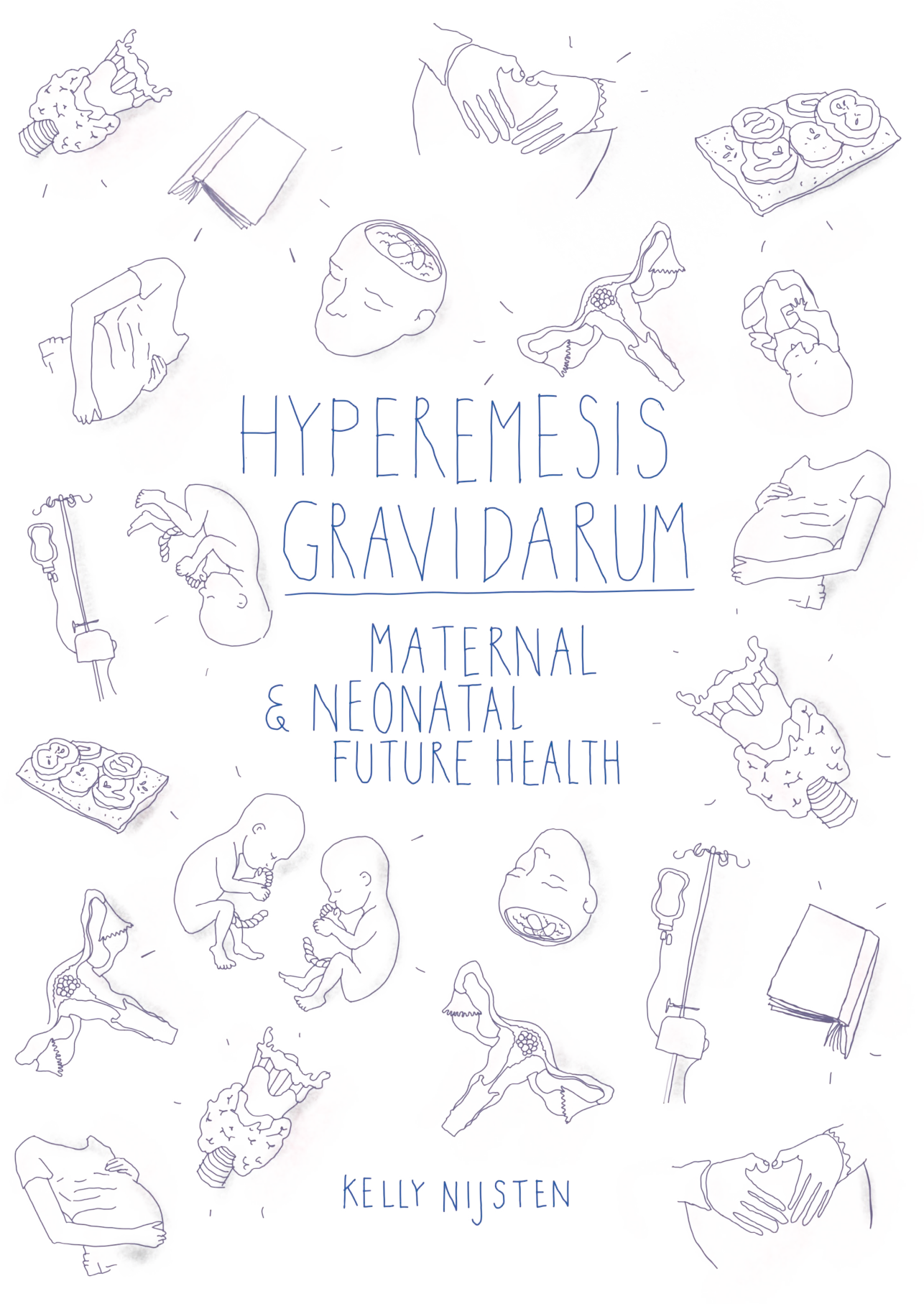


HYPEREMESIS GRAVIDARUM

MATERNAL & NEONATAL FUTURE HEALTH

KELLY NIJSTEN

Hyperemesis gravidarum: maternal and neonatal future health

Kelly Nijsten

Hyperemesis gravidarum: maternal and neonatal future health

Provided by thesis specialist Ridderprint, ridderprint.nl

Printing: Ridderprint
Cover design: Nadia Pepels, nadianena.com
Layout and design: Eduard Boxem, persoonlijkproefschrift.nl

Financial support for printing of this thesis was kindly provided and supported by the Amsterdam Reproduction & Development (AR&D) research institute, Nutricia Nederland B.V., Guerbet Nederland B.V., Goodlife Pharma and Bridea Medical B.V.

Hyperemesis gravidarum: maternal and neonatal future health

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek

ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 28 oktober 2022, te 16.00 uur

door Kelly Nijsten
geboren te Venlo

Promotiecommissie

Promotor:	prof. dr. T.J. Roseboom	AMC-UvA
Copromotor:	prof. dr. R.C. Painter	Vrije Universiteit Amsterdam
Overige leden:	prof. dr. M. Goddijn	AMC-UvA
	prof. dr. K.F.M. Joosten	Erasmus Universiteit Rotterdam
	dr. M.R. Soeters	AMC-UvA
	prof. dr. M.J.M. Serlie	AMC-UvA
	prof. dr. J. de Jonge	Vrije Universiteit Amsterdam
	prof. dr. K.W.M. Bloemenkamp	Universiteit Utrecht

Faculteit der Geneeskunde

Table of contents

CHAPTER 1

General introduction

Hyperemesis gravidarum: what is it?

Nausea and vomiting symptoms are common in early pregnancy and occur in the majority of pregnancies.[1] Hyperemesis gravidarum (HG) is a more serious condition, consisting of intractable vomiting in pregnancy often leading to dehydration, electrolyte disturbances and weight loss.[1] HG prevalence estimates range from 0.3 to 3%, depending on definition, which would amount to each year around 2000 pregnant persons suffering with HG in the Netherlands and up to 1.5 million worldwide .[2-4] Persons with HG frequently require hospital admission, with HG being the most common reason for hospitalisation in the first half of pregnancy.[5, 6] Symptoms often subside by about 20 weeks gestation, but some people suffer with HG until delivery.[7, 8] Until recently, there was no consensus on the definition of HG, leading to a diversity in patient characteristics and symptom severity in various studies, which frustrated attempts to aggregate data in meta-analyses as well as leading to lack of uniformity in clinical care.[9] Subsequently, this has led to research waste due to the inability to aggregate available evidence. In 2021 an international consensus definition was published, developed by a multi-stakeholder group, including health care professionals, researchers and HG patients and their families, using a modified Delphi procedure.[10] HG is now defined as a condition characterized by severe nausea and/or vomiting, an inability to eat and/or drink normally and which strongly limits daily activities, in which symptoms started in early pregnancy (before a gestational age of 16 weeks).[10] The agreement on a uniform definition of HG might eventually lead to a larger proportion of the pregnant population being diagnosed with HG.

Aetiology

Back in the 19th and early 20th century, HG was considered to be psychological disease. Although this stigma still remains nowadays, several other theories about the aetiology of HG have been studied since the end of the 20th century. Human chorionic gonadotropin (hCG) is most often hypothesized to be causally involved with HG, since hCG levels increase during the first trimester, similar to the peak incidence of nausea and vomiting in pregnancy.[11] A meta-analysis however, found the association between hCG and HG to be inconsistent.[11] Thyroid dysfunction has also been thought to contribute to the etiological pathway to HG, due to the similarity between HG and some symptoms of hyperthyroidism, as well as the fact that hCG can lead to thyroid stimulation ('Gestational transient thyrotoxicosis').[12, 13] Furthermore, Helicobacter Pylori has been suggested to play a role in the aetiology of HG.[14] Additionally, there are studies suggesting a genetic cause: an association was found between two appetite and cachexia genes, GDF-15 and GFBP7, and HG.[15, 16] These findings add to the evidence of a role for genetics in studies that showed a high hereditary rate for HG, with sisters of HG sufferers having an

18-fold increased risk for developing HG.[17] However, since none of these pathways offer an adequate explanation for HG, it is likely that HG has a multifactorial aetiology.[11]

Treatment

There is a lack of a curative treatment for HG, and research for finding a cure is complicated by the multifactorial origin of the disease. Current treatment for HG consists of a symptomatic approach, including anti-emetics, vitamin supplementation, intravenous rehydration and, in more severe or long lasting cases, tube feeding.[1] Little evidence is available on the effects of these currently available treatments and many of them are not evidence based.

A Cochrane systematic review from 2016 summarized evidence on available treatments for HG.[18] One of their important findings was that, besides the limited studies available on non-pharmacological interventions for HG, outpatient and inpatient care had similar maternal and neonatal outcomes. It also concluded that none of the available pharmacological treatments could be considered superior to any other. High quality evidence was lacking, with only a few placebo controlled trials available, leading to insufficient evidence available to draw any firm conclusions. However, studies evaluating ondansetron showed that ondansetron has less side effects than for example metoclopramide, while their effectiveness seems comparable, an important matter that should be taken into account in clinical practice.[19-21] Additionally, the use of anti-emetics among HG patients is challenged by concerns of the risk of congenital abnormalities. For example, a recently published study from the United States of America showed a small increased risk of the foetus developing oral clefts in mothers using ondansetron during the first trimester of pregnancy compared to offspring of mothers without ondansetron use (0.14% versus 0.11%). Subsequently, the Dutch Association Of Obstetrics and Gynaecology (NVOG) sent out a warning that ondansetron should not be advised as treatment for HG in the first trimester, despite the fact that this increased risk is very small and that little other (anti-emetic) treatments are available.[22]

Currently, existing pharmacological treatments are being evaluated for their effectiveness and safety in HG. Two small randomized controlled trials showed that gabapentin and clonidine led to improvement of symptoms,[23, 24] and at the moment an RCT is being carried out, evaluating whether mirtazapine would be an effective treatment for HG.[25] Nonetheless, these treatment options again are aimed at symptom control, and do not amount to a cure for HG. The Priority Setting project assisted by the James Lind Alliance, published in 2021, states that 'Finding a Cure' is the number one priority in future HG research.[26]

HG's impact on pregnant people

HG has an enormous impact on the wellbeing and quality of life of HG sufferers as HG can lead to both physical and psychological health problems.[30, 31] Maternal physical complications are mostly due to vomiting, subsequently leading to undernutrition and dehydration, and can consist of hypokalaemia, hyponatraemia or anaemia.[32] Additionally, HG has been associated with an increased risk of patients developing a venous thromboembolism, probably caused by dehydration and immobilization, a severe condition that is the leading cause of deaths among pregnant people in the United Kingdom.[33, 34] In rare cases, Wernicke Encephalopathy, a neurological condition due to vitamin B1 deficiency, can occur.[35] HG has, next to severe physical discomfort, also an enormous impact on maternal wellbeing.[30] Social life, work and other day to day activities are limited by the disease, which affects the quality of life of HG sufferers.[30, 31] This is an important aspect of the disease, illustrated by the fact that 'HG symptoms strongly limiting daily activities' is included as one of the required criteria in the newly developed definition of HG.

Depression and anxiety symptoms are frequently reported while suffering from HG.[36] The literature on this topic strongly suggests that poor mental health can persist until after pregnancy, illustrated by the fact that approximately 20% of people go on to develop posttraumatic stress disorder (PTSD) after a pregnancy complicated by HG.[37, 38] The severe effect of HG on sufferers' wellbeing is perhaps best demonstrated by the fact that an estimated 5-10% of pregnancies, despite being planned and wanted, are terminated with HG stated as the only reason for termination.[39, 40] Altogether, evidence so far shows worrying high numbers of HG sufferers terminating otherwise wanted pregnancies and maintaining or developing mental health problems in later life. However, only little evidence is available and more research is needed, for example to investigate whether there is an association between HG severity or other predictive factors and higher HG recurrence or termination rates or the risk of people with HG developing mental health issues to focus on preventative treatments in the future.

Health outcomes of offspring born to people with HG

Animal studies as well as human studies of people born in periods of famine globally have previously shown that undernutrition in pregnancy can lead to adverse long-term health effects in offspring, such as diabetes and cardiovascular diseases.[41-45] There is also mounting evidence that ties HG to adverse health outcomes in offspring. A systematic review and meta-analysis from 2012 showed that offspring born to people with HG are more likely to be born preterm, to be small for gestational age and to have lower birth weights.[46] There are also indications that HG can lead to long term health effects in offspring. HG exposure in utero might lead to

an increased cardiovascular and diabetes disease risk in later life,[47-49] and offspring of people with HG are also more likely to develop neurodevelopment and mental health disorders.[50, 51] Additionally, there are studies suggesting that HG exposure in utero leads to an increased risk for male offspring developing testis cancer.[52, 53] Many studies published on long-term effects however consisted of small study populations. In order to draw definitive conclusions on the effects of HG on offspring long-term health, there is need for a more up-to-date review that gives an overview of currently available (and lack of) evidence and whether there is evidence available that investigates a possible dose-effect response.

Aims and outline of this thesis

Research on the topic of HG has seen an upward trend, both in quality, quantity, and in its alignment with patients' priorities. However, there are still many gaps in research:

- First, it is unclear how much evidence is available for the top 10 priorities in HG research that have recently been listed by patients and clinicians.

- Secondly, prognostic tools or markers for HG severity and prognosis, which can help patients and health care professionals to individualize treatment, have yet to be identified.

- Third, the recurrence rate in current literature varies from 15 to 81%, which hampers informative counselling during preconception consultation for people who suffered from HG in a previous pregnancy. More research is needed to narrow this range and to identify risk factors for an increased HG recurrence risk in subsequent pregnancies.

- Fourth, although there is overwhelming evidence that HG is associated with depression and anxiety during pregnancy, little is known about maternal mental health in the postpartum period and in later life.

- Finally, the effects of HG exposure on the fetus (both perinatal and long-term effects) remain unclear, with the latest systematic review summarizing evidence on this topic being outdated, as it was performed almost ten years ago.

Therefore, this thesis will address the following research questions:

Part I includes *chapter 2* that describes available evidence for the Top 10 most urgent research questions in HG developed by patients, carers and health care professionals, in order to easily identify current knowledge gaps in HG to prevent research waste in the future.

Part II studies the effects of HG on maternal health during and after HG. In *chapter 3* we assessed whether thyroid function can be used as a marker and predictor for HG disease severity. *Chapter 4* addresses the recurrence rate of HG in a prospective cohort study. The proportion of people to postpone pregnancies due to their history of HG, and the number of terminations of pregnancy due to recurrent HG is described. In *chapter 5,* in the same cohort, we report on the proportion of people with HG with depression, anxiety and PTSD symptoms on average 4.5 years after their index pregnancy with HG. *Chapter 6* provides an overview of literature reporting on vitamin K deficiency during HG and the corresponding maternal and neonatal complications, including maternal haemorrhage and congenital anomalies.

Part III describes the effects of HG on perinatal outcomes and offspring health. *Chapter 7* is a systematic review and meta-analysis that gives an overview of fetal and neonatal outcomes among offspring born to people with HG, while *chapter 8* is a systematic review of available evidence on offspring health outcomes beyond the perinatal period. *Chapter 9* is an observational cohort study that answers the question whether HG can induce changes in cardiometabolic markers in offspring cord blood and the possible beneficial effect of maternal early enteral tube feeding on these cardiometabolic markers.

REFERENCES

1. Niebyl JR. Clinical practice. Nausea and vomiting in pregnancy. The New England journal of medicine. 2010;363(16):1544-50.
2. Einarson TR, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: a meta analysis. J Popul Ther Clin Pharmacol. 2013;20(2):e171-83.
3. Centraal Bureau voor de Statistiek. Geboorte; kerncijfers. 2020.
4. United Nations. Demographic Yearbook 70th Issue. New York. 2019.
5. Adams MM, Harlass FE, Sarno AP, Read JA, Rawlings JS. Antenatal hospitalization among enlisted servicewomen, 1987-1990. Obstet Gynecol. 1994;84(1):35-9.
6. Gazmararian JA, Petersen R, Jamieson DJ, Schild L, Adams MM, Deshpande AD, et al. Hospitalizations during pregnancy among managed care enrollees. Obstet Gynecol. 2002;100(1):94-100.
7. Fejzo MS, Poursharif B, Korst LM, Munch S, MacGibbon KW, Romero R, et al. Symptoms and pregnancy outcomes associated with extreme weight loss among women with hyperemesis gravidarum. J Womens Health (Larchmt). 2009;18(12):1981-7.
8. Bolin M, Åkerud H, Cnattingius S, Stephansson O, Wikström A. Hyperemesis gravidarum and risks of placental dysfunction disorders: a population-based cohort study. BJOG: An International Journal of Obstetrics & Gynaecology. 2013;120(5):541-7.
9. Koot MH, Boelig RC, Van't Hooft J, Limpens J, Roseboom TJ, Painter RC, et al. Variation in hyperemesis gravidarum definition and outcome reporting in randomised clinical trials: a systematic review. BJOG : an international journal of obstetrics and gynaecology. 2018;125(12):1514-21.
10. Jansen LAW, Koot MH, Van't Hooft J, Dean CR, Bossuyt PMM, Ganzevoort W, et al. The windsor definition for hyperemesis gravidarum: A multistakeholder international consensus definition. Eur J Obstet Gynecol Reprod Biol. 2021;266:15-22.
11. Niemeijer MN, Grooten IJ, Vos N, Bais JM, van der Post JA, Mol BW, et al. Diagnostic markers for hyperemesis gravidarum: a systematic review and metaanalysis. American journal of obstetrics and gynecology. 2014;211(2):150.e1-15.
12. Tsuruta E, Tada H, Tamaki H, Kashiwai T, Asahi K, Takeoka K, et al. Pathogenic role of asialo human chorionic gonadotropin in gestational thyrotoxicosis. The Journal of clinical endocrinology and metabolism. 1995;80(2):350-5.
13. Goodwin TM, Montoro M, Mestman JH, Pekary AE, Hershman JM. The role of chorionic gonadotropin in transient hyperthyroidism of hyperemesis gravidarum. The Journal of clinical endocrinology and metabolism. 1992;75(5):1333-7.
14. Ng QX, Venkatanarayanan N, De Deyn M, Ho CYX, Mo Y, Yeo WS. A meta-analysis of the association between Helicobacter pylori (H. pylori) infection and hyperemesis gravidarum. Helicobacter. 2018;23(1).
15. Fejzo MS, Myhre R, Colodro-Conde L, MacGibbon KW, Sinsheimer JS, Reddy M, et al. Genetic analysis of hyperemesis gravidarum reveals association with intracellular calcium release channel (RYR2). Mol Cell Endocrinol. 2017;439:308-16.
16. Fejzo MS, Sazonova OV, Sathirapongsasuti JF, Hallgrimsdottir IB, Vacic V, MacGibbon KW, et al. Placenta and appetite genes GDF15 and IGFBP7 are associated with hyperemesis gravidarum. Nature communications. 2018;9(1):1178.
17. Zhang Y, Cantor RM, MacGibbon K, Romero R, Goodwin TM, Mullin PM, et al. Familial aggregation of hyperemesis gravidarum. American journal of obstetrics and gynecology. 2011;204(3):230.e1-7.

18. Boelig RC, Barton SJ, Saccone G, Kelly AJ, Edwards SJ, Berghella V. Interventions for treating hyperemesis gravidarum. The Cochrane database of systematic reviews. 2016(5):Cd010607.

19. Abas MN, Tan PC, Azmi N, Omar SZ. Ondansetron compared with metoclopramide for hyperemesis gravidarum: a randomized controlled trial. Obstet Gynecol. 2014;123(6):1272-9.

20. Kashifard M, Basirat Z, Kashifard M, Golsorkhtabar-Amiri M, Moghaddamnia A. Ondansetrone or metoclopromide? Which is more effective in severe nausea and vomiting of pregnancy? A randomized trial double-blind study. Clin Exp Obstet Gynecol. 2013;40(1):127-30.

21. Sullivan CA, Johnson CA, Roach H, Martin RW, Stewart DK, Morrison JC. A pilot study of intravenous ondansetron for hyperemesis gravidarum. American journal of obstetrics and gynecology. 1996;174(5):1565-8.

