ‘complaints and symptoms’ throughout the text.
What differentiates Hyperemesis gravidarum from regular morning sickness?

Please provide brief definition.

We added to page 5 (line 11-13): “In its extreme form, NVP or ‘morning sickness’ can manifest itself as hyperemesis gravidarum (HG), a potentially life threatening but rare condition (0.5-2%) which is characterised by nausea, vomiting, severe dehydration and weight loss resulting in hospitalisation (Ebrahimi et al., 2009).”

Be more descriptive on what morphological factors may influence likelihood of CTS.

Carpal tunnel size has been reported to be a risk factor for CTS. However, the evidence is not very abundant in the literature. Therefore, after reconsidering the literature, we decided to omit ‘morphological factors’ as a possible risk factor for CTS.

Is there any previous evidence of the effects of CTS on quality of life? This is an important pathway in your model.

Previous studies have provided evidence for a negative effect of CTS on the quality of life. To support this with literature, we added two references about this topic (page 5, line 24):


Moreover, the Boston CTS questionnaire (BCTQ) especially asks for functional impairment, interfering with quality of life (e.g. writing, bathing, dressing, opening a jar).

The literature review of anxiety and depression in pregnancy is severely lacking with little description and an unawareness of the studies showing how distress is higher in pregnancy and it’s fluctuations across pregnancy (Newham and Martin, 2013). No references have been provided.

We agree with the reviewer and included additional information and references regarding previous studies on anxiety and depression during pregnancy.

Page 6 (line 14-22): “The occurrence of distress during pregnancy is subject to much debate: is there an increasing number of women suffering from depression and anxiety throughout gestation? The literature is inconclusive, mostly because there are very few studies that repeatedly assessed anxiety and depression. Most studies measured these symptoms only once during pregnancy, a few studies twice, but the systematic review of Bennett et al. (2004) and the review of
Alder et al. (2007) found no appropriate studies that assessed symptoms of depression and anxiety during all trimesters of pregnancy. Furthermore, some researchers suggested symptoms of depression and anxiety to be stable throughout pregnancy (Heron et al., 2004), while others stated that maternal mood and well-being fluctuate across pregnancy (Newham & Martin, 2013).

(9) Pg7, line 17-19. This is an inefficient description of the link to infant development. This is a prime area to highlight the importance of the study as links may be due to thyroid dysfunction causing biological alterations to the child independent of maternal mood or vice versa. This should be the core of literature discussion to justify the postpartum follow up.

We agree with the reviewer, and we added to the text – however, in the thyroid section, page 9 - : “Moreover, adding a biological variable (thyroid function) to a model with predominantly psychological variables (maternal depression and anxiety during pregnancy), may elucidate on the multifactorial origin of impaired infant development as has been published before by our research group (Pop et al., 1995; 2003)”. (Page 9, line 2-5).

(10) Pg7, Line 24. More explanation of what thyroid peroxidase enzyme (TPO-Ab) is and its implications is needed.

In order to explain it more clear we added to page 8 (line 4-5): “…antibodies against thyroid peroxidase enzyme (TPO-Ab). The TPO-Ab are involved in the autoimmune thyroiditis process resulting in an inactive thyroid”.

(11) The section ‘Antenatal Wellbeing’ begins by describing physiological conditions of nausea and carpal tunnel syndrome which seems incongruent as wellbeing is primarily about psychological mood.

The section ‘antenatal well-being’ is about both physical and psychological well-being, but we agree with the reviewer that we should state this clear. Therefore, this section is now called “Antenatal physical and mental well-being” (page 5, line 2).

(12) Overall I really think you should start the introduction with a brief description of the model and then describe each construct of the model so have an initial idea of how components are inter-related. So parts of Study objectives could go higher up. Unfortunately due to some of the sparse information on certain links (as described in comments above) the model doesn’t seem very cohesive so I think the argument of their inter-relations needs reinforcing early on.