22. Huybrechts KF, Hernández-Díaz S, Straub L, Gray KJ, Zhu Y, Patorno E, et al. Association of Maternal First-Trimester Ondansetron Use With Cardiac Malformations and Oral Clefts in Offspring. JAMA. 2018;320(23):2429-37.

23. Maina A, Arrotta M, Cicogna L, Donvito V, Mischinelli M, Todros T, et al. Transdermal clonidine in the treatment of severe hyperemesis. A pilot randomised control trial: CLONEMESI. BJOG : an international journal of obstetrics and gynaecology. 2014;121(12):1556-62.

24. Guttuso T, Messing S, Mullin P, Strittmatter C, Saha S, Thornburg LL. Gabapentin's Effects on Hyperemesis Gravidarum: A Randomized Controlled Trial [17N]. Obstetrics & Gynecology. 2020;135.

25. Ostenfeld A, Petersen TS, Futtrup TB, Andersen JT, Jensen AK, Westergaard HB, et al. Validating the effect of Ondansetron and Mirtazapine In Treating hyperemesis gravidarum (VOMIT): protocol for a randomised placebo-controlled trial. BMJ Open. 2020;10(3):e034712.

26. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient-clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ open. 2021;11(1):e041254-e.

27. Dean C. Does the historical stigma of hyperemesis gravidarum impact health care professionals' attitudes towards and treatment of women with the condition today? A review of recent literature. MIDIRS Midwifery Digest. 2016;26(2):186-93.

28. van Vliet R, Bink M, Polman J, Suntharan A, Grooten I, Zwolsman SE, et al. Patient Preferences and Experiences in Hyperemesis Gravidarum Treatment: A Qualitative Study. J Pregnancy. 2018;2018:5378502.

29. Havnen GC, Truong MB-T, Do M-LH, Heitmann K, Holst L, Nordeng H. Women's perspectives on the management and consequences of hyperemesis gravidarum – a descriptive interview study. Scandinavian Journal of Primary Health Care. 2019;37(1):30-40.

30. Lacasse A, Rey E, Ferreira E, Morin C, Bérard A. Nausea and vomiting of pregnancy: what about quality of life? BJOG : an international journal of obstetrics and gynaecology. 2008;115(12):1484-93.

31. Poursharif B, Korst LM, Fejzo MS, MacGibbon KW, Romero R, Goodwin TM. The psychosocial burden of hyperemesis gravidarum. J Perinatol. 2008;28(3):176-81.

32. Fejzo MS, Trovik J, Grooten IJ, Sridharan K, Roseboom TJ, Vikanes Å, et al. Nausea and vomiting of pregnancy and hyperemesis gravidarum. Nature Reviews Disease Primers. 2019;5(1):62.

33. Virkus RA, Løkkegaard E, Lidegaard Ø, Langhoff-Roos J, Nielsen AK, Rothman KJ, et al. Risk factors for venous thromboembolism in 1.3 million pregnancies: a nationwide prospective cohort. PLoS One. 2014;9(5):e96495.

34. de Swiet M. Maternal mortality: Confidential Enquiries into Maternal Deaths in the United Kingdom. American journal of obstetrics and gynecology. 2000;182(4):760-6.

35. Oudman E, Wijnia JW, Oey M, van Dam M, Painter RC, Postma A. Wernicke's encephalopathy in hyperemesis gravidarum: A systematic review. Eur J Obstet Gynecol Reprod Biol. 2019;236:84-93.

36. Mitchell-Jones N, Gallos I, Farren J, Tobias A, Bottomley C, Bourne T. Psychological morbidity associated with hyperemesis gravidarum: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2017;124(1):20-30.

37. Christodoulou-Smith J, Gold JI, Romero R, Goodwin TM, Macgibbon KW, Mullin PM, et al. Posttraumatic stress symptoms following pregnancy complicated by hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2011;24(11):1307-11.

38. Kjeldgaard HK, Vikanes Å, Benth J, Junge C, Garthus-Niegel S, Eberhard-Gran M. The association between the degree of nausea in pregnancy and subsequent posttraumatic stress. Arch Womens Ment Health. 2019;22(4):493-501.

39. Poursharif B, Korst LM, Macgibbon KW, Fejzo MS, Romero R, Goodwin TM. Elective pregnancy termination in a large cohort of women with hyperemesis gravidarum. Contraception. 2007;76(6):451-5.

40. Nana M, Tydeman F, Bevan G, Boulding H, Kavanagh K, Dean C, et al. Hyperemesis gravidarum is associated with increased rates of termination of pregnancy and suicidal ideation: results from a survey completed by >5000 participants. American journal of obstetrics and gynecology. 2021;224(6):629-31.

41. Edwards LJ, McMillen IC. Periconceptional nutrition programs development of the cardiovascular system in the fetal sheep. Am J Physiol Regul Integr Comp Physiol. 2002;283(3):R669-79.

42. Liu H, Chen X, Shi T, Qu G, Zhao T, Xuan K, et al. Association of famine exposure with the risk of type 2 diabetes: A meta-analysis. Clin Nutr. 2020;39(6):1717-23.

43. Hult M, Tornhammar P, Ueda P, Chima C, Bonamy AK, Ozumba B, et al. Hypertension, diabetes and overweight: looming legacies of the Biafran famine. PLoS One. 2010;5(10):e13582.

44. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Schroeder-Tanka JM, et al. Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. Heart (British Cardiac Society). 2000;84(6):595-8.

45. de Rooij SR, Painter RC, Roseboom TJ, Phillips DI, Osmond C, Barker DJ, et al. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. Diabetologia. 2006;49(4):637-43.

46. Veenendaal MV, van Abeelen AF, Painter RC, van der Post JA, Roseboom TJ. Consequences of hyperemesis gravidarum for offspring: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2011;118(11):1302-13.

47. Ayyavoo A, Derraik JG, Hofman PL, Biggs J, Bloomfield FH, Cormack BE, et al. Severe hyperemesis gravidarum is associated with reduced insulin sensitivity in the offspring in childhood. The Journal of clinical endocrinology and metabolism. 2013;98(8):3263-8.

48. Grooten IJ, Painter RC, Pontesilli M, van der Post JA, Mol BW, van Eijsden M, et al. Weight loss in pregnancy and cardiometabolic profile in childhood: findings from a longitudinal birth cohort. BJOG : an international journal of obstetrics and gynaecology. 2015;122(12):1664-73.

49. Poeran-Bahadoer S, Jaddoe VWV, Gishti O, Grooten IJ, Franco OH, Hofman A, et al. Maternal vomiting during early pregnancy and cardiovascular risk factors at school age: the Generation R Study. J Dev Orig Health Dis. 2020;11(2):118-26.

50. Fejzo M, Kam A, Laguna A, MacGibbon K, Mullin P. Analysis of neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum reveals increased reporting of autism spectrum disorder. Reprod Toxicol. 2019;84:59-64.

51. Getahun D, Fassett MJ, Jacobsen SJ, Xiang AH, Takhar HS, Wing DA, et al. Autism Spectrum Disorders in Children Exposed in Utero to Hyperemesis Gravidarum. Am J Perinatol. 2021;38(3):265-72.

52. Depue RH, Pike MC, Henderson BE. Estrogen exposure during gestation and risk of testicular cancer. J Natl Cancer Inst. 1983;71(6):1151-5.

53. Henderson BE, Benton B, Jing J, Yu MC, Pike MC. Risk factors for cancer of the testis in young men. Int J Cancer. 1979;23(5):598-602.

PART I
CURRENT AVAILABLE HG EVIDENCE

CHAPTER 2

A systematic evidence map of evidence addressing the top ten priority research questions for hyperemesis gravidarum

Caitlin R. Dean, Kelly Nijsten, René Spijker, Margaret E. O'Hara,
Tessa J. Roseboom, Rebecca C. Painter

Accepted for publication in The British Medical Journal Open.

ABSTRACT

Objective

Knowledge gaps regarding hyperemesis gravidarum (HG) are substantial. We aimed to systematically identify and map recent evidence addressing the top ten priority questions for HG, as published in 2021 in a James Lind Alliance Priority Setting Partnership.

Design

Systematic evidence map

Methods

We searched MEDLINE and EMBASE on 12th January 2021 and CINAHL on 22nd February 2021 with search terms hyperemesis gravidarum, pernicious vomiting in pregnancy and their synonyms. Results were limited to 2009 onwards. Two reviewers independently screened titles and abstracts to assess whether the studies addressed a top ten priority questions for HG. Differences were discussed until consensus was reached. Publications were allocated to one or more top ten research questions. Study design was noted, as was patient or public involvement. Two reviewers extracted data synchronously and both cross-checked 10%. Extracted data was imported into EPPI Reviewer software to create an evidence map.

Outcome measures

The number and design of studies in the search yield, displayed per the published ten priority questions.

Results

Searches returned 4338 results for screening; 406 publications were included in the evidence map. 136 publications addressed multiple questions. Numerous studies address the immediate and long-term outcomes or possible markers for HG (question 8 and 9, respectively 164 & 82 studies). Very few studies seek a possible cure for HG (question 1, 8 studies), preventative treatment (question 4, 2 studies) or how to achieve nutritional requirements of pregnancy (question 10, 17 studies). Case reports/series were most numerous with 125 (30.7%) included. Few qualitative studies (9, 2.2%) were identified. 25 (6.1%) systematic reviews addressed eight questions, or aspects of them. 31 (7.6%) studies included patient involvement.

Conclusions

There are significant gaps and overlap in the current HG literature addressing priority questions. Researchers and funders should direct their efforts at addressing the gaps in the top ten questions.

Strengths and limitations of this study

- The ten questions mapped were identified and prioritised by patients and clinicians using the James Lind Alliance method, thereby ensuring their relevance and importance.
- A broad overview of the research landscape of the top ten priority questions for hyperemesis gravidarum is provided, and gaps can be easily identified in this visual presentation.
- We translated 21 articles but were unable to translate 18 foreign language studies, particularly those in Arabic and Iranian, which may be seen as a limitation.
- Many of the excluded studies were abstracts which had not been published as an article, highlighting the need for researchers to ensure their research is published.
- Search results were limited to 2009 onwards. Older literature may be informative but was not in the scope of this review.

BACKGROUND

Hyperemesis gravidarum (HG) affects approximately 1.1% of pregnant women globally.[1] The condition is characterised by extreme levels of nausea and vomiting leading to complications such as dehydration and malnutrition.[2] HG accounts for severe physical and psychological morbidity for women affected,[3-5] and where symptoms lead to malnutrition in the mother, there may be lifelong consequences for the exposed offspring.[6, 7]

Prior to the rapid advance of Intravenous (IV) therapies in the early 1960's, HG was well documented and researched as it was a common cause of death in early pregnancy.[8, 9] However, with the dawn of the psychosomatic era and with the invention of IV therapy, interest in the condition declined and HG patients were commonly mislabelled as psychiatric; an attitude which has persisted into the current century.[10] The incorrect psychiatric labelling of HG can lead to further stigmatisation and consequently a lack of interest in HG research.[11] Additionally, the little research that has been done into HG has been hampered by factors such as a lack of definition and standard outcomes rendering research too heterogeneous and unfit for meta-analysis.[12] Two recent systematic reviews of treatments for nausea and vomiting in pregnancy (NVP) and HG were unable to draw conclusions due to the heterogeneity of the studies included.[13, 14] However, researchers are now seeking to lay solid foundations for future research such as with an internationally agreed definition and by mandating a set of core study outcomes required for publication, which could each contribute to limit heterogeneity of individual studies.[15]

The chasm between the questions patients and clinicians want answers to and the questions research has been seeking to answer, is an important underlying factor for research waste.[16] The recent introduction of patient and public involvement (PPI) in the research process from question development to outcome dissemination is aimed at closing this chasm.[17] Ensuring that research funding is directed to the most important and useful projects can significantly reduce research waste.[18] A recent James Lind Alliance (JLA) Priority Setting Partnership for HG, which benefited from thorough PPI throughout, identified the top 10 priority questions for researchers to address over the coming years.[19]

The aim of the present evidence mapping project was to systematically identify the number and design of published literature addressing the top ten priority questions for HG. The systematic evidence map (SEM) we aimed to produce should help researchers and funders identify the areas of greatest need and potential benefit thereby reducing research waste and maximising value. We additionally aimed to map patient and public involvement in HG research.

METHODS AND ANALYSIS

Study design

Systematic Evidence Map methodology is a systematic approach to identifying gaps in knowledge and future research needs of a particular topic using transparent and robust methods.[20, 21] It aims to create a visual matrix of current evidence without further appraisal of the quality of the evidence.[22]

A preliminary search of Prospero and key databases for published studies and protocols was conducted to rule out other SEM projects or similar scoping review for HG; no such studies or protocols were identified.

A methodological framework combining the Arksey and O'Malley[23] scoping studies approach and the Campbell Collaboration[24] protocol template for evidence maps was adopted, which at the time of development was a pragmatic approach to use the most established methods to date for SEMs. The Campbell Collaboration is an organisation promoting the use of systematic evidence synthesis for positive social and economic policy and practice change.[25] It has produced standards for SEMs which we have incorporated into this project such as defining the search strategy, objectives, inclusion criteria, eligibility criteria, categories, restrictions and so on, in advance in order to expand the below stages.[24]

The Arksey and O'Malley steps which were incorporated in our approach are as follows:[23]

Stage 1: identifying the research question(s)
Stage 2: identifying relevant studies, i.e. conducting the searches
Stage 3: study selection i.e. screening and selecting those fitting the eligibility criteria
Stage 4: charting the data
Stage 5: collating, summarising, and reporting the results

Patient and Public Involvement (PPI)

The authors of this SEM are fully committed to patient involvement in HG research which adds significant relevance to research findings.[26] The lead author (CD) and MEO are both HG patients themselves and are experienced advocates of patients with HG. Patients have co-created this work with clinicians and academics; An HG patient and advocate, created the concept, conducted the research and wrote the manuscript, while experienced academics and clinicians acted as supervisors and collaborators. Additional patients and clinicians

were consulted throughout the process (See acknowledgements section): researchers and clinicians opinions were sought during online meetings and phone calls and patients from England were shown early versions of the SEM during informal volunteer online meet-ups on the usefulness of the map and the categories for them. Trustees of the charities Pregnancy Sickness Support (United Kingdom (UK)) and Hyperemesis Ireland, whose boards consist of patient representatives, clinicians and/or researchers, provided feedback after the final SEM was presented during an online meeting.

Ethics approval statement

The SEM project does not require ethical approval. Patient and Public Involvement for this research did not require ethical approval. See **Supplementary File 1** for Health Research Authority confirmation on non-research status.

Stage 1: Identifying the research questions

A James Lind Alliance (JLA) Priority Setting Partnership (PSP) for HG was conducted between 2017-2019, which brought together patients, their caregivers and offspring and healthcare professionals involved in HG care, to identify the top ten most pressing, unanswered research questions.[19] **Table 1** provides the resulting top ten questions which form the basis for this project. For full details of the JLA project, including methods, data collection, participant recruitment, countries represented, the prioritisation process and how the questions were developed, please refer to the published research available at: doi:10.1136/bmjopen-2020-041254.

Table 1. Top ten unanswered research priority questions* for hyperemesis gravidarum, in ranked order of importance from one as the most important[19]

Ranking	Question
1	Can we find a cure? What novel or new treatments are being developed/tested/used elsewhere which could have a curative effect and to address all the symptoms of HG rather than just the vomiting?
2	How can we most effectively manage HG? What clinical support measure is most important to people who have had hyperemesis and what did they find most beneficial? E.g. medical management, pharmaceutical review, nutrition support, rehydration, psychological support
3	What causes HG?
4	Is HG preventable? What is the effect of preventative treatment or early intervention on the severity and duration of HG in a subsequent pregnancy?
5	What are the immediate[a] and long-term effects[b] of HG (including malnutrition[c] and dehydration[d], stress[e]) on the developing fetus (offspring)?

Table 1. Continued

Ranking	Question
6	What are the immediate[a] and long-term effects[b] of the various medications/treatments on the developing fetus (offspring) throughout the various stages of pregnancy and in varying doses or combinations of treatments?
7	What are relative[f] efficacies of the current medications and treatment options available? What is the optimal dose, route, timing and combination of the medications and what are the related side effects?
8	What are the immediate[a] and long term[b], physical, mental and social consequences and complications of HG (including malnutrition and dehydration) on the pregnant person's body? (ie. Metabolic impact, DVT, depression, effects of dehydration)
9	What clinical measurements and markers are most useful in assessing, diagnosing, managing and monitoring HG?
10	What are the nutritional requirements of the 1st, 2nd and 3rd trimesters and how can people with HG achieve these goals? i.e. Oral supplements, fortifying food, dietary measures

**The phrasing of the questions was established using the JLA consensus method therefore we were not able to alter the phrasing in the writing of this manuscript. [a] Immediate effects relates to those during the perinatal period. [b] long-term effects relates to any time after the perinatal period. [c] example indicators of malnutrition include weight loss or nutritional intake. [d] example indicators of dehydration include need for IV rehydration or urine output. [e] Stress could be measured with questionnaires. [f] relative to each other.*

Stage 2: identifying relevant studies

Search Strategy

The original search strategy was devised and conducted electronically by a Medical Information Specialist (author RS) in 2019 as part of the JLA PSP evidence check process. This search sought to identify if any questions could be considered answered with enough evidence of sufficient quality. This has been described in detail previously and includes the protocol.[19] For this SEM, the searches were repeated. MEDLINE and EMBASE were searched from inception to the 12th of January 2021 by RS using the following broad terms:

> hyperemesis gravidarum/ or (("excessive vomiting" or (pernicious adj3 vomiting) or hyperemesis) and (gravid* or pregn* or gestation or antenatal)).mp.

Additionally, the Cochrane Library was searched electronically by CD in collaboration with RS, using the term "hyperemesis gravidarum". A further CINAHL search was conducted with the same strategy by RS on the 22nd February 2021, as a deviation from the protocol, because the reviewers noticed that certain papers from nursing and midwifery journals had not been returned in the original searches.

The searches are detailed in **Supplementary File 2**.

Eligibility criteria

Inclusion and exclusion criteria

To be eligible, publications had to study women with HG, or their offspring, and address any aspect of any of the top ten questions. All study designs and languages were eligible. We did not apply a minimum study size for eligibility. The steering group for the PSP agreed that due to the paucity of HG research and changing attitudes to HG treatment in the last decade, a ten year limit was most appropriate.[19] Since the first search was performed in 2019, we excluded articles published prior to 2009.

Abstracts (from conference oral presentations or posters) were excluded if full texts for the same study were not found. Review articles in which search methods were not described were excluded as narrative reviews. Reviews which describe their study as a systematic review were included as such. Reviews which described their methods including databases and search terms used, but did not fit generally accepted criteria for a systematic review such as following a protocol, screening and data extracting in duplicate or assessing risk of bias,[27] were labelled as literature reviews. Protocols for systematic reviews, randomised controlled trials, cohort or case-controlled studies and qualitative studies, which were published in peer-review journals and did not yet have a corresponding publication of results, were included. The addition of protocols was aimed at offering researchers information about research currently underway in order to reduce research waste and potentially aid collaboration.

Defining Hyperemesis Gravidarum

Other systematic review protocols have highlighted the challenge of defining the condition of HG itself.[28] A historic lack of clinical definition has hampered HG research efforts globally and made meta-analysis difficult due to heterogeneity within the studies.[12, 29] This review took the same approach as other systematic reviews which have included articles that describe HG, regardless of how that is defined. We excluded studies which only included mild to moderate NVP, but included studies where severe NVP was explicitly described.

Stage 3: Study Selection

Two reviewers (CD and KN) independently screened titles and abstracts to establish if they may be eligible for inclusion according to the criteria above, using Rayyan software,[30] and met to discuss differences, such as whether the study relates to a top ten question or not. In case of disagreement, a third reviewer was consulted (RCP). The reviewers labelled relevant references with the top ten priority question they related to. Full texts were retrieved for full screening. Foreign language papers were translated using Google Translate online where possible and

authors were contacted where full texts were unfindable via online library sources. Full text eligibility screening was conducted independently by the two reviewers for 50% of the texts each, followed by checking 10% of each other's. Discrepancies were discussed until consensus was reached. Publications meeting the inclusion criteria detailed above were included.

Stage 4: Charting the data

The two reviewers independently extracted data from the full texts, completing 50% each, again checking 10% of each other's. An Excel data charting form was used to extract key information which included:

- The top ten question it addressed
- Author(s), full reference, year of publication
- Country
- Abstract
- Aims of the study
- Study design
- Reporting of PPI
- Outcome measures
- Results

Defining Patient and Public Involvement (PPI) and Patient Authorship

INVOLVE is a UK government funded programme established to support active PPI in medical, health service and social care research; it defines PPI in research as "research being carried out 'with' or 'by' members of the public rather than 'to', 'about' or 'for' them".[31] They further define the term "public" as patients, potential patients, carers and people who use health and social care services as well as people from organisations that represent people who use services.[31]

PPI was considered to be included if a manuscript explicitly describes how it was included, if one of the author's affiliation was a patient organisation for HG, or if an author specified that they experienced HG within the manuscript.[32]

Stage 5: Collating, Summarising and reporting the results

Research was categorised according to the top ten priority questions and some questions were further labelled with subcategories which were identified and constructed from the studies during the data extraction process, see **Table 2**. These categories were discussed with clinical research colleagues to ensure they were relevant and reflective of the research. Studies with

ambiguity around its categorisation were also discussed with clinical research colleagues and a patient representative. Research was next categorised according to study designs: Reviews, randomised control trials (RCT), cohort studies, case-control studies, qualitative studies, surveys, and case reports/series. These were further categorised as either systematic or literature reviews, either prospective or retrospective for cohorts and case-control studies, and either case reports or case series. Other designs did not require further subcategories.

Table 1. Question subcategories which emerged during data extraction

Question*	**Subcategories**
Q2: How can we most effectively manage HG?	• Outpatient treatment • Intravenous treatment • Tube feeding • Other treatments
Q3: What causes HG?	• Genetic studies • Helicobacter pylori • Laboratory studies of other factors e.g. hCG • Psychosocial factors • Other causes
Q5: What are the immediate and long-term effects of HG on the fetus?	• Perinatal outcomes • Long-term outcomes
Q7: What are the relative efficacies of current treatments?	• Anti-emetics • Steroids • Other treatments
Q8: What are the immediate and long-term effects of HG on pregnant people?	• Psychosocial effects • Wernicke's encephalopathy • Other maternal complications due to HG • Long term maternal health • Metabolic impact (laboratory results) • Other outcomes
Q9: What clinical measurements and markers in HG are available and most useful in assessing, diagnosing, managing and monitoring HG?	• Psychosocial measurements • Helicobacter pylori as marker • Other laboratory markers • HG assessment questionnaires • Other assessments

**Questions 1, 6 and 10 did not require subcategories*

Data was then imported into the EPPI-Reviewer software[33] which is an online tool designed to generate a bubble map. Bubble maps present evidence visually with circles whose size represents the number of studies. The top ten questions and their subcategories are on the X-axis and methodologies used in the studies are on the Y-axis. Inclusion of PPI in research categories were assigned colours for a third-dimension representation within the map.

Additionally, the country of the studies was labelled for convenient identification of research output by country through the filter function of the software.

Where there was potential ambiguity over categorisation articles were discussed with a third author (RCP) and external justification for labelling sought, e.g. A treatment was considered new or novel if it does not currently appear in national guidelines in the UK, USA or Netherlands (i.e. gabapentin, clonidine, cannabis and mirtazapine).

RESULTS

Identification of studies

Figure 1 shows the Prisma flow chart of the selection process. The combined searches yielded 5821 eligible citations, of which 2435 were excluded because of being published before 2009. After initial screening of title and abstract, 624 remained for full text assessment.

Reasons for exclusions

Of the 624 a further 218 studies were excluded (See**Supplementary File 3**). While 21 included articles were translated, we were unable to translate a further 18 articles which were predominantly written in Persian or Arabic. 126 studies were presented only as an abstract or poster and were therefore excluded, and we were unable to obtain full texts for a further 17 articles despite requests to the authors. 16 of the 624 studies were deemed not to be about HG when the full text was reviewed. A further 17 articles were commentaries or letters referring to other research and 24 were excluded for other reasons such as being a general discussion, background article or narrative review.

A total of 406 studies were included in the final SEM. See **Supplementary File 4** for the full list with labels. Interactive spreadsheets are available at: https://www.hgresearch.org/hgmapfiles

Bubble Map

The Interactive map is available online at www.hgresearch.org/hgmapfiles, and as **Figure 2** as a static image without expanding subcategories.

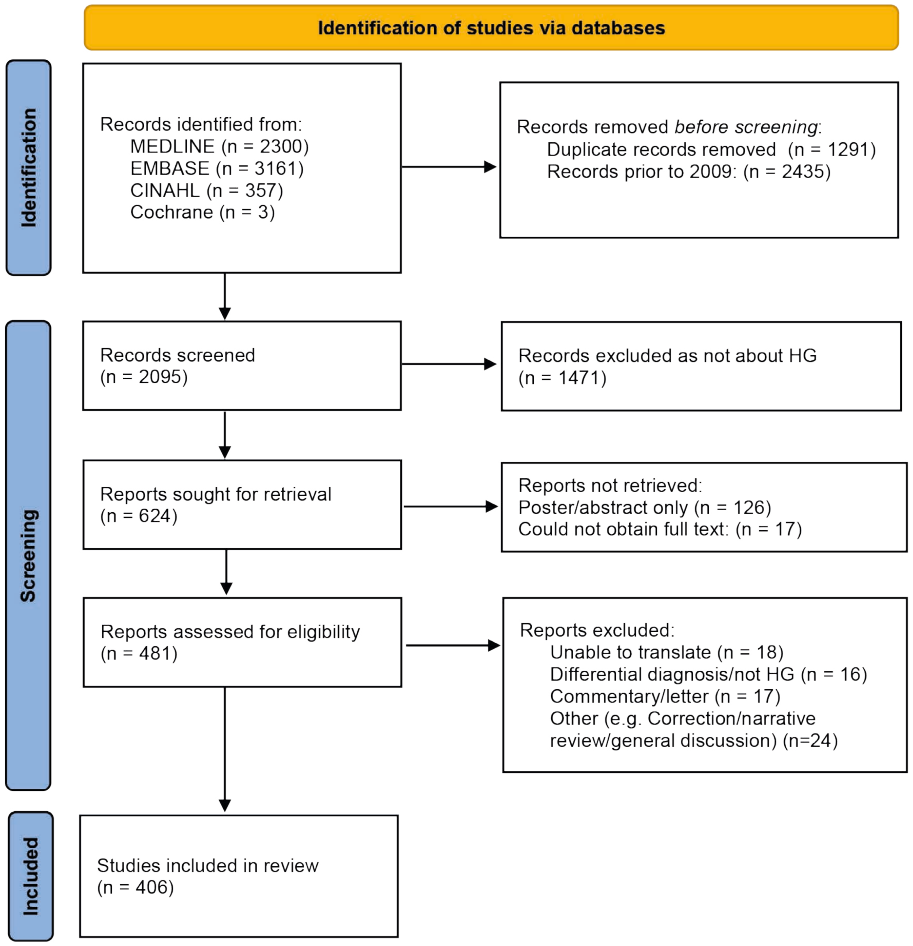


Figure 1. PRISMA Flow chart of inclusions and exclusions

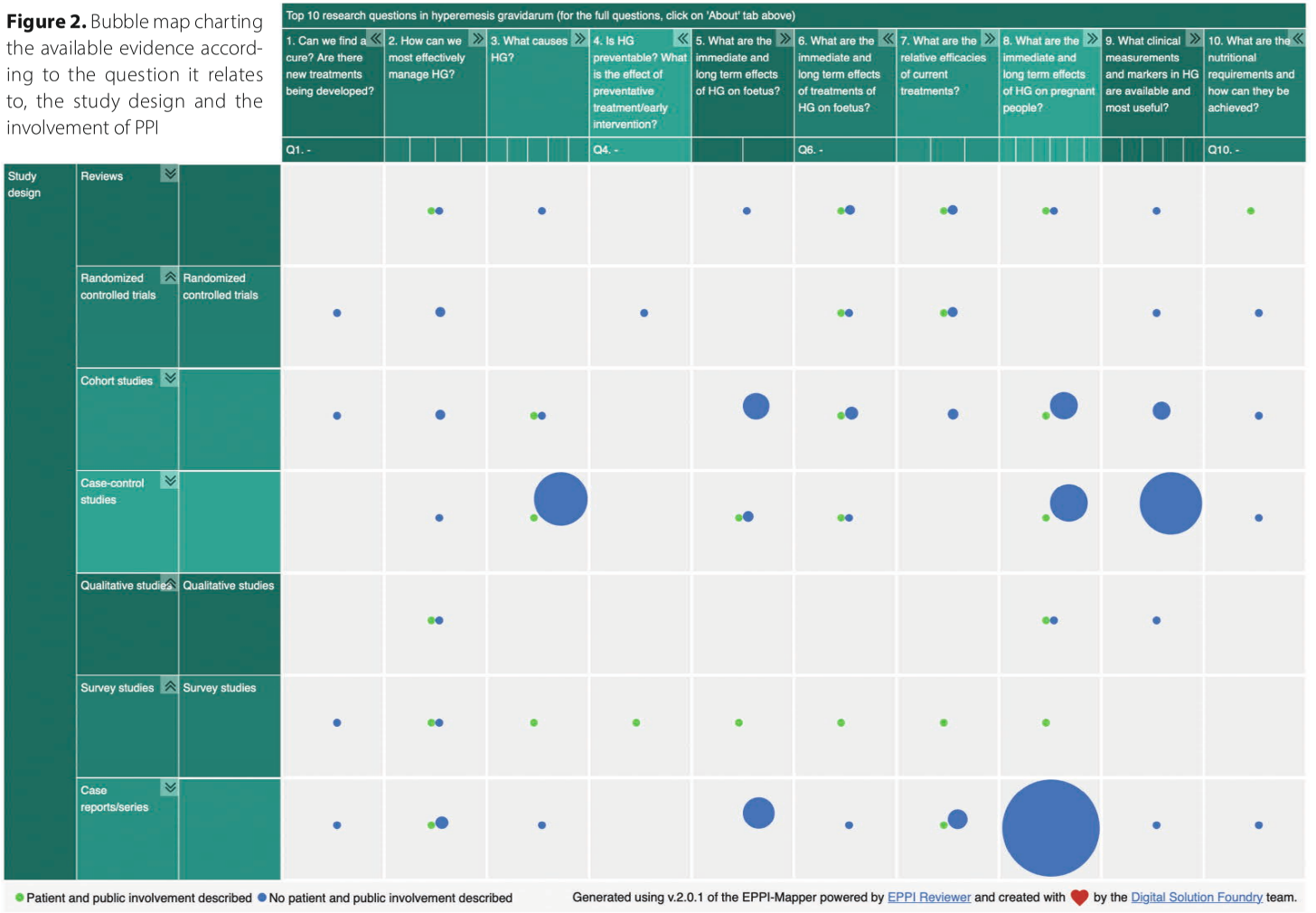


Figure 2. Bubble map charting the available evidence according to the question it relates to, the study design and the involvement of PPI

Characteristics of studies

Table 3 shows the number of studies according to study design. There was some overlap as five studies incorporated more than one design, for example a prospective case-control study which also included a qualitative element.[34] Case reports/series and case-control studies were most numerous with 125 and 124 identified respectively. Of the 406 included papers, only 25 were systematic reviews and 21 were RCTs. The majority of the studies included originated from Europe and the USA (66%, 275/406).

Table 3. Studies included in systematic evidence map for hyperemesis gravidarum, presented according to their method

Method**	Number of studies	Method Subcategory	Number of studies
Reviews	34	Systematic Review	25
		Literature Review	9
RCT*	21		
Cohort Studies	85	Prospective Cohort Study	41
		Retrospective Cohort Study	44
Case-Control Studies	124	Prospective Case-Control Study	109
		Retrospective Case-Control Study	15
Qualitative Study*	9		
Surveys*	13		
Case Reports/Series	125	Case reports	115
		Case Series	10
Total	**411**		

**Randomised Control Trials, Qualitative studies and Surveys did not require subcategories.*
***Method categories were not mutually exclusive.*

Results per question

Figure 3 shows the number of studies identified per question which ranged from two studies addressing preventing HG (question 4) to 164 studies addressing the effects and complications of HG (question 8). 136 studies addressed more than one question. Where more than nine references are described please refer to **Supplementary File 4**.

Question 1 *– Can we find a cure? What novel or new treatments are being developed/tested/used elsewhere which could have a curative effect and to address all the symptoms of HG rather than just the vomiting?*

Eight studies assessed whether four different novel treatments could have a beneficial or curative effect on HG. Of these, two were RCTs (one assessing transdermal clonidine[35] and one assessing gabapentin[36]), two were prospective cohort studies also assessing clonidine[37] and gabapentin,[38] one was a survey study of cannabis use in pregnancy for sickness,[39] and three were case reports (one of cannabis[40] and two of mirtazapine).[41, 42] The RCTs and cohort studies also reported fetal outcomes while the other studies did not.

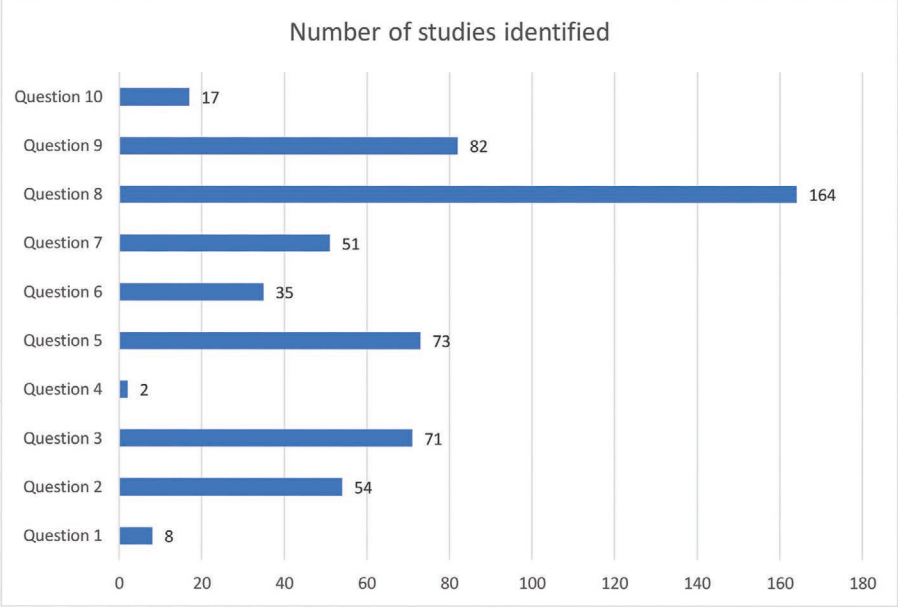


Figure 3. The number of studies identified per top ten question included in the evidence map for hyperemesis gravidarum.

Question 2 *– How can we most effectively manage HG? What clinical support measure is most important to people who have had hyperemesis and what did they find most beneficial? E.g. medical management, pharmaceutical review, nutrition support, rehydration, psychological support*

A total of 54 studies were identified regarding management of HG. Five systematic reviews were identified, four of which were almost identical assessing interventions for HG.[13, 14, 43, 44] The last was a systematic review of the effect of acustimulation on NVP and HG.[45] Nine RCTs were identified on a variety of topics regarding how best to manage HG, including three studies assessing outpatient care,[46-48] one on tube feeding,[49] two assessed intravenous therapies,[50, 51] and three studies reported other types of clinical support measures, including a 12-hour fasting approach and relaxation methods.[52-54]

Question two contained the most qualitative studies with six identified describing women's experiences of the condition and its treatments.[55-60]

Question 3 *– What causes HG?*

71 studies have attempted to find a cause for HG. Of these, eight have sought to identify genetic causes,[61-68] 17 researched the role of Helicobacter pylori (H. pylori) in the aetiology of HG, 23 looked at a variety of laboratory markers, 18 studies assess psychosocial factors as a cause and five studies assessed other possible causes such as nervous system dysfunction, dietary factors and the vestibular system. Laboratory studies were included under question three if the authors of the study stated that they were specifically looking at possible aetiology, rather than for diagnostics, assessment or monitoring purposes.

Of the 17 studies assessing H. pylori, three systematic reviews have been published[69-71] and the remaining 14 studies are all prospective case control studies.

To date, no systematic review has been conducted of the published reports regarding genetic factors, laboratory markers and possible psychosocial causes of HG, although one systematic review assessed polyunsaturated fatty acids in HG.[72]

Question 4 *– Is HG preventable? What is the effect of preventative treatment or early intervention on the severity and duration of HG in a subsequent pregnancy?*

Only two studies sought to assess if HG is preventable, either during a pregnancy or prior to a subsequent one. Of these, one was an RCT assessing the effect of pre-emptive medication on the incidence and severity of HG in a subsequent pregnancy.[73] The other was a survey study exploring the experiences of HG in a subsequent pregnancy and how factors such as increased support, or early treatment affected symptoms.[74]

Question 5 *– What are the immediate and long-term effects of HG (including malnutrition and dehydration, stress) on the developing fetus (offspring)?*

We identified 73 studies assessing perinatal and/or long-term offspring outcomes following HG. Of these, 60 assessed perinatal outcomes and 15 assessed health in later life among offspring; two studies assessed both.[6, 75] Three systematic reviews have been conducted which describe

perinatal outcomes for the fetus, one of which also reported long-term outcomes.[6, 76, 77] Among the 60 studies describing perinatal outcomes, 28 were case reports.

Question 6 – *What are the immediate and long-term effects of the various medications/treatments on the developing fetus (offspring) throughout the various stages of pregnancy and in varying doses or combinations of treatments?*

35 studies reported fetal outcomes following HG treatment with a range of medications and interventions, of which nine were systematic reviews[13, 14, 43, 78-83] and one was a literature review.[84] Four of these assessed the safety of ondansetron specifically.[78, 80, 81, 84] Six were RCTs,[35, 36, 46, 49, 85, 86] 10 were retrospective cohort studies and two were prospective cohort studies.[37, 38] There were also three retrospective case-control studies,[87-89] two surveys[90, 91] and two case reports/series.[92, 93]

Question 7 – *What are relative efficacies of the current medications and treatment options available? What is the optimal dose, route, timing and combination of the medications and what are the related side effects?*

51 studies assessed the efficacy of treatments, of which eight were systematic reviews[13, 14, 43, 44, 79, 82, 83, 94] and ten were RCTs. Of the studies assessing treatments, 30 assessed the efficacy of anti-emetics, five assessed corticosteroids specifically[79, 95-98] and 20 studied "other treatments". Other treatments included gabapentin, clonidine, cannabis, ginger, antacids, diazepam, mirtazapine, B vitamins, Chinese medicines and Japanese herbal Kampo medicines, as well as routes of administration including peripheral central catheters and transdermal application.

Question 8 – *What are the immediate and long-term, physical, mental and social consequences and complications of HG (including malnutrition and dehydration) on the pregnant person's body? (ie. metabolic impact, DVT, depression, effects of dehydration)*

164 studies addressed aspects of this question, however, 90 of these are case reports of serious complications such as thyrotoxicosis, refeeding syndrome, cardiac arrest and hepatorenal failure. In total, there were 56 case reports and one systematic review on Wernicke's encephalopathy.[77] 40 studies assessed the psychosocial effects of HG on women, including two systematic reviews: one of quantitative studies[29] and one of qualitative studies.[3]

Question 9 *- What clinical measurements and markers are most useful in assessing, diagnosing, managing and monitoring hyperemesis?*

82 studies sought to assess clinical measurements and markers for HG, of which 59 were searching for altered serum levels of a vast array of markers, predominately with prospective case-control studies (n= 42). In addition to laboratory markers, six studies sought to validate assessment questionnaire tools[34, 99-103] and six studies looked at the effect that HG had on other assessments conducted during pregnancy, such as screening for gestational diabetes, urinary tract infections and the triple test screen.[104-109] There was one systematic review which summarized diagnostic laboratory markers for HG in general[110] as well as two systematic reviews on H. pylori and HG[69, 70] and one specifically on nucleic acids in pregnancy complications.[111]

Question 10 *- What are the nutritional requirements of the 1st, 2nd and 3rd trimesters and how can people with HG achieve these goals? i.e. Oral supplements, fortifying food, dietary measures*

The effect of HG on nutritional intake and methods for addressing deficiencies were addressed by 17 studies, including a scoping review on the nutritional intake of women with HG [112] and an RCT to assess the effect of early enteral tube feeding.[49]

PPI and patient authorship

PPI was included in 31 studies, of which 12 explicitly described how patients were involved in the development of the research and 25 had an author who was also a patient. Of the studies with patient authors, 19 did not describe the scope of the PPI in the development, design, or production of the research, beyond listing the affiliation. The remainder of the included studies did not mention PPI or explicitly stated that it was not included. See **Supplementary File 1** for the full reference list for each category.