We added to the statistic section (page 20-22) some sentences with regard to the different models that will be used for data analysis.

Methods

(13) Give full description for acronym of HAPPY in the manuscript as only given in the title.

Following the reviewer’s suggestion, we added an introductory sentence
including the acronym of HAPPY (page 12, line 6/7): “The HAPPY study (Holistic Approach to Pregnancy and the first Postpartum Year) contains several aims with regard to both pregnancy and the postpartum period.”

(14) Will cases of pregnancy loss be excluded if later in the pregnancy? Blood samples at 10-12 weeks may be informative for identifying possible biomarkers but I think this needs explaining even if ethics meant it was not possible.

Women with pregnancy loss are lost to follow-up. However, if the number of women with pregnancy loss will be sufficient for statistical power, post-hoc analysis will be performed using 12-weeks thyroid data as a possible predictor of pregnancy loss. It should be remembered that up to 80% of pregnancy loss is occurring before 12 weeks of gestation (before blood sampling).

(15) It is unclear whether it is the same blood sample that is routinely collected or an additional blood samples taken at the same time point (e.g. ‘Because all additional blood samples were obtained during regular blood assessments as part of regular obstetric care, the board confirmed that no additional approval of a medical committee was needed.’ / ‘Furthermore, at standardized blood assessment around 10-12 weeks and 26-30 weeks of gestation, an additional tube of blood is withdrawn for thyroid function assessment and HCG analysis’). This needs to be transparent and consistent.

We agree with the reviewer, and we further clarified our logistics:

Page 15 (line 8/9): “During regular blood assessments (10-12 weeks and 26-30 weeks of gestation) an additional tube of blood is withdrawn for thyroid function and HCG analysis.

(16) I think the response rate of 70% is optimistic if additional blood samples are required -even if taken at the same time points. Also it does rely on midwives actively recruiting and I think there should be some scepticism as it is not a single research midwife committed to data collection.

Furthermore also the reliance of postal responses will cause a high drop out rate. To simply state ‘Based on previous pregnancy-related research experiences of the last 20 years in this area, the response rate is expected to be high (70%)’ seems a bit optimistic and referencing is required to justify this.

We wish to disagree, based on over 25 years of experience in perinatal research, with special interest in thyroid function during gestation, we know that the response rate is 70-79%. (Pop et al., 1995; 2006; 2011; Bergink et al., 2011).

However, this refers to research in which one or even up to three additional venipunctures are asked (Kuppens et al., 2011). We know from these studies that a reason for not participating is an additional venipuncture. However, in the HAPPY-study no additional venipunctures are requested, which means that is it realistic to expect even higher response rates.

(17) The authors give an admirable calculation of expected sample sizes but I think the authors should try to anticipate for lower numbers, and possible non-normality of the data, by considering boot-strapping statistical techniques
which will robustly allow to make inferences when numbers do not meet expected levels and when wanting to combine parametric and non-parametric data in the regression analysis.

We agree with the reviewer and will use the appropriate statistical analyses when power problem emerge due to low numbers of cases (e.g. the prevalence of pre-eclampsia, or preterm birth).

(18) The term ‘all kinds of somatic symptoms’ is not very scientific and needs further clarification.

We agree with the reviewer. At page 16 (line 22/23) we changed it in “somatic symptoms which are frequently mentioned during physiological gestation.

(19) EPDS cut offs are lower than actually have been established and these need to be raised or at least rationalised in the context of research in this area(http://www.ncbi.nlm.nih.gov/pubmed/19298573).

We agree with the reviewer and added the EPDS cut-off according to Gibson’s review to page 19.

“A cut-off value of 9 will be used to detect a possible postpartum depression (Gibson et al.,2009).” (page 19, line 5/6).

(20) Also move the section of the EPDS in the postpartum to when discussing it in the antenatal and give appropriate cut-offs in the postpartum.

We agree with the reviewer and moved the section of the EPDS after the section of the EDS (page 19, line 1-6).