Of those that included PPI, four were systematic reviews[3, 14, 94, 112] (of which one was a systematic review of qualitative studies), four were survey studies,[55, 74, 113, 114] two were prospective case-control studies,[115, 116] one was a protocol for an RCT[86] and one was a qualitative study.[60] Of the remaining 20 studies that included a patient author, but did not report PPI, seven were survey studies,[75, 90, 91, 117-120] four were cohorts,[63, 121-123] three were case-control studies,[61, 87, 124] two were case reports,[125, 126] two were qualitative studies,[56, 58] and one was a literature review.[11]

DISCUSSION

We systematically searched the literature for studies on HG and identified 406 studies, addressing the top ten unanswered research priorities for HG and mapped them according to study design and patient involvement. While all the questions have at least two papers addressing them, the JLA PSP found all questions "remain unanswered by sufficiently robust and conclusive systematic review" and were thus included in the prioritisation process.[19] Where many individual (small) reports exist, a systematic review can help provide robust summary answers to questions. Identifying a presence of a wealth of small individual studies, in the absence of a systematic review could trigger future systematic review and meta-analysis development. For many other questions there simply is a dearth of evidence, and primary research is needed. To our knowledge, this is the first time a SEM has followed a James Lind Alliance Priority Setting Partnership and it is the first SEM for HG. Systematic evidence maps are a relatively new type of evidence synthesis product but are increasingly recognised for their ability to identify gaps in the literature and informing future research efforts, thereby addressing need and reduce research waste.[127, 128]

Gaps in the literature

In this SEM, substantial gaps in the literature were identified as well as duplicate systematic reviews. For example, only two studies were identified for question four, regarding prevention of HG, suggesting a serious need for original research to address the effect of early or preventative treatment. In total, there were only 25 (5.9%) systematic reviews included, of which 13 contained meta-analysis addressing various topics including H. pylori, infant outcomes, diagnostic markers, interventions and medications, psychosocial factors and traditional Chinese therapies. By comparison, an evidence map of social, behavioural and community engagement interventions for reproductive, maternal, newborn and child health, conducted by the World Health Organisation in 2017[129] found systematic reviews accounted for 23% of their 612 included studies. Our SEM also identified notable overlap on systematic review topics, specifically for H. pylori in association with HG (two of which were less than three years apart[69, 70]), and on treatments for NVP/HG, including multiple reviews published in the same year or within one year of each other.[14, 43, 44, 79, 82, 94, 130] Five separate recent systematic reviews on treatments or interventions for HG assessed the efficacy of medications and all found that trials to date were small and of low quality and high heterogeneity, and all concluded with the need for large, high quality trials with consistent outcome measures.[14, 43, 44, 130, 131] This suggests that researchers are not assessing what is already known and where the gaps are before embarking on new systematic reviews or original research which is a necessary step in reducing research waste.[132]

Yet since the publication of these systematic reviews, only one large trial of prednisolone verses placebo has been published and one RCT protocol for mirtazapine verses ondansetron and no others are currently registered with clinicaltrials.gov.[86, 96]

Methods underpinning the literature

Some gaps within the bubble map are unlikely to be filled as some methods would not be appropriate to answer the question, such as using qualitative methods to address the aetiology, while other gaps quite clearly need to be filled with future research, such as randomised control trials for treatment efficacy. There is a demand for large, well designed RCTs rather than yet more systematic reviews of the same heterogeneous, low-quality studies. However, two more systematic reviews are now registered on PROSPERO to assess the effectiveness of acupuncture for HG, despite all five of the recent systematic reviews published including acupuncture as a treatment and finding conflicting results from low quality studies.[133, 134] This problem of redundancy of reviews is not limited to HG. A survey of 73 randomly selected meta-analyses from 2010 found that two thirds (67%) had at least one duplicate meta-analysis published concurrently within three years of the original meta-analysis. This survey found that, on four topics, there were more than eight overlapping meta-analyses with the same subject.[135] Whilst some overlap can be justified and, indeed, necessary for updating and independent replication, the degree of overlap we found in the course of this SEM likely reflects substantial wasted efforts and funds.[135] Registering systematic reviews on PROSPERO, which has been established since 2011, may help to reduce unnecessary duplication.[136] However only two of the five systematic reviews on treatments we identified had been registered on PROSPERO, which hampers authors in their ability to gain timely awareness of concurrent duplicate efforts.[14, 44] A survey of authors who published a systematic review and/or meta-analysis between 2010-2016 found almost half (44.9%) did not register their protocols, primarily due to a lack of knowledge of the need and importance of protocol registration.[137] Increasingly journals are requiring registration of protocols, which should begin to raise awareness of the importance of this practice.

Patient and public involvement tracking

Like ours, some other SEMs have included PPI in their design, methods, conduction and publication.[138, 139] However, they did not extract information on whether the studies they mapped had included PPI and to the best of our knowledge, no other evidence map has specifically extracted data on PPI. The body of evidence in support of PPI in research is ever growing and positive impacts have been found throughout the research process.[26, 140, 141] It is particularly notable for reducing research waste by ensuring that questions are meaningful

and methods appropriate to answer them, improving recruitment and aiding dissemination.[18, 26] We hope that this will act as a stimulant for future research to include PPI.

Strengths and limitations

This was the first SEM on HG and it was conducted from a patient-centred perspective with an innovative approach to evidence synthesis combining two established methods. Another strength of our study is the broad terms used to conduct the search in multiple databases, ensuring a wide net for studies to fall into. We were able to translate 22 foreign language articles, however we were unable to translate a further 18 articles. We also limited our study to 2009 onwards which on the one hand ensures the map is current, but conversely means that some key studies published prior to 2009 were excluded. While we contacted many authors for full texts that were otherwise unavailable, 17 authors did not reply. Furthermore, we did not contact the authors of conference oral/poster abstracts (n=126) to request if full texts had been published due to resource limitations.

Due to the wide variety of study methods included it was not possible to extract data on population sizes in studies. Additionally, other SEMs extracted data on additional features, which we did not, such as "open access availability" of published studies which could be useful for researchers using the map.

Although we had clearly defined categories and two researchers conducted the labelling and checked each other to reduce bias, there was a degree of subjectivity when labelling many papers which could fit in multiple categories or did not describe methods explicitly enough to know exactly how to categorise it. Additionally, due to the broad nature of the top ten questions there was substantial overlap and potential for subjectivity.

Individual questions would benefit from wider searching with individually designed strategies and different methodology. For example, research addressing question 10 may exist within the wider field of pregnancy nutrition and epidemiological studies of anti-emetics that address question seven may not have shown up in our HG specific search. We took a pragmatic approach to inclusion of research where participants are described as having HG or clearly defined severe NVP. Differences in HG or severe NVP diagnosis leads to heterogeneity in included studies and hampers aggregation of evidence, as previously demonstrated .[142] Hopefully the publication of the internationally agreed Windsor definition for HG the next decade will enhance research homogeneity and reduce waste.[143]

Many of the individual questions would benefit from having the quality of their available evidence appraised, however, due to the nature of SEMs,[127] we did not attempt to quality assess included studies, which can be seen as a limitation

Conclusions

This SEM provides an overview of the current evidence addressing the top ten priority questions for HG. While all the questions have at least two papers addressing them, all questions remain unanswered and would benefit from either original research or systematic review. The SEM presents a useful, interactive tool for researchers seeking to address one of these questions and could save valuable finite resources to justify, or rule out, planned studies. The SEM highlights significant gaps in the literature, requiring original research, particularly in the fields of cure and prevention of HG as well how to address the nutritional challenges of HG. We aim for this SEM to be updated annually through the International Collaboration on Hyperemesis Gravidarum (ICHG).

Supplementary File 1. The evidence map; **Supplementary File 2**. References by study design and **Supplementary File 3**. References by Question Number and PPI inclusion, are available on: https://www.hgresearch.org/hgmapfiles

Acknowledgements

The authors would like to thank the Trustees and volunteers of Pregnancy Sickness Support and Dr Brian Cleary, from Hyperemesis Ireland, for consultation throughout the process and feedback on the map and categories. In particular, thanks to patient representatives Amy Armstrong and Emma Watford for input throughout the process.

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Contribution to authorship

CD and RP designed the study and protocol. RS designed the search strategy and conducted the searches as well as converting the data for upload to EPPI. CD and KN screened, data extracted and labelled all studies, with RCP providing third person input where required. CD and KN created the SEM and wrote the manuscript. MEO provided patient involvement, reviewed and gave feedback on the map and reviewed the manuscript. TR provided supervision and feedback on the manuscript. All authors reviewed and approved the final manuscript.

REFERENCES*

*The following references are cited in this report, for the full list of references included in the map see **Supplementary file 4**.

1. Einarson TR, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: A meta-analysis. J Popul Ther Clin Pharmacol. 2013;20(2):e171-e83.
2. Dean CR, Shemar M, Ostrowski GAU, et al. Management of severe pregnancy sickness and hyperemesis gravidarum. BMJ. 2018;363:410-2.
3. Dean C, Bannigan K, Marsden J. Reviewing the effect of hyperemesis gravidarum on women's lives and mental health. Br J Midwifery. 2018;26(2):109-19.
4. Royal College of Obstetricians and Gynaecologists. The Management of Nausea and Vomiting of Pregnancy and Hyperemesis Gravidarum. London: Royal College of Obstetricians and Gynaecologists; 2016.
5. MacGibbon K, Fejzo M, Mullin P. Mortality Secondary to Hyperemesis Gravidarum: A Case Report. Women's Health & Gynecology. 2015;1(2):[Online at: http://scientonline.org/open-access/mortality-secondary-to-hyperemesis-gravidarum-a-case-report.pdf].
6. Veenendaal MV, van Abeelen AF, Painter RC, et al. Consequences of hyperemesis gravidarum for offspring: a systematic review and meta-analysis. Bjog. 2011;118(11):1302-13.
7. Grooten I, Painter R, Pontesilli M, et al. Weight loss in pregnancy and cardiometabolic profile in childhood: findings from a longitudinal birth cohort. Bjog. 2015;122(12):1664-73.
8. Fairweather DV. Nausea and vomiting in pregnancy. Am J Obstet Gynecol. 1968;102(1):135-75.
9. Donald I. Practical Obstetric Problems. 4th ed. London: Luke Lloyd; 1974.
10. Karpel L, de Gmeline C. [Psychological approach to hyperemis gravidarum]. L'approche psychologique des vomissements incoercibles gravidiques. J Gynecol Obstet Biol Reprod. 2004;33(7):623-31.
11. Dean C. Does the historical stigma of hyperemesis gravidarum impact healthcare professional's attitudes and treatment towards women with the condition today? A review of recent literature. MIDIRS Midwifery Digest. 2016;26(2):186-94.
12. Grooten I, Roseboom T, Painter R. Barriers and Challenges in Hyperemesis Gravidarum Research. Nutr Metab Insights. 2015;8(Suppl 1):33-9.
13. Boelig RC, Barton SJ, Saccone G, et al. Interventions for treating hyperemesis gravidarum. Cochrane Database Syst Rev. 2016;5:DOI: 10.1002/14651858.CD010607.pub2.
14. O'Donnell A, McParlin C, Robson SC, et al. Treatments for hyperemesis gravidarum and nausea and vomiting in pregnancy: a systematic review and economic assessment. Health Technol Assess. 2016;20(74):1-268.
15. Painter C, Boelig R, kelly A, et al. Hyperemesis Gravidarum - Why we need consensus on definition and core outcomes. First World Colloquium on Hyperemesis Gravidarum; 21st October 2015; Bergen, Norway: No Hype - The Norwegian Hyperemesis Initiative; 2015.
16. The James Lind Alliance. The James Lind Alliance Guidebook Version 6. Southampton: National Institute for Health Research; 2016.
17. Department of Health. Going the extra mile: Improving the nation's health and wellbeing through public involvement in research. In: (NIHR) NIfHR, editor. Online: Department of Health; 2015.

18. Chalmers I, Bracken MB, Djulbegovic B, et al. How to increase value and reduce waste when research priorities are set. The Lancet. 2014;383(9912):156-65.

19. Dean C., Bierma H., Clarke R., et al. A Patient-Clinician James Lind Alliance Partnership to Identify Research Priorities for Hyperemesis Gravidarum. BMJ Open. 2021;11(1):e041254.

20. Clapton J, Rutter D, Sharif N. SCIE Systematic Mapping Guidance April 2009. Social Care Institue for Excellence; 2009.

21. White H, Albers B, Gaarder M, et al. Guidance for producing a Campbell evidence and gap map. Campbell Systematic Reviews. 2020;16(4).

22. Systematic evidence maps as a novel tool to support evidence-based decision-making in chemicals policy and risk management. Environ Int. 2019;130:104871.

23. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.

24. Campbell Colaboration. Evidence and gap map protocol template. Oslo, Norway: Campbell Collaboration; 2018. p. 8.

25. Collaboration C. Campbell's vision, mission and key principles Oslo: Campbell Collaboration; 2022 [Available from: https://www.campbellcollaboration.org/about-campbell/vision-mission-and-principle.html.

26. Brett J, Staniszewska S, Mockford C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect. 2014;17(5):637-50.

27. Amstar. AMSTAR - Assessing the Methodological Quality of Systematic Reviews 2021 [Available from: https://amstar.ca/Amstar_Checklist.php.

28. Dean C, Bannigan K, O'Hara M, et al. Recurrence rates of hyperemesis gravidarum in pregnancy: a systematic review protocol. JBI Database System Rev Implement Rep. 2017;15(11):2659-65.

29. Mitchell-Jones N, Gallos I, Farren J, et al. Psychological morbidity associated with hyperemesis gravidarum: a systematic review and meta-analysis. Bjog. 2017;124(1):20-30.

30. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan - A web and mobile app for systematic reviews. Syst Rev. 2016;5(210):1-10.

31. INVOLVE. Frequently asked questions - What is public involvement in research? London: NIHR; 2011 [updated 2011-04-19. Available from: https://www.invo.org.uk/frequently-asked-questions/.

32. INVOLVE. Public Information Pack (PIP) 1, How to get involved in NHS, public health and social care research - A quick guide. Southampton: INVOLVE; 2019.

33. Thomas J, Graziosi S, Brunton J, et al. EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis. EPPI-Centre Software. London: UCL Social Research Institute; 2020.

34. Power Z, Campbell M, Kilcoyne P, et al. The Hyperemesis Impact of Symptoms Questionnaire: Development and validation of a clinical tool. Int J Nurs Stud. 2010;47(1):67-77.

35. Maina A, Arrotta M, Cicogna L, et al. Transdermal clonidine in the treatment of severe hyperemesis. A pilot randomised control trial: CLONEMESI. Bjog. 2014;121(12):1556-62.

36. Guttuso T, Jr., Messing S, Tu X, et al. Effect of gabapentin on hyperemesis gravidarum: a double-blind, randomized controlled trial. Am J Obstet Gynecol MFM. 2021;3(1):100273.

37. Maina A, Todros T. A novel approach to hyperemesis gravidarum: evaluation by a visual analogue scale score and treatment with transdermal clonidine. Obstet Med. 2011;4(4):156-9.

38. Guttuso T, Robinson LK, Amankwah KS. Gabapentin use in hyperemesis gravidarum: A pilot study. Early Hum. Dev. 2010;86(1):65-6.

39. Westfall RE, Janssen PA, Lucas P, et al. Survey of medicinal cannabis use among childbearing women: patterns of its use in pregnancy and retroactive self-assessment of its efficacy against 'morning sickness'. Complement Ther Clin Pract. 2006;12(1):27-33.

40. Koren G, Cohen R. The use of cannabis for Hyperemesis Gravidarum (HG). J Cannabis Res. 2020;2(1):4.

41. Kul A, Gundogmus I, Aydin M. Rapid efficacy of mirtazapine in the treatment of hyperemesis gravidarum with esophagus perforation and ketonuria in a normoglycemic patient: a case report. Düşünen Adam. 2020;33:206-209

42. Spiegel D, Ramchandani J, Spiegel A, et al. A Case of Treatment-Refractory Hyperemesis Gravidarum Responsive to Adjunctive Mirtazapine in a Patient With Anxiety Comorbidity and Severe Weight Loss. J Clin Psychopharmacol. 2020;40(5):509-12.

43. Festin M. Nausea and Vomiting in Early Pregnancy. Clin. Evid. 2014;03(1405):1-35.

44. Sridharan K, Sivaramakrishnan G. Interventions for treating hyperemesis gravidarum: a network meta-analysis of randomized clinical trials. J Matern Fetal Neonatal Med. 2020;33(8):1405-11.

45. Van den Heuvel E, Goossens M, Vanderhaegen H, et al. Effect of acustimulation on nausea and vomiting and on hyperemesis in pregnancy: a systematic review of Western and Chinese literature. BMC Complement Altern Med. 2016;16:13.

46. McParlin C, Carrick-Sen D, Steen IN, et al. Hyperemesis in Pregnancy Study: a pilot randomised controlled trial of midwife-led outpatient care. Eur J Obstet Gynecol Reprod Biol. 2016;200:6-10.

47. Mitchell-Jones N, Farren JA, Tobias A, et al. Ambulatory versus inpatient management of severe nausea and vomiting of pregnancy: a randomised control trial with patient preference arm. BMJ Open. 2017;7(12):e017566.

48. Murphy A, McCarthy FP, McElroy B, et al. Day care versus inpatient management of nausea and vomiting of pregnancy: cost utility analysis of a randomised controlled trial. Eur J Obstet Gynecol Reprod Biol. 2016;197:78-82.

49. Grooten IJ, Koot MH, van der Post JA, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. Am. J. Clin. Nutr. 2017;106(3):812-20.

50. Budi G, Chalid S, Tiro E. Comparison between Blood Electrolyte and Ketonuria Preand Post- 5% Dextrose—Ringer's Lactate Rehydration Compared with Ringer's Lactate on Grade II Hyperemesis Gravidarum. South Asian Fed. Obstet. Gynecol. 2020;12(4):230-4.

51. Tan PC, Norazilah MJ, Omar SZ. Dextrose saline compared with normal saline rehydration of hyperemesis gravidarum: a randomized controlled trial. Obstet Gynecol. 2013;121(2 Pt 1):291-8.

52. Tan PC, Abdussyukur SA, Lim BK, et al. Twelve-hour fasting compared with expedited oral intake in the initial inpatient management of hyperemesis gravidarum: a randomised trial. Bjog. 2020;127(11):1430-7.

53. Gawande S, Vaidya M, Tadke R, et al. Progressive Muscle Relaxation in Hyperemesis Gravidarum. South Asian Fed. Obstet. Gynecol. 2011;3(1):28-32.

54. Shakiba M, Parsi H, Pahlavani Shikhi Z, et al. The Effect of Psycho-Education Intervention Based on Relaxation Methods and Guided Imagery on Nausea and Vomiting of Pregnant Women. J Family Reprod Health. 2019;13(1):47-55.

55. Dean C, Marsden J. Women's experiences of treatment for hyperemesis gravidarum in day case settings compared to hospital admissions. Midirs Digest. 2017;27(2):177-86.

56. Nicholson M. Women's experiences of the therapeutic value of writing about pregnancy sickness. Couns Psychother Res. 2018;18(1):26-34.

57. Power Z, Thomson A, Waterman H. Understanding the Stigma of Hyperemesis Gravidarum: Qualitative Findings from an Action Research Study. Birth-Issue Perinat Care. 2010;37(3):237-44.

58. Sykes C, Swallow B, Gadsby R, et al. Seeking medical help for Nausea and Vomiting in Pregnancy (NVP) and Hyperemesis Gravidarum (HG) in primary care. Midirs. 2013;9:13-5.

59. Tsalkitzi E, Nikolakopoulos P, Karanikas E. "Which baby?": a qualitative study exploring the fantasies about the unborn baby among women with hyperemesis gravidarum. J Psychosom Obstet Gynaecol. 2020:1-8.

60. van Vliet R, Bink M, Polman J, et al. Patient Preferences and Experiences in Hyperemesis Gravidarum Treatment: A Qualitative Study. J Pregnancy. 2018;2018:5378502.

61. Fejzo MS, Arzy D, Tian R, et al. Evidence GDF15 Plays a Role in Familial and Recurrent Hyperemesis Gravidarum. Geburtshilfe Frauenheilkd. 2018;78(9):866-70.

62. Fejzo MS, Ching C, Schoenberg FP, et al. Change in paternity and recurrence of hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2012;25(8):1241-5.

63. Fejzo MS, Fasching PA, Schneider MO, et al. Analysis of GDF15 and IGFBP7 in Hyperemesis Gravidarum Support Causality. Geburtshilfe Frauenheilkd. 2019;79(4):382-8.

64. Fejzo MS, Myhre R, Colodro-Conde L, et al. Genetic analysis of hyperemesis gravidarum reveals association with intracellular calcium release channel (RYR2). Mol Cell Endocrinol. 2017;439:308-16.

65. Fejzo MS, Sazonova OV, Sathirapongsasuti JF, et al. Placenta and appetite genes GDF15 and IGFBP7 are associated with hyperemesis gravidarum. Nat. Commun. 2018;9(1178):1-9.

66. Mullin PM, Bray A, Vu V, et al. No increased risk of psychological/behavioral disorders in siblings of women with hyperemesis gravidarum (HG) unless their mother had HG. J Dev Orig Health Dis. 2012;3(5):375-9.

67. Petry CJ, Ong KK, Burling KA, et al. Associations of vomiting and antiemetic use in pregnancy with levels of circulating GDF15 early in the second trimester: A nested case-control study. Wellcome Open Res. 2018;3:123.

68. Zhang YF, Cantor RM, MacGibbon K, et al. Familial aggregation of hyperemesis gravidarum. AJOG. 2011;204(3).

69. Li L, Li L, Zhou X, et al. Helicobacter pylori Infection Is Associated with an Increased Risk of Hyperemesis Gravidarum: A Meta-Analysis. Gastroenterol Res Pract. 2015;2015:278905.

70. Ng QX, Venkatanarayanan N, De Deyn M, et al. A meta-analysis of the association between Helicobacter pylori (H. pylori) infection and hyperemesis gravidarum. Helicobacter. 2018;23(1).

71. Sandven I, Abdelnoor M, Nesheim BI, et al. Helicobacter pylori infection and hyperemesis gravidarum: a systematic review and meta-analysis of case-control studies. Acta Obstet Gynecol Scand. 2009;88(11):1190-200.

72. Lindberg R, Lindqvist M, Trupp M, et al. Polyunsaturated Fatty Acids and Their Metabolites in Hyperemesis Gravidarum. Nutrients. 2020;12(11).

73. Maltepe C, Koren G. Preemptive treatment of nausea and vomiting of pregnancy: results of a randomized controlled trial. Obstet Gynecol Int. 2013;2013:809787.

74. O'Hara ME. Experiences of hyperemesis gravidarum in a subsequent pregnancy. Midirs. 2017;27(3):309-18.

75. Fejzo M, Poursharif B, Korst L, et al. Symptoms and Pregnancy Outcomes Associated with Extreme Weight Loss among Women with Hyperemesis Gravidarum. J Womens Health. 2009;18(12):1981-7.

76. Dinberu MT, Mohammed MA, Tekelab T, et al. Burden, risk factors and outcomes of hyperemesis gravidarum in low-income and middle-income countries (LMICs): systematic review and meta-analysis protocol. BMJ Open. 2019;9(4):e025841.

77. Oudman E, Wijnia JW, Oey M, et al. Wernicke's encephalopathy in hyperemesis gravidarum: A systematic review. Eur J Obstet Gynecol Reprod Biol. 2019;236:84-93.

78. Carstairs SD. Ondansetron Use in Pregnancy and Birth Defects: A Systematic Review. Obstet Gynecol. 2016;127(5):878-83.

79. Grooten IJ, Vinke ME, Roseboom TJ, et al. A Systematic Review and Meta-Analysis of the Utility of Corticosteroids in the Treatment of Hyperemesis Gravidarum. Nutr Metab Insights. 2015;8(Suppl 1):23-32.

80. Kaplan YC, Richardson JL, Keskin-Arslan E, et al. Use of ondansetron during pregnancy and the risk of major congenital malformations: A systematic review and meta-analysis. Reprod Toxicol. 2019;86:1-13.

81. Lavecchia M, Chari R, Campbell S, et al. Ondansetron in Pregnancy and the Risk of Congenital Malformations: A Systematic Review. JOGC. 2018;40(7):910-8.

82. Matthews A, Haas DM, O'Mathuna DP, et al. Interventions for nausea and vomiting in early pregnancy. Cochrane Database Syst. Rev. 2015(9):CD007575.

83. Yan R, Zhan J, Liu G, et al. A comparison of the efficacy and safety of traditional Chinese medicine external treatment for the hyperemesis gravidarum: A protocol for systematic review and network meta-analysis. Medicine (Baltimore). 2020;99(45):e23019.

84. Andrade C. Major Congenital Malformation Risk After First Trimester Gestational Exposure to Oral or Intravenous Ondansetron. J Clin Psychiatry. 2020;81(3).

85. Basirat Z, Barat S, Moghadamnia A. Comparing the effects of prednisolone and promethazine in the treatment of hyperemesis gravidarum: a double-blind, randomized clinical trial. Feyz. 2012;16(5):414-9.

86. Ostenfeld A, Petersen TS, Futtrup TB, et al. Validating the effect of Ondansetron and Mirtazapine In Treating hyperemesis gravidarum (VOMIT): protocol for a randomised placebo-controlled trial. BMJ Open. 2020;10(3):e034712.

87. Fejzo MS, Magtira A, Schoenberg FP, et al. Neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2015;189:79-84.

88. Peled Y, Melamed N, Hiersch L, et al. The impact of total parenteral nutrition support on pregnancy outcome in women with hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2014;27(11):1146-50.

89. Shapira M, Avrahami I, Mazaki-Tovi S, et al. The safety of early pregnancy exposure to granisetron. Eur J Obstet Gynecol Reprod Biol. 2020;245:35-8.

90. Fejzo M, Kam A, Laguna A, et al. Analysis of neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum reveals increased reporting of autism spectrum disorder. Reprod Toxicol. 2019;84:59-64.

91. Fejzo MS, Magtira A, Schoenberg FP, et al. Antihistamines and other prognostic factors for adverse outcome in hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2013;170(1):71-6.

92. Carrasco M, Rao SC, Bearer CF, et al. Neonatal Gabapentin Withdrawal Syndrome. Pediatr Neurol. 2015;53(5):445-7.

93. Ferreira E, Gillet M, Lelièvre J, et al. Ondansetron use during pregnancy: a case series. J. Popul. Ther. Clin. 2012;19(1):e1-e10.

94. McParlin C, O'Donnell A, Robson SC, et al. Treatments for Hyperemesis Gravidarum and Nausea and Vomiting in Pregnancy: A Systematic Review. JAMA. 2016;316(13):1392-401.

95. Al-Ozairi E, Waugh JJS, Taylor R. Termination is not the treatment of choice for severe hyperemesis gravidarum: Successful management using prednisolone. Obstet. Med. 2009;2(1):34-7.

96. Asmat A, Yasin I, Hamid I, et al. Is Prednisolone Useful in Treatment of Hyperemesis Gravidarum? Cureus. 2020;12(10):e11128.

97. Corona G, Simonetti L, Giuliani C, et al. A case of osmotic demyelination syndrome occurred after the correction of severe hyponatraemia in hyperemesis gravidarum. BMC Endo Dis. 2014;14(34):1-7.

98. Mamdouh Abdeldayem T, Samy El-agwany A, Elsayed Kholeif A. Long acting corticosteroids for the control of hyperemsis gravidarum and its effect on blood chloride level. Prog Obstet Ginecol. 2016;59(6).

99. Birkeland E, Stokke G, Tangvik RJ, et al. Norwegian PUQE (Pregnancy-Unique Quantification of Emesis and nausea) identifies patients with hyperemesis gravidarum and poor nutritional intake: a prospective cohort validation study. PLoS One. 2015;10(4):e0119962.

100. Dochez V, Dimet J, David-Gruselle A, et al. Validation of specific questionnaires to assess nausea and vomiting of pregnancy in a French population. Int J Gynaecol Obstet. 2016;134(3):294-8.

101. Fletcher SJ, Waterman H, Nelson L, et al. Holistic assessment of women with hyperemesis gravidarum: A randomised controlled trial. Int J Nurs Stud. 2015;52(11):1669-77.

102. Koot MH, Grooten IJ, van der Post JAM, et al. Determinants of disease course and severity in hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2020;245:162-7.

103. Morris ZH, Azab AN, Harlev S, et al. Developing and validating a prognostic index predicting re-hospitalization of patients with Hyperemesis Gravidarum. Eur J Obstet Gynecol Reprod Biol. 2018;225:113-7.

104. Madendag Y, Sahin E, Madendag Col I, et al. The effect of hyperemesis gravidarum on the 75 g oral glucose tolerance test screening and gestational diabetes mellitus. J Matern Fetal Neonatal Med. 2018;31(15):1989-92.

105. Morgan SR, Long L, Johns J, et al. Are early pregnancy complications more common in women with hyperemesis gravidarum? J Obstet Gynaecol. 2017;37(3):355-7.

106. Ohara R, Obata-Yasuoka M, Abe K, et al. Effect of hyperemesis gravidarum on gestational diabetes mellitus screening. Int J Gynaecol Obstet. 2016;132(2):156-8.

107. Peled Y, Melamed N, Krissi H, et al. The impact of severe hyperemesis gravidarum on the triple test screening results. J Matern Fetal Neonatal Med. 2012;25(6):637-8.

108. Tan PC, King AS, Omar SZ. Screening for urinary tract infection in women with hyperemesis gravidarum. J Obstet Gynaecol Res. 2012;38(1):145-53.

109. Tulek F, Kahraman A, Taskin S, et al. Changes in first trimester screening test parameters in pregnancies complicated by placenta previa and association with hyperemesis gravidarum. J Turk Ger Gynecol Assoc. 2014;15(4):212-6.

110. Niemeijer MN, Grooten IJ, Vos N, et al. Diagnostic markers for hyperemesis gravidarum: a systematic review and metaanalysis. Am J Obstet Gynecol. 2012;211(2):150.e1-.e15.

111. Carbone IF, Conforti A, Picarelli S, et al. Circulating Nucleic Acids in Maternal Plasma and Serum in Pregnancy Complications: Are They Really Useful in Clinical Practice? A Systematic Review. Mol Diagn Ther. 2020;24(4):409-31.

112. Maslin K, Shaw V, Brown A, et al. What is known about the nutritional intake of women with Hyperemesis Gravidarum?: A scoping review. Eur J Obstet Gynecol Reprod Biol. 2021;257:76-83.

113. Dean C, Marsden J. Satisfaction for treatment of hyperemesis gravidarum in day case settings compared to hospital admissions. Midirs. 2017;27(1):11-20.

114. Havnen GC, Truong MB, Do MH, et al. Women's perspectives on the management and consequences of hyperemesis gravidarum - a descriptive interview study. Scand J Prim Health Care. 2019;37(1):30-40.

115. Mitchell-Jones N, Lawson K, Bobdiwala S, et al. Association between hyperemesis gravidarum and psychological symptoms, psychosocial outcomes and infant bonding: a two-point prospective case–control multicentre survey study in an inner city setting. BMJ Open. 2020;10(10).

116. Mullin P, Bray A, Schoenberg F, et al. Prenatal exposure to hyperemesis gravidarum linked to increased risk of psychological and behavioral disorders in adulthood. J Dev Orig Health Dis. 2011;2(4):200-4.

117. Christodoulou-Smith J, Gold J, Romero R, et al. Posttraumatic stress symptoms following pregnancy complicated by hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2011;24(11):1307-11.

118. Dean CR, O'Hara ME. Ginger is ineffective for hyperemesis gravidarum, and causes harm: an internet based survey of sufferers. MIDIRS. 2015;25(4):449-55.

119. Mullin PM, Ching C, Schoenberg F, et al. Risk factors, treatments, and outcomes associated with prolonged hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2012;25(6):632-6.

120. Saleh A, Sykes C. The impact of online information on health related quality of life amongst women with nausea and vomiting in pregnancy and hyperemesis gravidarum. MIDIRS. 2014;24(2):179-85.

121. Fejzo MS, MacGibbon KW, Mullin PM. Ondansetron in pregnancy and risk of adverse fetal outcomes in the United States. Reprod Toxicol. 2016;62:87-91.

122. Heitmann K, Nordeng H, Havnen GC, et al. The burden of nausea and vomiting during pregnancy: severe impacts on quality of life, daily life functioning and willingness to become pregnant again - results from a cross-sectional study. BMC Pregnancy Childbirth. 2017;17(1):75.

123. Magtira A, Schoenberg FP, MacGibbon K, et al. Psychiatric factors do not affect recurrence risk of hyperemesis gravidarum. J Obstet Gynaecol Res. 2015;41(4):512-6.

124. Tian R, MacGibbon K, Martin B, et al. Analysis of pre- and post-pregnancy issues in women with hyperemesis gravidarum. Auton Neurosci. 2017;202:73-8.

125. Davies R. Constant sickness is not good news. BMJ. 2018;363:k4208.

126. Dean C. A patient experience of hyperemesis gravidarum and how the midwife can support her care. Essentially MIDIRS. 2014;5(2):32-6.

127. Miake-Lye IM, Hempel S, Shanman R, et al. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. Syst Rev. 2016;5:28.

128. Peter Bragge, Ornella Clavisi, Tari Turner, et al. The Global Evidence Mapping Initiative: Scoping research in broad topic areas. BMC Med Res Methodol. 2011;11:1-12.

129. World Health Organization and International Initiative for Impact Evaluation. An evidence map of social, behavioural and community engagement interventions for reproductive, maternal, newborn and child health. Geneva: World health Organization; 2017.

130. Boelig RC, Barton SJ, Saccone G, et al. Interventions for treating hyperemesis gravidarum: a Cochrane systematic review and meta-analysis. J Matern Fetal Neonatal Med. 2018;31(18):2492-505.

131. Matthews A, Haas DM, O'Mathuna DP, et al. Interventions for nausea and vomiting in early pregnancy. Cochrane Database Syst. Rev. 2015;3:CD007575.

132. Salman RA-S, Beller E, Kagan J, et al. Increasing value and reducing waste in biomedical research regulation and management. The Lancet. 2014;383(9912):176-85.

133. Yingju Wang, Mengyi Lv, Liu. S. The effectiveness and safety of acupuncture in the treatment of nausea and vomiting of pregnancy and hyperemesis gravidarum: a meta-analysis of randomized controlled studies. PROSPERO: International prospective register of systematic reviews. 2018;CRD42018118919.

134. Hai-Zhen Lu, Zheng C-W. Effectiveness of acupuncture in treatment of hyperemesis gravidarum : A systematic review and meta-analysis. PROSPERO: International prospective register of systematic reviews. 2021;2021:CRD42021232187.

135. Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. BMJ. 2013;347:f4501.

136. Moher D, Booth A, Stewart L. How to reduce unnecessary duplication: use PROSPERO. Bjog. 2014;121(7):784-6.

137. Tawfik GM, Giang HTN, Ghozy S, et al. Protocol registration issues of systematic review and meta-analysis studies: a survey of global researchers. BMC Med Res Methodol. 2020;20(1):213.

138. Marshall Z, Welch V, Minichiello A, et al. Documenting Research with Transgender, Nonbinary, and Other Gender Diverse (Trans) Individuals and Communities: Introducing the Global Trans Research Evidence Map. Transgend Health. 2019;4(1):68-80.

139. Gonzalez AG, Schmucker C, Nothacker J, et al. Health-related preferences of older patients with multimorbidity: an evidence map. BMJ Open. 2019;9(12):e034485.

140. Blackburn S, McLachlan S, Jowett S, et al. The extent, quality and impact of patient and public involvement in primary care research: a mixed methods study. Res Involv Engagem. 2018;4:16.

141. Skovlund PC, Nielsen BK, Thaysen HV, et al. The impact of patient involvement in research: a case study of the planning, conduct and dissemination of a clinical, controlled trial. Res Involv Engagem. 2020;6:43.

142. Koot MH, Boelig RC, Van't Hooft J, et al. Variation in hyperemesis gravidarum definition and outcome reporting in randomised clinical trials: a systematic review. Bjog. 2018;125(12):1514-21.

143. Jansen LAW, Koot MH, Van't Hooft J, et al. The windsor definition for hyperemesis gravidarum: A multistakeholder international consensus definition. Eur J Obstet Gynecol Reprod Biol. 2021;266:15-22.

PART II
MATERNAL FUTURE HEALTH

CHAPTER 3

Thyroid-stimulating hormone and free thyroxine fail to predict the severity and clinical course of hyperemesis gravidarum: a prospective cohort study

Kelly Nijsten, Marjette H. Koot, Joris A.M. van der Post, Joke M.J. Bais, Carrie Ris-Stalpers, Christiana Naaktgeboren, Henk A. Bremer, David P. van der Ham, Wieteke M. Heidema, Anjoke Huisjes, Gunilla Kleiverda, Simone M. Kuppens, Judith O.E.H. van Laar, Josje Langenveld, Flip van der Made, Dimitri Papatsonis, Marie-José Pelinck, Paula J. Pernet, Leonie van Rheenen-Flach, Robbert J. Rijnders, Hubertina C.J. Scheepers, Sarah E. Siegelaar, Tatjana Vogelvang, Ben W. Mol, Prof. Tessa J. Roseboom, Iris J. Grooten, Rebecca C. Painter

Acta Obstetricia et Gynecologica Scandinavica, 2021; 100: 1419-1429.

ABSTRACT

Introduction

Little is known about the pathophysiology of hyperemesis gravidarum (HG). Proposed underlying causes are multifactorial and thyroid function is hypothesized to be causally involved. In this study, we aim to assess the utility of thyroid stimulating-hormone (TSH) and free thyroxine (FT4) as a marker and predictor for the severity and clinical course of HG.

Material and methods

We conducted a prospective cohort study including women admitted for HG between 5 and 20 weeks gestation in 19 hospitals in the Netherlands. Women with a medical history of thyroid disease were excluded. TSH and FT4 were measured at study entry. To adjust for gestational age, we calculated TSH Multiples of the Median (MoM). We assessed HG severity at study entry as severity of nausea and vomiting (by the Pregnancy Unique Quantification of Emesis and nausea score), weight change compared to pre pregnancy weight and quality of life. We assessed the clinical course of HG as severity of nausea and vomiting and quality of life 1 week after inclusion, duration of hospital admissions and readmissions. We performed multivariable regression analysis with absolute TSH, TSH MoMs and FT4.

Results

Between 2013 and 2016, 215 women participated in the cohort. TSH, TSH MoM and FT4 were available for respectively 150, 126 and 106 of these women. Multivariable linear regression analysis showed that lower TSH MoM was significantly associated with increased weight loss or lower weight gain at study entry (ΔKg; β=2.00, 95% CI: 0.47 - 3.53), whereas absolute TSH and FT4 were not. Lower TSH, not lower TSH MoM or FT4, was significantly associated with lower nausea and vomiting scores one week after inclusion (β=1.74, 95% CI: 0.36 - 3.11). TSH and FT4 showed no association with any of the other markers of the severity or clinical course of HG. Twenty-one out of 215 (9.8%) women had gestational transient thyrotoxicosis. Women with GTT had a lower quality of life 1 week after inclusion then women with no gestational transient thyrotoxicosis (*P*=0.03).

Conclusions

Our findings show an inconsistent role for TSH, TSH MoM or FT4 at time of admission and provide little guidance on the severity and clinical course of HG.