(21) Why will the psychometric properties of the new anhedonia subscale be assessed in only a sub-sample and not the entire sample? I’m against the idea of just fusing two measures together, and assessing the psychometrics of this should be a central aspect of the project.

We wish to disagree. Although there has been a recent paper using highly sophisticated analyses to confirm the three-dimensional structure of the E(P)DS (De Cock et al.,2011), we follow Tabachnick & Fidell (2007) that subscales of less than four items are questionable. For evaluating the psychometric properties of an extended EPDS-scale (n=17 items) including several additional anhedonia-items, based on the MASQ (Mood and Anxiety Symptoms Questionnaire), it is generally accepted that sample size of up to 170 subjects (based on n=17 items) is largely sufficient to conduct proper statistical analyses. Therefore, using subsamples of 300-500 subjects (for explorative as well as confirmative factor analyses) is appropriate. There is no question of fusing two measures together.

(22) Pg. 20, line 20 – occurrence of stressful life events. No description of how this will be measured.

We added to the text (page 20): “Furthermore, women are asked whether a negative event did occur during the first trimester of pregnancy or during the period between the previous and next assessment until twelve months
With the statistics I really think some form of pathway analysis needs to be performed, or mediation analysis to examine the inter-relations between the components of the model. I think the model in Figure 1 is central to the paper and rationale for the study. Without it, it can come across a little like a ‘data fishing’ exercise by how so many factors are examined.

We agree with the reviewer and we added to the text at page 21 (line 9-11): “In order to evaluate possible causality between different variables path analysis will be performed using structural equation modelling (SEM).”

- Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

(23) The issue of high risk cases is quite important in this protocol as it is described in Pg. 14, Line 5-9; ‘It should be noticed that in The Netherlands, 84% [24] of all pregnant women contact the primary care community midwife for a first antenatal control (usually between six to ten weeks of gestation). The other 16% who immediately contact the obstetrician at the hospital represents a relatively high risk group of women (with gemelli pregnancy or suffering from chronic diseases such as diabetes, thyroid dysfunction, psychiatric disorders) and by definition are not eligible for this study.’

First of all I am a little unclear at when a woman is defined as high risk. Does this occur at the first midwife session mentioned as I am unclear when they would be referred based on this description. International differences in practices require this is fully explained to the readers.

We agree with the reviewer and we added to the text (page 14/15): "The Dutch obstetric care system is organised in primary care - represented by independent midwives providing care to women with low-risk pregnancies - and secondary care, represented by hospital midwives and gynaecologists who are responsible for high-risk pregnancies. Management of 84% of all pregnant women starts in midwifery practices [The Netherlands Perinatal Registry, 2013]. The other 16% are high-risk. The high-risk population consists of women with gemelli pregnancy or with chronic diseases such as diabetes, thyroid dysfunction, hypertension, or psychiatric disorders; all exclusion criteria for participation. As 84% of all pregnant women will have at least one prenatal visit in a primary care midwifery practice (usually a first antenatal control between six to ten weeks of gestation), recruitment takes place by the primary care community midwives.” (page 14/15)

(24) Secondly, and more importantly, is that by excluding these high risk cases you are getting rid of those who are most likely to show variation in thyroid function. Similarly by getting rid of those with psychiatric disorders means you are unlikely to identify those with high anhedonia scores in your sample. I think that this needs serious consideration as if no association is found it is likely to be because you are excluding the cases where this is an identified problem.

We wished to disagree. Anhedonia is one of the key symptoms of depression, and there is no exclusion criteria mentioned in the text with regard to depression/anhedonia. Only those women with severe psychiatric disease postpartum.” (page 20, line 3-5).
(schizophrenia, borderline, or bipolar disorder) (page 14, line 18-19) are excluded. Therefore, there is no reason to suggest that we exclude women with severe anhedonia scores as this is part of the depression syndrome rather than schizophrenia, borderline or bipolar disorder.