BACKGROUND

Hyperemesis gravidarum (HG) is a severe form of nausea and vomiting in pregnancy (NVP) affecting 0.3-3.6% of pregnancies.[1, 2] HG is the most common reason for hospital admission in early pregnancy, but evidence based effective treatment options are currently limited.[3, 4] Because of the major impact of HG on maternal wellbeing and quality of life, a marker that could help identify the severity and clinical course of HG would be of value for assessing a patient's prognosis and individualizing patient care.[5-8]

Little is known about the pathophysiology of HG. Proposed underlying causes are multifactorial and related to maternal endocrine and placental function as well as gastrointestinal conditions, although recently genetic causes, including involvement of Growth differentiation factor-15 (GDF 15), a cachexia gene, and its receptor have been implicated.[9, 10]

Increased Human Chorionic Gonadotropin (hCG) has been hypothesized to be causally involved in NVP and HG: hCG rises in the first trimester, which coincides with the peak in occurrence of NVP and HG.[11] However, the putative association of increased hCG with NVP symptoms was not confirmed upon systematic review, which found an association in only half of the included studies.[12] An explanation for the apparent discrepancy in findings could lie in the fact that hCG and Thyroid Stimulating Hormone (TSH) have biosimilar effects on the TSH receptor, which can result in hCG induced thyroid stimulation leading to a clinically relevant rise in free thyroxine (FT4) and subsequent suppression of TSH levels, a condition known as gestational transient thyrotoxicosis (GTT).[13] Hyperthyroidism can produce nausea, vomiting and weight loss. Therefore, FT4 and TSH might be important factors in the etiology of HG.[14] A systematic review from Niemeijer *et al.*[12] found few studies assessing the effect of TSH on the severity and clinical course of HG. We updated this systematic review[12] in order to put our findings in context, as shown in **Appendix A**. Available evidence concerning possible associations between TSH and FT4 and measures of HG severity and clinical course shows conflicting results. None of the studies took gestational fluctuations of TSH into account. Niemeijer *et al.*[12] called for better investigation of the use of TSH adjusted for gestational duration in future studies and currently no adjusted reference interval is available for gestational fluctuations of TSH.

In the present study we aim to assess the association between absolute TSH (without adjustments for gestational fluctuations), TSH Multiple of the Median (MoM) and FT4 and the severity and clinical course of HG to evaluate whether it is useful to measure thyroid function in order to predict the severity and clinical course of HG.

MATERIALS AND METHODS

Our study used data of the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) trial and associated cohort and is a prospective observational cohort study carried out between 2013 and 2016, that included women admitted for HG in 19 hospitals in the Netherlands.[15] The MOTHER trial, a multicenter open-label randomized controlled trial (RCT), aimed to evaluate the effect of early enteral tube feeding in addition to standard care for hyperemesis gravidarum patients requiring hospital inpatient care, including intravenous fluid and anti-emetic medication. Women who were eligible for the RCT but who declined participation were asked to participate in the cohort. Since early enteral tube feeding had no effect on perinatal and maternal outcomes we deemed it appropriate to combine data of the RCT and cohort into one study population for the present study.[15] Participants of both trial and cohort provided informed consent.

We included women of 18 years and older who had been admitted to the hospital for HG between 5 and 20 weeks' gestation. Women were diagnosed with HG if they had severe nausea and vomiting necessitating admission. More detailed description of study methodology has been previously published.[15]

Thyroid Stimulating Hormone (TSH) and Free Thyroxine (FT4)

According to local protocol maternal blood was taken during routine laboratory on the first day of admission and analyzed as well as stored frozen in the biobank, as reported in the previously published MOTHER protocol.[16] If TSH was included in local routine work-up for HG, TSH was assessed in local laboratory at baseline. In August 2019, available frozen stored maternal samples from baseline were used to assess FT4 in one central laboratory.

We assessed TSH unadjusted for gestational fluctuations (absolute TSH), but we also calculated TSH Multiples of the Medians (MoMs) in order to account for physiological fluctuation of TSH during pregnancy. To do so, we used data from the CATS study from Bestwick *et al.*,[17] the largest available study in Europe which assessed TSH medians in healthy pregnant women between 7 and 15 weeks gestation. We calculated Multiples of the Medians (MoMs) by dividing the participants' observed absolute TSH concentration by the expected TSH median for corresponding gestational age, as published by Bestwick *et al.*[17] GTT was defined as FT4 above 22.0 pmol/L according to the central laboratory's reference interval.[18]

Data collection

Trained research staff completed a Case Report Form (CRF) to extract and report information from medical and obstetric antenatal files including age, parity, gestational age, weight and comorbidity at study entry. Pre-pregnancy weight, ethnicity and education level were self-reported at baseline and if not reported, extracted from medical file, if available.

HG severity was measured at baseline by evaluating weight change, symptom severity and quality of life. Weight change was defined as weight at baseline minus prepregnancy weight. Symptom severity and quality of life were measured by three self-reported validated questionnaires, filled out on the first day of inclusion. The Pregnancy Unique Quantification of Emesis and nausea (PUQE-24) measures the severity of nausea and vomiting: a higher PUQE-24 score (PUQE-24 ≥13) indicates severe vomiting.[19] The Hyperemesis Impact of Symptoms (HIS) questionnaire determines the impact of nausea and vomiting.[20] The Nausea and Vomiting in Pregnancy Quality of Life questionnaire (NVPQoL) measures the impact of nausea and vomiting specific quality of life.[21] A higher NVPQoL- or HIS-score indicates a lower quality of life or higher impact on maternal wellbeing.

The clinical course of HG was measured as symptom severity and quality of life one week after inclusion, duration of hospital admissions and readmissions. We measured symptom severity by PUQE-24 score one week after inclusion and quality of life by NVPQoL and HIS score one week after inclusion. Duration of hospital admissions and readmissions were collected from medical files. Duration of hospital admissions was measured in days whereas the day of admission and discharge both counted as one day.

Statistical analyses

Normally distributed continuous variables are presented as means with standard deviations (SDs) and skewed distributions as medians with inter quartile ranges (IQRs). Dichotomous and categorical variables are presented as frequencies with percentages.

We performed univariable and multivariable linear regression analysis to assess the association between absolute TSH, TSH MoM and FT4 and continuous outcome variables. NVPQoL one week after inclusion and total days of hospital admission were not normally distributed and were logarithmically transformed to achieve normality. These logarithmically transformed variables were back-transformed and reported in proportionate differences and 95% confidence intervals (95% CIs, expressed as percentages); normally distributed variables were reported in differences (β) and 95% CI. Dichotomous outcome variables were analyzed using univariable

and multivariable logistic regression analysis and are reported in odds ratios (OR) with 95% CI. In the first model we performed a univariable regression analysis. In the second model we performed a multivariable regression analysis adjusted for age, prepregnancy BMI, ethnicity and maternal education. Weight change was added as confounder in the multivariable analysis for clinical course of HG based on available literature.[22]

A sensitivity analysis was used for assessing whether baseline characteristics and measures of the severity and clinical course of HG differed between included and excluded participants in our study. We also assessed whether baseline characteristics and measures of the severity and clinical course of HG differed between women with and without GTT, using chi-square test, Fisher's exact test, Mann–Whitney U test and independent Student's *t* test. *P*-values <0.05 were considered statistically significant. SPSS Statistics 25.0 for Windows (IBM Corp., Armonk, NY, USA) was used for all analysis.

Ethical approval

The MOTHER trial and cohort were registered at www.trialregister.nl (NTR4197) and have been approved by the research ethics committee of the Academic Medical Centre on the third of April 2013. (NL41011.018.12)

RESULTS

215 women were included in the combined cohort: 115 women in the RCT and another 100 women in the associated observational cohort. Nine women with a medical history of hypo- or hyperthyroidism or with a clinical hypothyroidism (TSH ≥ 4.0 mE/L) at time of inclusion were excluded from the analysis **(Figure 1)**. TSH at time of inclusion was measured in 150 women. We were able to calculate TSH MoMs for 126 women, based on TSH medians available from Bestwick *et al.*[17], and we were able to measure FT4 in 106 women, based on the number frozen stored blood samples available.

Baseline characteristics and measurements of the clinical course of HG and disease severity of participants are shown in **Table 1**. First admission at inclusion significantly differed between women with (92.7%) and without absolute TSH levels available (80.0%, *P*=0.01) **(Table S2).** Gestational age (*P*<0.01) and first admission (*P*=0.03) at inclusion significantly differed between women with and without TSH MoM available, but measurements of the severity and clinical course of HG did not (**Table S3**). We did not find any significant differences in baseline characteristics and outcome variables between women with and without FT4 available **(Table S4)**.

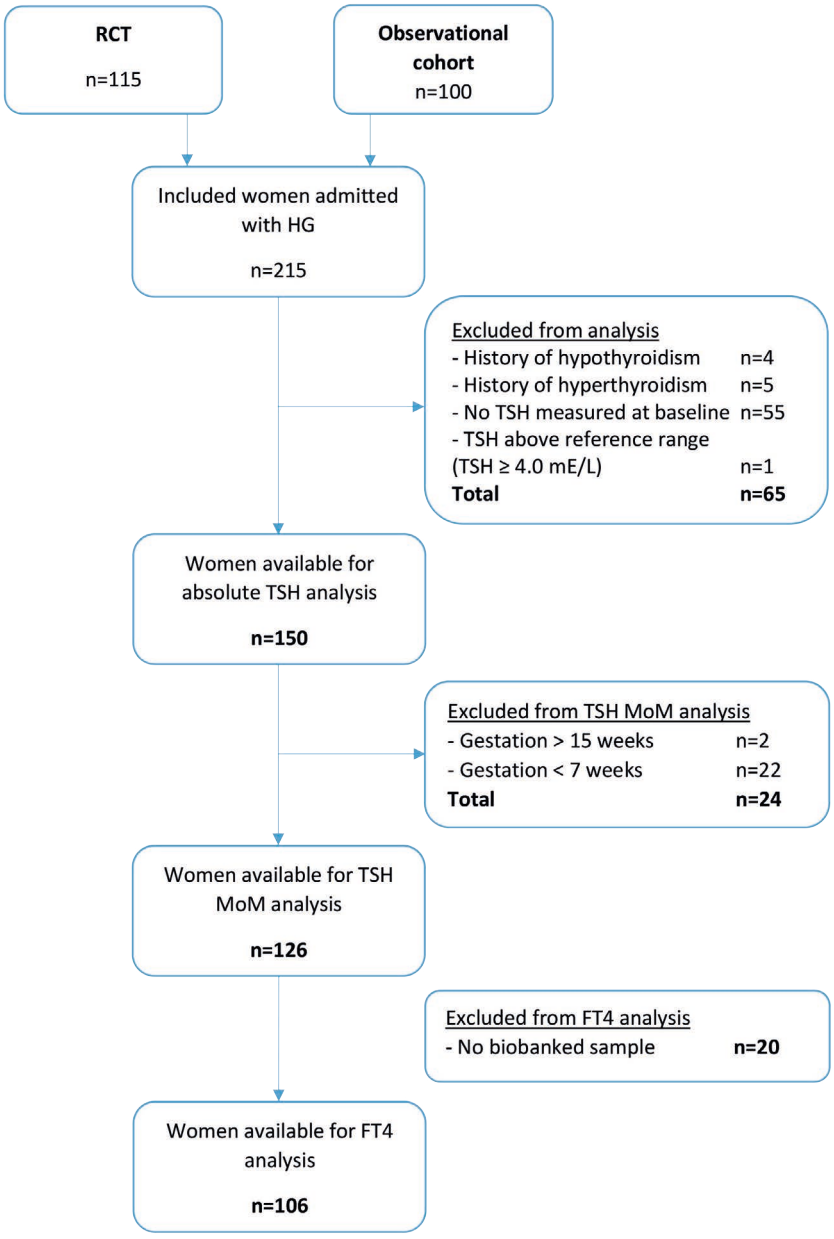


Figure 1. Flowchart in- and exclusions study population

Table 1. Baseline characteristics for women admitted for HG included in this cohort

	N=215	% missing
Demographics		
Age (years)	28.83±4.83	0.0%
Prepregnancy weight (kg)	71.11±15.02	2.3%
Prepregnancy BMI (kg/m^2)	25.12±4.89	3.7%
Ethnic origin		18.1%
- Western	123 (57.2%)	
- Non-western	53 (24.7%)	
Education level		34.0%
- Primary or secondary	86 (40.0%)	
- Higher	56 (26.0%)	
Mental health disorder in medical history[1]	41 (19.1%)	0.0%
HG in previous pregnancy[2]	68 (45.0%)	15.2%
HG in previous pregnancy requiring hospital admission[2]	37 (24.5%)	10.3%
Thyroid Stimulating Hormone (TSH)	0.78±0.61	30.2%
Free thyroxine (FT4)	19.26±4.76	50.7%
Pregnancy characteristics		
Primigravida	64 (29.8%)	0.0%
Twin pregnancy	5 (2.3%)	0.0%
Gestational age at onset of symptoms of HG (weeks)	6.00 (5.25-7.00)	23.3%
Gestational age at inclusion	9.00 (7.00-11.00)	0.0%
First admission at study entry	191 (88.8%)	0.0%
Outcomes		
HG severity at baseline		
- Weight change (kg)	-2.92±4.07	2.8%
- PUQE-24	10.01±3.30	37.2%
- NVPQoL	173.44±23.43	34.9%
- HIS	27.77±3.86	34.4%
Clinical course of HG		
- PUQE-24 one week after inclusion	9.00 (6.00-11.00)	45.6%
- NVPQoL one week after inclusion	76.00 (61.00-100.50)	49.3%
- HIS one week after inclusion	25.71±3.82	49.3%
- Duration first admission (days)	4.00 (3.00-5.00)	0.0%
- Total days of hospital admission for HG	5.00 (3.00-8.00)	0.0%
- Readmitted	71 (33.0%)	0.0%
- Readmitted ≥ two times	29 (13.5%)	0.0%

Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). [1] Mental health disorder consists of an eating disorder, anxiety disorder or a depressive disorder. [2] Percentage shown is frequency divided by number of multigravidas. BMI: body mass index. FT4: free thyroxine. HG: Hyperemesis Gravidarum. TSH: Thyroid Stimulating Hormone. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HIS: Hyperemesis Impact of Symptoms. NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life. A higher PUQE-24, HIS- or NVPQoL-score indicates more severe symptoms or lower quality of life.

Lower TSH MoM was significantly associated with increased weight loss at baseline compared to prepregnancy weight (Kg; β=2.00, 95% CI: 0.47 - 3.53, *P*=0.01), whereas absolute TSH and FT4 were not (**Table 2 and 3**). Neither TSH, nor TSH MoM, nor FT4 were significantly associated with other markers of HG severity at baseline including the PUQE-24-, HIS- and NVPQoL-score.

Regarding the association between TSH and FT4 and the clinical course of HG, we found that lower absolute TSH was significantly associated with lower nausea and vomiting scores one week after inclusion in multivariable regression analysis (β=1.74, 95% CI: 0.36 - 3.11, *P*=0.01) (**Table 2a)**. Lower TSH MoM was only significantly associated with lower nausea and vomiting scores one week after inclusion in univariable regression analysis (β=1.41, 95% CI: 0.11 - 2.72, *P*=0.03) **(Table 2b).** FT4 was not associated with the severity of nausea and vomiting one week after inclusion **(Table 3)**. No significant association was found between TSH, TSH MoM or FT4 and quality of life one week after inclusion, duration of hospital admissions and readmissions.

Comparing women with and without GTT at baseline, we found that women with GTT (n=21) had a higher HIS score one week after inclusion then women with no GTT (*P*=0.03) as shown in **Table 4**. No significant differences in baseline characteristics or other outcome variables were found between women with and without GTT.

Table 2. Multivariable linear and logistic regression to assess the association between absolute TSH (Table 2a) and TSH MoM (Table 2b) and measures of the severity and clinical course of HG

Table 2a. Absolute TSH	Model 1			Model 2		
HG severity at baseline	***ß***	**95% CI**	***P***	***ß***	**95% CI**	***P***
-Weight change	0.82	-0.20 to 1.84	0.12	0.94	-0.43 to 2.32	0.18
-PUQE-24	-0.15	-1.21 to 0.91	0.78	-0.43	-1.60 to 0.74	0.47
-NVPQoL	-5.71	-12.19 to 0.78	0.08	-5.63	-12.90 to 1.65	0.13
-HIS	0.49	-0.67 to 1.65	0.40	1.03	-0.35 to 2.41	0.14
Clinical course of HG						
-PUQE-24 one week after inclusion	1.02	-0.03 to 2.06	0.06	1.74	0.36 - 3.11	0.01
-NVPQoL one week after inclusion ∫	8.65	-5.92 to 25.36	0.26	1.11	-15.55 to 21.17	0.90
-HIS one week after inclusion	0.09	-1.30 to 1.48	0.90	1.04	-0.69 to 2.76	0.23
-Total days of hospital admission for HG ∫	2.43	-14.02 to 21.90	0.79	-0.20	-22.89 to 29.30	0.99
	OR	**95% CI**	***P***	**OR**	**95% CI**	***P***
-Readmitted yes or no	1.55	0.90 - 2.68	0.11	1.87	0.85 - 4.13	0.12
-Readmitted ≥ two times	1.14	0.55 - 2.38	0.72	0.76	0.27 - 2.09	0.59

Table 2b. TSH MoM	Model 1			Model 2		
HG severity at baseline	***ß***	**95% CI**	***P***	***ß***	**95% CI**	***P***
-Weight change	1.54	0.35 - 2.73	0.01	2.00	0.47 - 3.53	0.01
-PUQE-24	0.24	-0.96 to 1.43	0.70	0.13	-1.26 to 1.52	0.85
-NVPQoL	-2.71	-9.80 to 4.38	0.45	-4.28	-12.31 to 3.76	0.29
-HIS	0.75	-0.64 to 2.14	0.29	1.33	-0.39 to 3.05	0.13
Clinical course of HG						
-PUQE-24 one week after inclusion	1.41	0.11 - 2.72	0.03	1.74	-0.02 to 3.50	0.05
-NVPQoL one week after inclusion ∫	4.50	-11.22 to 23.00	0.59	-2.27	-20.71 to 20.32	0.82
-HIS one week after inclusion	-0.07	-1.68 to 1.54	0.93	0.90	-1.08 to 2.88	0.37
-Total days of hospital admission for HG ∫	-0.50	-19.99 to 23.74	0.96	-4.97	-33.30 to 35.53	0.77
	OR	**95% CI**	***P***	**OR**	**95% CI**	***P***
-Readmitted yes or no	1.27	0.64 - 2.50	0.50	1.72	0.62 - 4.80	0.30
-Readmitted ≥ two times	1.49	0.62 - 3.59	0.38	1.23	0.36 - 4.23	0.75

A *P*-value <0.05 was considered significant, represented with the corresponding 95% confidence interval (95% CI). *β* is the unstandardized regression coefficient. OR is the odds ratio. ∫ is log transformed, back transformed and expressed in % of difference. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HG: Hyperemesis gravidarum, HIS: Hyperemesis Impact Score, NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life, PUQE: 24-hour Pregnancy Unique Quantification of Emesis and nausea. A higher HIS- or NVPQoL-score indicates lower quality of life or higher impact on maternal wellbeing. A higher PUQE-24 score indicates more severe symptoms.

Measures of HG severity at baseline as outcome
Model 1: univariable regression analysis
Model 2: multivariable regression analysis adjusted for age, prepregnancy BMI, ethnicity (western or not) and education level

Measures of the clinical course of HG as outcome
Model 1: univariable regression analysis
Model 2: multivariable regression analysis adjusted for age, prepregnancy BMI, weight change at baseline, ethnicity (western or not) and education level

Table 3. Multivariable linear and logistic regression to assess the association between **FT4** and measures of the severity and clinical course of HG

	Model 1			Model 2		
HG severity at baseline	**ß**	**95% CI**	***P***	**ß**	**95% CI**	***P***
-Weight change	-0.13	-0.29 to 0.04	0.13	-0.14	-0.35 to 0.06	0.16
-PUQE-24	-0.05	-0.19 to 0.10	0.52	-0.05	-0.23 to 0.12	0.54
-NVPQoL	0.65	-0.40 to 1.70	0.22	0.55	-0.63 to 1.74	0.35
-HIS	-0.02	-0.19 to 0.15	0.81	-0.06	-0.26 to 0.14	0.55
Clinical course of HG						
-PUQE-24 one week after inclusion	-0.03	-0.21 to 0.15	0.72	0.01	-0.21 to 0.23	0.94
-NVPQoL one week after inclusion ∫	-0.20	-2.37 to 2.12	0.87	0.80	-1.69 to 3.36	0.53
-HIS one week after inclusion	0.15	-0.05 to 0.36	0.13	0.07	-0.15 to 0.29	0.50
-Total days of hospital admission for HG ∫	2.22	-0.50 to 5.02	0.10	3.05	-0.90 to 7.04	0.13
	OR	**95% CI**	***P***	**OR**	**95% CI**	***P***
-Readmitted yes or no	1.03	0.95 - 1.12	0.48	1.02	0.91 - 1.14	0.72
-Readmitted ≥ two times	1.02	0.92 - 1.14	0.67	1.09	0.96 - 1.23	0.19

A *P*-value <0.05 was considered significant, represented with the therefore corresponding 95% confidence interval (95% CI). β is the unstandardized regression coefficient. OR is the odds ratio. **∫** is log transformed, back transformed and expressed in % of difference. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HG: Hyperemesis gravidarum, HIS: Hyperemesis Impact Score, NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life, PUQE: 24-hour Pregnancy Unique Quantification of Emesis and nausea. A higher HIS- or NVPQoL-score indicates lower quality of life or higher impact on maternal wellbeing. A higher PUQE-24 score indicates more severe symptoms.

Measures of HG severity at baseline as outcome
Model 1: univariable regression analysis
Model 2: multivariable regression analysis adjusted for gestational age, age, prepregnancy BMI, ethnicity (western or not) and education level

Measures of clinical course of HG as outcome
Model 1: univariable regression analysis
Model 2: multivariable regression analysis adjusted for gestational age, age, prepregnancy BMI, weight change at baseline, ethnicity (western or not) and education level

Table 4. Baseline characteristics and outcomes between women admitted for HG included in this study with and without Gestational Thyrotoxicosis (GTT)

	GTT N=21	No GTT N=85	*P*
Demographics			
Age (years)	29.67±4.72	28.45±4.90	0.31
Prepregnancy weight (kg)	69.40±15.32	71.68±15.41	0.55
Prepregnancy BMI (kg/m^2)	25.12±5.41	25.54±5.17	0.75
Ethnic origin			0.22
- Western	9 (42.9%)	50 (58.8%)	
- Non-western	8 (38.1%)	23 (27.1%)	
Education level			0.13
- Primary or secondary	5 (23.8%)	39 (45.9%)	
- Higher	9 (42.9%)	24 (28.2%)	
Mental health disorder in medical history[1]	2 (9.5%)	17 (20.0%)	0.35
HG in previous pregnancy[2]	9 (60%)	19 (34.5%)	0.08
HG in previous pregnancy requiring hospital admission[2]	4 (26.7%)	10 (18.2%)	1.00
Thyroid Stimulating Hormone (TSH)	0.33±0.43	0.89±0.67	0.00
Free thyroxine (FT4)	26.37±5.59	17.50±2.26	0.00
Pregnancy characteristics			
Primigravida	6 (28.6%)	30 (35.3%)	0.56
Twin pregnancy	1 (4.8%)	2 (2.4%)	0.49
Gestational age at onset of symptoms of HG (weeks)	6.00 (5.00-6.50)	6.00 (5.00-7.00)	0.60
Gestational age at baseline	8.00 (7.50-9.50)	8.00 (7.00-10.00)	0.89
First admission at study entry	18 (85.7%)	79 (92.9%)	0.38
Outcomes			
HG severity at baseline			
- Weight change (kg)	-3.87±3.07	-2.96±4.08	0.35
- PUQE-24	9.93±3.77	10.36±3.02	0.65
- NVPQoL	176.21±14.28	173.15±23.58	0.64
- HIS	26.80±2.83	28.13±3.96	0.23
Clinical course of HG			
- PUQE-24 one week after inclusion	9.00 (7.50-10.00)	9.00 (6.00-12.00)	1.00
- NVPQoL one week after inclusion	73.50 (57.25-84.75)	79.00 (58.75-107.25)	0.47
- HIS one week after inclusion	27.70±1.64	24.98±3.65	0.03
- Duration first admission (days)	4.00 (3.00-5.00)	4.00 (3.00-5.00)	0.37
- Total days of hospital admission for HG	6.00 (4.00-10.50)	5.00 (3.00-7.50)	0.20
- Readmitted	9 (42.9%)	32 (37.6%)	0.66
- Readmitted ≥ two times	4 (19.0%)	11 (12.9%)	0.49

Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). A *P*-value <0.05 was considered significant. [1] Mental health disorder consists of an eating disorder, anxiety disorder or a depressive disorder. [2] Percentage shown is frequency divided by number of multigravidas. BMI: body mass index. GTT: Gestational transient thyrotoxicosis: defined as women with a free thyroxine (FT4) level above 22.0 pmol/L. HG: Hyperemesis Gravidarum. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HIS: Hyperemesis Impact of Symptoms. NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life. A higher PUQE-24, HIS- or NVPQoL-score indicates more severe symptoms or lower quality of life.

COMMENTS

Main findings

We found that, in line with expectations, women who had lower TSH MoM, had more weight loss upon hospitalization with HG. Furthermore, 10% of women hospitalized with HG had concurrent GTT, associated with higher impact of symptoms scores (HIS) one week after inclusion. In contrast to literature, and not in line with expectations, lower TSH was also associated with markedly lower nausea and vomiting scores one week after baseline. Thyroid function showed no association with any of the other outcome measures of disease severity or course of HG. The fact that our findings present an inconsistent role for thyroid function in HG does not support use of thyroid measurements as marker or predictor for HG disease severity and clinical course.

Interpretation

In our search to identify markers of disease severity among newly diagnosed HG patients, our study found that lower TSH MoM is associated with increased weight loss in women admitted for HG, but FT4 is not. This is in contrast with existing literature that hypothesized that hyperthyroidism and thus increased FT4 leads to weight loss.[14, 23, 24] However, an alternative explanation may be more likely: in healthy pregnant women without HG, lower maternal weight gain is also associated with lower TSH, when taking into account the graded decrease in TSH over the first trimester, suggesting low TSH and not FT4 may simply be a marker of low maternal weight gain, and not a cause. Literature showing decreasing TSH levels in obese non-pregnant patients who lost weight after caloric restriction or bariatric surgery supports this theory.[25-27] Since our study did not track TSH and maternal weight over time before HG symptoms developed, we were unable to test this hypothesis.

We also sought to identify whether TSH or FT4 could be helpful in identifying women with HG that were to have a more severe or prolonged course of illness. Contrary to what had been suggested by the literature, our study found that only a lower TSH at baseline was associated with lower nausea and vomiting scores (PUQE-24<13) one week after inclusion, whereas FT4 was not. An explanation could be that an increased energy intake during this period led to an increase and normalization of TSH levels, since there is evidence that there is a direct relation between TSH and energy intake.[28]

Previous studies did not find an association between TSH or FT4 and readmission rates or the duration of inpatient hospital stay.[29-31] Our study confirms these findings. Unlike the three

currently available studies, our study is the first that performed a multivariable regression analysis and used TSH MoMs in order to correct for gestational fluctuations.

Another reason why clinicians may want to be informed about thyroid function in women with HG, is in order to rule out clinical hyperthyroidism as an alternative explanation for severe NVP symptoms, which may require thyreostatic therapy.[32] One study in particular, which used PUQE-24 scores to quantify NVP severity, focused on ruling out thyroid dysfunction among women with HG. They found no association between hyperthyroidism (3 out of 63 women) and the PUQE-24 score and could therefore not support this alternative explanation for severe nausea and vomiting symptoms. In our study, we also found no association between TSH as well as FT4 and the PUQE-24 score at baseline. Also no differences in PUQE-24 score at baseline between women with and without GTT were found. Therefore, measuring thyroid function to rule out clinical hyperthyroidism as an explanation for severe nausea and vomiting seems unnecessary.

Strengths and limitations

Our study is one of the few prospective cohort studies available including women admitted with HG using multiple measurements to assess HG severity and the clinical course of HG by using validated questionnaires. Furthermore, the women included in this study reflect a geographically representative sample of the Dutch population, because of data collection from 19 different hospitals across the Netherlands. Another strength is that we used TSH MoM to allow us to adjust for pregnancy related fluctuations of TSH related to rise and fall of placental hCG as pregnancy progresses.

Our study has some limitations. First and foremost, we had TSH values available of only two thirds of the included women in our study, since TSH was not always included in local routine laboratory work-up for HG. We also had to exclude women for TSH MoM and FT4 analysis due to lack of TSH medians or due to lack of available frozen stored blood samples. Together with loss to follow up, including a high rate of missing data, despite multiple efforts to retrieve this data as described in the original MOTHER study, this may have limited our power to detect associations. Potentially, there may have been selective loss to follow up, with women with severe symptoms being too unwell to complete follow up questionnaires. However, we found no evidence for selective participation: only the baseline characteristics gestational age and first admission of HG differed between the included and excluded women. Further baseline characteristics and measures of the severity and clinical course of HG did not differ and therefore it is unlikely to have altered our results. Secondly, in this study TSH and FT4 levels were not followed up and therefore we were unable to investigate whether thyroid suppression normalized without

treatment later in pregnancy or after delivery. Previous literature suggested that TSH as well as FT4 levels return to normal by the second trimester in women with HG and GTT.[33]

Conclusion

Based on inconsistent findings from our study as well as from earlier research, there seems little utility of thyroid measurement as marker or predictor for the severity and clinical course of HG. The clinical relevance of measuring TSH in women with HG therefore seems low, making the likelihood that the thyroid plays an important role in the etiology of HG also less likely. As advised in the HG guideline of the Nordic Federation of Societies of Obstetrics and Gynecology[34], thyroid function assessment should be reserved to rule out thyroid disease in women with clinical signs, such as goiter, or women with symptoms of clinical hyperthyroidism, such as marked tachycardia or prolonged atypical symptoms. Further research in pursuit of a biomarker for diagnosis or predictor of disease course, or monitoring treatment effect in patients with HG is needed in order to optimize patient care and treatment.

Acknowledgements

First of all, we thank all participating women in this cohort and all medical staff from the participating hospitals for their efforts. Secondly, we thank Dr. J.P. Bestwick (employed at Queen Mary University of London, London, United Kingdom) and Professor dr. J.H. Lazarus (employed at Cardiff School of Medicine, Cardiff, United Kingdom) for providing us TSH medians from their study in the United Kingdom. Dr. J.P. Bestwick and Professor dr. Lazarus have nothing to disclose.

Funding

This prospective cohort study was supported by a research grant from North West Hospital Group, Alkmaar, the Netherlands (Grant number: 2013T085) and by a research grant from the Amsterdam Reproduction and Development (AR&D) research institute, Amsterdam UMC, the Netherlands (Project number: 23346).

Contribution to authorship

IJG, TJR and RCP conceived and designed the study. KN performed the statistical analysis under supervision of MHK and RCP. All authors critically reviewed the manuscript and approved the final version.

REFERENCES

1. Abramowitz A, Miller ES, Wisner KL. Treatment options for hyperemesis gravidarum. Archives of women's mental health. 2017.
2. Matsuo K, Ushioda N, Nagamatsu M, Kimura T. Hyperemesis gravidarum in Eastern Asian population. Gynecologic and obstetric investigation. 2007;64(4):213-6.
3. Boelig RC, Barton SJ, Saccone G, Kelly AJ, Edwards SJ, Berghella V. Interventions for treating hyperemesis gravidarum. The Cochrane database of systematic reviews. 2016(5):Cd010607.
4. Gazmararian JA, Petersen R, Jamieson DJ, Schild L, Adams MM, Deshpande AD, et al. Hospitalizations during pregnancy among managed care enrollees. Obstetrics and gynecology. 2002;100(1):94-100.
5. Poursharif B, Korst LM, Fejzo MS, MacGibbon KW, Romero R, Goodwin TM. The psychosocial burden of hyperemesis gravidarum. Journal of perinatology : official journal of the California Perinatal Association. 2008;28(3):176-81.
6. Wood H, McKellar LV, Lightbody M. Nausea and vomiting in pregnancy: blooming or bloomin' awful? A review of the literature. Women and birth : journal of the Australian College of Midwives. 2013;26(2):100-4.
7. Lacasse A, Rey E, Ferreira E, Morin C, Berard A. Nausea and vomiting of pregnancy: what about quality of life? BJOG : an international journal of obstetrics and gynaecology. 2008;115(12):1484-93.
8. Bailit JL. Hyperemesis gravidarium: Epidemiologic findings from a large cohort. American journal of obstetrics and gynecology. 2005;193(3 Pt 1):811-4.
9. Fejzo MS, Arzy D, Tian R, MacGibbon KW, Mullin PM. Evidence GDF15 Plays a Role in Familial and Recurrent Hyperemesis Gravidarum. Geburtshilfe und Frauenheilkunde. 2018;78(9):866-70.
10. Fejzo MS, Sazonova OV, Sathirapongsasuti JF, Hallgrimsdottir IB, Vacic V, MacGibbon KW, et al. Placenta and appetite genes GDF15 and IGFBP7 are associated with hyperemesis gravidarum. Nature communications. 2018;9(1):1178.
11. Niebyl JR. Clinical practice. Nausea and vomiting in pregnancy. The New England journal of medicine. 2010;363(16):1544-50.
12. Niemeijer MN, Grooten IJ, Vos N, Bais JM, van der Post JA, Mol BW, et al. Diagnostic markers for hyperemesis gravidarum: a systematic review and metaanalysis. American journal of obstetrics and gynecology. 2014;211(2):150.e1-15.
13. Yoshimura M, Hershman JM. Thyrotropic action of human chorionic gonadotropin. Thyroid : official journal of the American Thyroid Association. 1995;5(5):425-34.
14. Melish JS. Thyroid Disease. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. Boston: Butterworths. Butterworth Publishers, a division of Reed Publishing.; 1990.
15. Grooten IJ, Koot MH, van der Post JA, Bais JM, Ris-Stalpers C, Naaktgeboren C, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. The American journal of clinical nutrition. 2017;106(3):812-20.
16. Grooten IJ, Mol BW, van der Post JA, Ris-Stalpers C, Kok M, Bais JM, et al. Early nasogastric tube feeding in optimising treatment for hyperemesis gravidarum: the MOTHER randomised controlled trial (Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding). BMC pregnancy and childbirth. 2016;16:22.

17. Bestwick JP, John R, Maina A, Guaraldo V, Joomun M, Wald NJ, et al. Thyroid stimulating hormone and free thyroxine in pregnancy: expressing concentrations as multiples of the median (MoMs). Clinica chimica acta; international journal of clinical chemistry. 2014;430:33-7.

18. Roche. Elecsys Systems 1010/2010. Reference interval for children and adults (TSH, FT4, FT3, T-Uptake, FT-4 index, Anti-TPO, anti-TG, Tg). 2009.

19. Koren G, Piwko C, Ahn E, Boskovic R, Maltepe C, Einarson A, et al. Validation studies of the Pregnancy Unique-Quantification of Emesis (PUQE) scores. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2005;25(3):241-4.

20. Power Z, Campbell M, Kilcoyne P, Kitchener H, Waterman H. The Hyperemesis Impact of Symptoms Questionnaire: development and validation of a clinical tool. International journal of nursing studies. 2010;47(1):67-77.

21. Lacasse A, Berard A. Validation of the nausea and vomiting of pregnancy specific health related quality of life questionnaire. Health and quality of life outcomes. 2008;6:32.

22. Fejzo MS, Poursharif B, Korst LM, Munch S, MacGibbon KW, Romero R, et al. Symptoms and pregnancy outcomes associated with extreme weight loss among women with hyperemesis gravidarum. Journal of women's health (2002). 2009;18(12):1981-7.

23. Asakura H, Watanabe S, Sekiguchi A, Power GG, Araki T. Severity of hyperemesis gravidarum correlates with serum levels of reverse T3. Archives of gynecology and obstetrics. 2000;264(2):57-62.

24. Tsuruta E, Tada H, Tamaki H, Kashiwai T, Asahi K, Takeoka K, et al. Pathogenic role of asialo human chorionic gonadotropin in gestational thyrotoxicosis. The Journal of clinical endocrinology and metabolism. 1995;80(2):350-5.

25. Janssen IM, Homan J, Schijns W, Betzel B, Aarts EO, Berends FJ, et al. Subclinical hypothyroidism and its relation to obesity in patients before and after Roux-en-Y gastric bypass. Surgery for obesity and related diseases : official journal of the American Society for Bariatric Surgery. 2015;11(6):1257-63.

26. Lips MA, Pijl H, van Klinken JB, de Groot GH, Janssen IM, Van Ramshorst B, et al. Roux-en-Y gastric bypass and calorie restriction induce comparable time-dependent effects on thyroid hormone function tests in obese female subjects. European journal of endocrinology. 2013;169(3):339-47.

27. Abu-Ghanem Y, Inbar R, Tyomkin V, Kent I, Berkovich L, Ghinea R, et al. Effect of sleeve gastrectomy on thyroid hormone levels. Obesity surgery. 2015;25(3):452-6.

28. Matzen LE, Kvetny J. The influence of caloric deprivation and food composition on TSH, thyroid hormones and nuclear binding of T3 in mononuclear blood cells in obese women. Metabolism: clinical and experimental. 1989;38(6):555-61.

29. Tan PC, Jacob R, Quek KF, Omar SZ. Readmission risk and metabolic, biochemical, haematological and clinical indicators of severity in hyperemesis gravidarum. The Australian & New Zealand journal of obstetrics & gynaecology. 2006;46(5):446-50.

30. Tan PC, Jacob R, Quek KF, Omar SZ. Indicators of prolonged hospital stay in hyperemesis gravidarum. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2006;93(3):246-7.

31. Derbent AU, Yanik FF, Simavli S, Atasoy L, Urun E, Kuscu UE, et al. First trimester maternal serum PAPP-A and free beta-HCG levels in hyperemesis gravidarum. Prenat Diagn. 2011;31(5):450-3.

32. Mestman JH. Hyperthyroidism in pregnancy. Best Practice & Research Clinical Endocrinology & Metabolism. 2004;18(2):267-88.

33. Sun S, Qiu X, Zhou J. Clinical analysis of 65 cases of hyperemesis gravidarum with gestational transient thyrotoxicosis. The journal of obstetrics and gynaecology research. 2014;40(6):1567-72.

34. Vikanes A, Trovik J, Tellum T, Lomsdal S, Stensløkken A, Nesheim B. Emesis & hyperemesis gravidarum 2014 [Available from: http://www.nfog.org/files/guidelines/7%20NGF%20Obst%20hyperemisis%20Vikanes.pdf.

APPENDIX A

Research in context - updated systematic review

Search strategy and study selection

In order to place our research in contest we updated the search of Niemeijer *et al.*[1], who performed a systematic review and meta-analysis on diagnostic markers for HG in 2012. We updated this search according to the method developed by Bramer *et al.*[2] using EndNote. Since we aimed to study only TSH as biomarker for HG, we modified our search in collaboration with an information specialist to optimize our search results as shown below in 'Updated search terms'. We searched Medline and Embase from inception through May 22, 2020 to identify articles that have reported on TSH or FT4 and the clinical course of HG and disease severity. Etiologic, prognostic, predictive and diagnostic studies that reported on TSH or FT4 in women with HG, written in English were included. Case reports were excluded and if one study population was used in two different studies, only the study with the most complete data was included. Two reviewers (KN and MHK) independently assessed whether studies were eligible. In case of disagreement, a third reviewer was consulted (RCP). We evaluated publication bias by using a funnel plot. We evaluated the quality of included studies by using the QUality In Prognosis Studies (QUIPS) tool .[3] The QUIPS tool assesses the potential risk of bias of each article on six domains: participation, attrition, prognostic factor measurement, confounding measurement, outcome measurement and analysis and reporting).[3]

Included articles and quality assessment

The search yielded 5260 unique articles (3442 articles from Niemeijer *et al.* and 1818 additional articles in our updated search) as shown in **Figure S1.** We included 15 articles (n= 3508 participants) reporting on the association between TSH or FT4 and disease severity (12 studies) and the clinical course of HG (3 studies) **(Table S1)**. In almost all domains, studies were rated as low or moderate risk of bias **(Figure S2)**. Twelve studies however, had a high risk of bias due to confounding, caused by unclear reporting of confounders or by lack of including confounders in statistical analysis. We found evidence of publication bias based on an asymmetric funnel plot **(Figure S3)**.

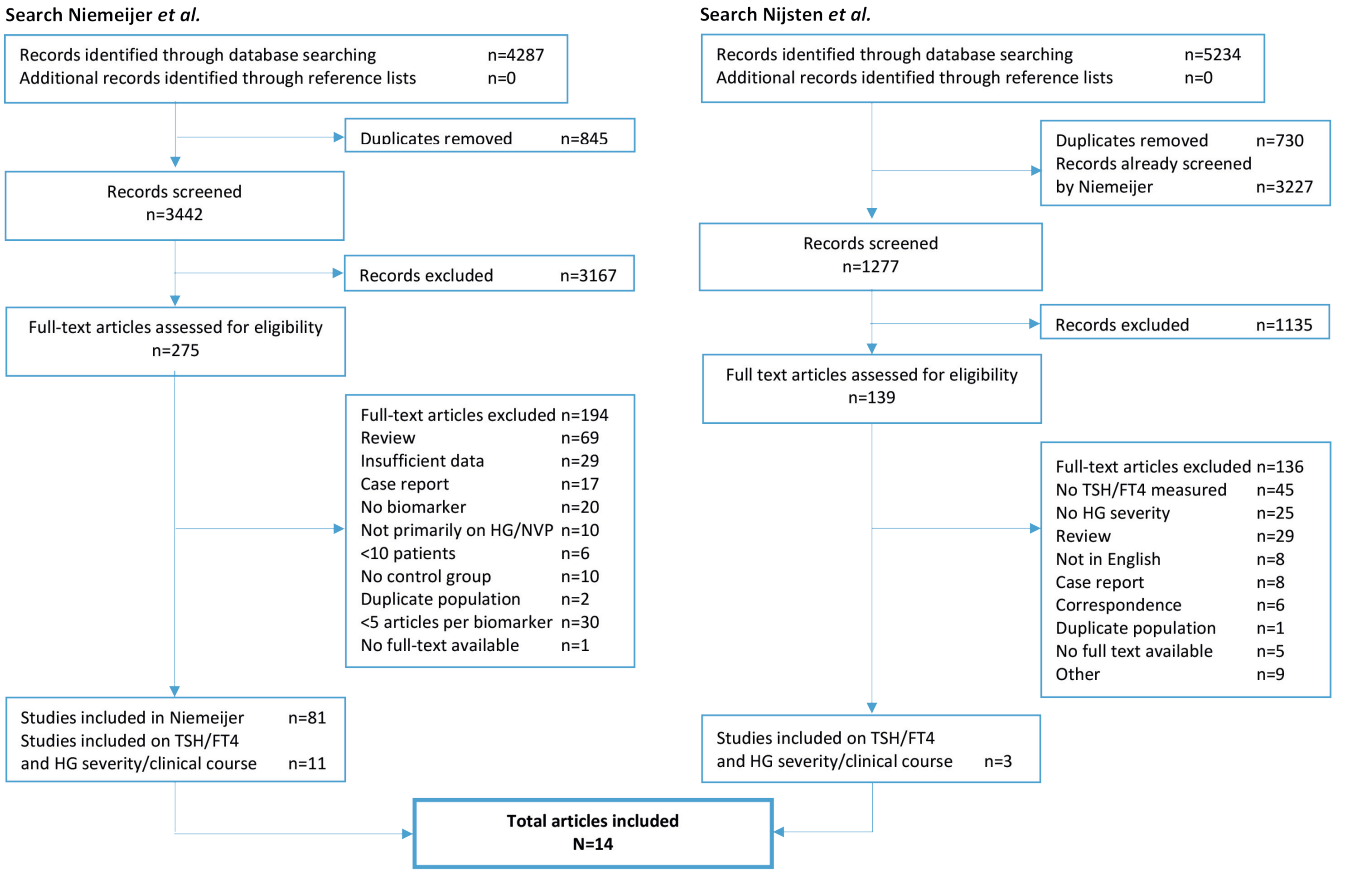


Figure S1. Flowchart screening systematic review

Supplementary Table S1. Studies included in updated systematic review

		Severity and clinical course of HG				Number of participants		Thyroid function	
Study	Year	Severity of vomiting	Weight loss	Readmission	Duration hospital admission	Cases	Control subjects	TSH measured	FT4 measured
Alizza et al	2018	X				30	30	X	X
Asakura et al	2000		X			24	20		X
Chihara et al	2003		X			17	37	X	
Dekkers et al	2020	X				318	1364	X	X
Derbent et al	2011				X	115	110	X	X
Evans et al	1986	X				342	0		X
Goodwin et al	1992	X				57	57	X	X
Güngören et al	2013	X				90	50	X	X
Malek et al	2017	X				63	0	X	X
Mori et al	1988	X				111	41	X	X
Murata et al	2006		X			44	53	X	
Ndungu et al	2009	X				72	0	X	X
Tan et al	2006				X	192	0	X	X
Tan et al	2006			X		192	0	X	X
Tsuruta et al	1995	X				55	24	X	X

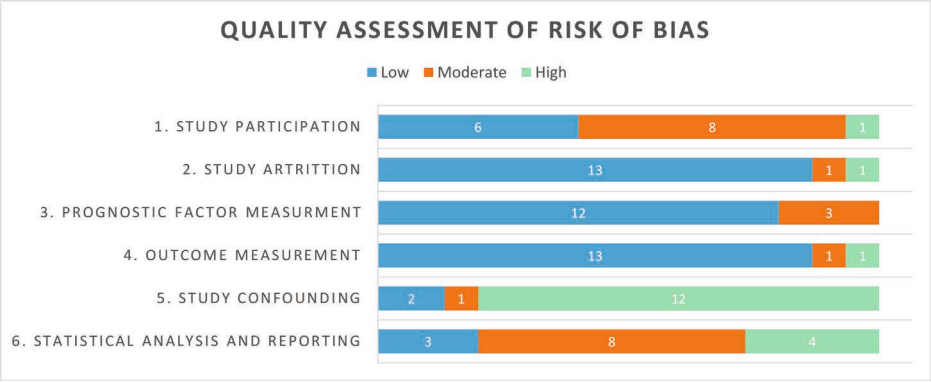


Figure S2. Quality assessment risk of bias (QUIPS)

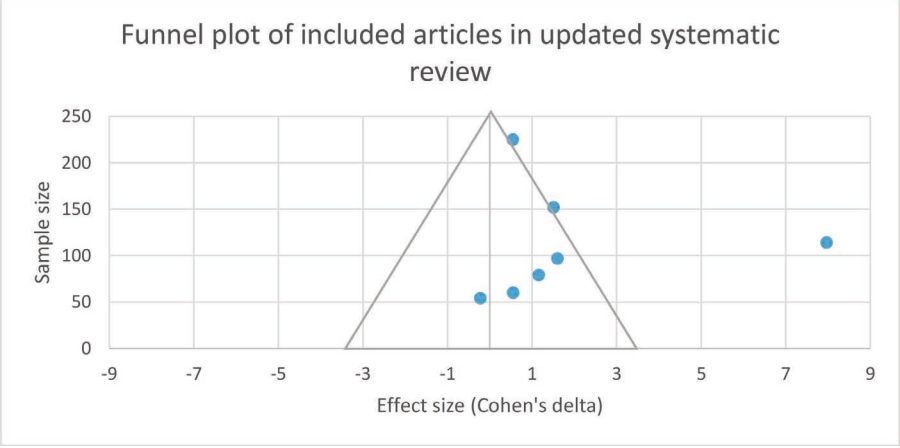


Figure S3. Funnel plot systematic review

TSH and FT4 and the severity and clinical course of HG

9 studies (n=2899 participants) reported on the association between TSH and/or FT4 and the severity of vomiting[4-12], of which 5 studies (n=2087 participants) found an association.[4-8] Two studies used the validated PUQE-24 score (n=1745 participants). The first study (n=63 participants) found that hyperthyroidism (defined as TSH<0.1mIU/L and FT4>26pmol/L) was only present in 3 patients, all with PUQE-24 score > 6, representing moderate to severe vomiting.[12] This study did not find an association between the PUQE-24 score and hyperthyroidism

and the association between TSH and PUQE-24 score was not reported.[12] The second study (n=1682 participants) found that PUQE scores were inversely related to TSH levels, but not with FT4 levels. PUQE scores did not differ significantly between euthyroid women, women with (sub)clinical hypo- or hyperthyroidism. Also, no association was found between (sub)clinical hyperthyroidism (TSH<0.21 mU/L, based on lower 2.5% CI of study population) and more severe NVP (defined as PUQE-24>9) in multivariable logistic regression.[8] Two studies (n= 224 participants) reported an association between lower TSH/ higher FT4 and increased weight loss across women with HG, NVP and no emesis.[13, 14] The third article (n=35 women) found that women with HG with increased FT4 had more weight loss (4.4 ± 0.4 kg) than women with HG without increased FT4 (0.4 ± 1.0 kg).[15] Three studies reporting on the association between TSH or FT4 and the clinical course of HG found no association between TSH or FT4 and the duration of admission (n=417 participants)[16, 17] or the readmission risk (n=192 participants)[18] in women with HG.

Conclusion based on currently available research

The little available evidence concerning the possible association between TSH and FT4 and the severity and clinical course of HG is conflicting. Lower TSH and higher FT4 were associated with increased gestational weight loss or less weight gain regardless of whether they had no, mild, or severe NVP. Lower TSH and higher FT4 were not associated with admission duration or readmissions in women with HG. None of the available studies took the variation in TSH levels according to duration of pregnancy and hCG concentrations into account.

Updated search terms

A. Embase – Updated Search Nijsten May 2020

1. exp etiology/ or exp marker/ or etiol*.ab,ti. or causal*.ab,ti. or cause*.ab,ti. or biomarker*.ab,ti. or marker*.ab,ti. or exp blood analysis/ or analysis.ab,ti. or exp thyrotropin/ or TSH*.ab,ti. or 'thyroid stimulating hormone'.ab,ti. Or exp thyroxine/ or thyroxine.ab,ti. or FT4.ab,ti.

2. exp hyperemesis gravidarum/ or 'hyperemesis gravidarum'.ti,ab.

3. ((pregnan* or gravidar* or gravidit* or gestat* or antenat* or prenat* or ante-nat* or pre-nat*) adj6 (nausea* or antinause* or vomit* or antivomit* or emes* or hypereme* or antiemet* or emetic*)).tw,kw.

4. 2 or 3

5. 1 and 4

B. Medline – Updated Search Nijsten May 2020

1. (morning sickness[MeSH Terms]) OR hyperemesis gravid*[Title/Abstract]

2. (pregnan*[Title/Abstract] OR gestation*[Title/Abstract] OR pregnancy[MeSH Terms]) AND (nausea[Title/Abstract] OR vomit*[Title/Abstract])

3. 1 or 2

4. causality[MeSH Terms] OR etiology[MeSH Subheading] OR etiolog*[Title/Abstract] OR causal*[Title/Abstract] OR cause[Title/Abstract] OR biomarker[Title/Abstract] OR marker[Title/Abstract] OR analysis[MeSH Subheading] OR thyrotropin[MeSH Terms] OR thyroid stimulating hormone[Title/Abstract] OR TSH*[Title/Abstract] OR thyroxine[MeSH Terms] OR thyroxine[Title/Abstract] OR FT4[Title/Abstract]

5. 3 and 4

REFERENCES APPENDIX

1. Niemeijer MN, Grooten IJ, Vos N, Bais JM, van der Post JA, Mol BW, et al. Diagnostic markers for hyperemesis gravidarum: a systematic review and metaanalysis. American journal of obstetrics and gynecology. 2014;211(2):150.e1-15.
2. Bramer W, Bain P. Updating search strategies for systematic reviews using EndNote. Journal of the Medical Library Association : JMLA. 2017;105(3):285-9.
3. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Annals of internal medicine. 2013;158(4):280-6.
4. Goodwin TM, Montoro M, Mestman JH, Pekary AE, Hershman JM. The role of chorionic gonadotropin in transient hyperthyroidism of hyperemesis gravidarum. The Journal of clinical endocrinology and metabolism. 1992;75(5):1333-7.
5. Mori M, Amino N, Tamaki H, Miyai K, Tanizawa O. Morning sickness and thyroid function in normal pregnancy. Obstetrics and gynecology. 1988;72(3 Pt 1):355-9.
6. Tsuruta E, Tada H, Tamaki H, Kashiwai T, Asahi K, Takeoka K, et al. Pathogenic role of asialo human chorionic gonadotropin in gestational thyrotoxicosis. The Journal of clinical endocrinology and metabolism. 1995;80(2):350-5.
7. Alizzi FJ, Abdul Abbas WAJ, Fawzi HA. Assessment of the role of cholecystokinin in hyperemesis gravidarum and correlation with its severity. Journal of Pharmaceutical Sciences and Research. 2018;10(2):272-5.
8. Dekkers GWF, Broeren MAC, Truijens SEM, Kop WJ, Pop VJM. Hormonal and psychological factors in nausea and vomiting during pregnancy. Psychological medicine. 2020;50(2):229-36.
9. Evans AJ, Li TC, Selby C, Jeffcoate WJ. Morning sickness and thyroid function. British journal of obstetrics and gynaecology. 1986;93(5):520-2.
10. Ndungu JR, Amayo A, Qureshi ZP, Kigondu CS. Gestational thyrotoxicosis associated with emesis in early pregnancy. East African medical journal. 2009;86(2):55-8.
11. Gungoren A, Bayramoglu N, Duran N, Kurul M. Association of Helicobacter pylori positivity with the symptoms in patients with hyperemesis gravidarum. Archives of gynecology and obstetrics. 2013;288(6):1279-83.
12. Malek NZH, Kalok A, Hanafiah ZA, Shah SA, Ismail NAM. Association of transient hyperthyroidism and severity of hyperemesis gravidarum. Hormone molecular biology and clinical investigation. 2017;30(3).
13. Asakura H, Watanabe S, Sekiguchi A, Power GG, Araki T. Severity of hyperemesis gravidarum correlates with serum levels of reverse T3. Archives of gynecology and obstetrics. 2000;264(2):57-62.
14. Murata T, Suzuki S, Takeuchi T, Takeshita T. Relation between plasma adenosine and serum TSH levels in women with hyperemesis gravidarum. Archives of gynecology and obstetrics. 2006;273(6):331-6.
15. Chihara H, Otsubo Y, Yoneyama Y, Sawa R, Suzuki S, Power GG, et al. Basal metabolic rate in hyperemesis gravidarum: comparison to normal pregnancy and response to treatment. American journal of obstetrics and gynecology. 2003;188(2):434-8.
16. Derbent AU, Yanik FF, Simavli S, Atasoy L, Urun E, Kuscu UE, et al. First trimester maternal serum PAPP-A and free beta-HCG levels in hyperemesis gravidarum. Prenat Diagn. 2011;31(5):450-3.
17. Tan PC, Jacob R, Quek KF, Omar SZ. Indicators of prolonged hospital stay in hyperemesis gravidarum. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2006;93(3):246-7.
18. Tan PC, Jacob R, Quek KF, Omar SZ. Readmission risk and metabolic, biochemical, haematological and clinical indicators of severity in hyperemesis gravidarum. The Australian & New Zealand journal of obstetrics & gynaecology. 2006;46(5):446-50.

Supplementary Table S2. Differences in baseline characteristics and outcome variables between women with HG with and without **absolute TSH** available

	Included for **absolute TSH** analysis	Missing (%)	Excluded for **absolute TSH** analysis	Missing (%)	***P***
n	150		65		
Demographics					
Age (years)	28.95±4.87	0.0%	28.55±4.76	0.0%	0.58
Prepregnancy weight (kg)	70.50±15.13	2.7%	72.51±14.80	1.5%	0.37
Prepregnancy BMI (kg/m^2)	25.00±5.06	4.0%	25.39±4.51	3.1%	0.60
Ethnic origin		18.0%		18.5%	0.47
- Western	88 (58.7%)		35 (53.8%)		
- Non-western	35 (23.3%)		18 (27.7%)		
Education level		35.3%		30.8%	0.33
- Primary or secondary	55 (36.7%)		31 (47.7%)		
- Higher	42 (28.0%)		14 (21.5%)		
Mental health disorder in medical history[1]	30 (20.0%)	0.0%	11 (16.9%)	0.0%	0.60
HG in previous pregnancy[2]	42 (42.0%)	17%	26 (51.0%)	11.8%	0.44
HG in previous pregnancy requiring hospital admission[2]	22 (22.0%)	4.8%	15 (29.4%)	19.2%	0.21
Pregnancy characteristics					
Primigravida	50 (33.3%)	0.0%	14 (21.5%)	0.0%	0.08
Twin pregnancy	4 (2.7%)	0.0%	1 (1.5%)	0.0%	1.00
Gestational age of onset of HG symptoms (weeks)	6.00 (5.00-7.00)	24.6%	6.00 (5.50-7.00)	22.7%	0.59
Gestational age at baseline	8.50 (7.00-11.00)	0.0%	9.00 (7.00-11.00)	0.0%	0.52
First admission at study entry	139 (92.7%)	0.0%	52 (80.0%)	0.0%	0.01
Outcomes					
HG severity at baseline					
-Weight change (kg)	-2.68±3.83	3.3%	-3.47±4.55	1.5%	0.19
-PUQE-24	9.92±3.29	34.7%	10.24±3.35	43.1%	0.61
-NVPQoL	173.88±20.89	32.0%	172.26±29.49	41.5%	0.72
-HIS	27.85±3.62	31.3%	27.53±4.50	41.5%	0.66
Clinical course of HG					
-PUQE-24 one week after inclusion	9.00 (6.00-11.00)	44.7%	8.50 (5.75-11.00)	47.7%	0.31
-NVPQoL one week after inclusion	78.50 (62.00-99.50)	49.3%	72.00 (58.50-129.50)	49.2%	0.99
-HIS one week after inclusion	25.51±3.54	49.3%	16.15±4.44	49.2%	0.43
-Duration first admission (days)	4.00 (3.00-5.00)	0.0%	4.00 (3.00-6.00)	0.0%	0.90
-Total days of hospital admission for HG	5.00 (4.00-8.00)	0.0%	4.00 (3.00-8.00)	0.0%	0.38
-Readmitted	53 (35.3%)	0.0%	18 (27.7%)	0.0%	0.27
-Readmitted ≥ two times	21 (14.0%)	0.0%	8 (12.3%)	0.0%	0.74

A *P*-value<0.05 is considered significant. Data represented with mean±SD and median (IQR), unless st1ated otherwise (frequency (%)). [1] Mental health disorder consists of an eating disorder, anxiety disorder or a depressive disorder. [2] Percentage shown is frequency divided by number of multigravidas. BMI: body mass index. HG: Hyperemesis Gravidarum. TSH: Thyroid Stimulating Hormone. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HIS: Hyperemesis Impact of Symptoms. NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life. A higher PUQE-24, HIS- or NVPQoL-score indicates more severe symptoms or lower quality of life.

Supplementary Table S3. Differences in baseline characteristics and outcome variables between women with HG with and without **TSH MoM** available

	Included for **TSH MoM** analysis	Missing (%)	Excluded for **TSH MoM** analysis	Missing (%)	***P***
n	126		89		
Demographics					
Age (years)	28.75±4.49	0.0%	28.96±5.30	0.0%	0.76
Prepregnancy weight (kg)	71.48±15.49	3.2%	70.61±14.42	1.1%	0.68
Prepregnancy BMI (kg/m^2)	25.15±5.30	4.0%	25.07±4.27	3.4%	0.91
Ethnic origin		16.7%		20.2%	0.62
- Western	75 (59.5%)		48 (53.9%)		
- Non-western	30 (23.8%)		23 (25.8%)		
Education level		34.1%		33.7%	0.28
- Primary or secondary	46 (36.5%)		40 (44.9%)		
- Higher	37 (29.4%)		19 (21.3%)		
Mental health disorder in medical history[1]	24 (19.0%)	0.0%	17 (19.1%)	0.0%	1.00
HG in previous pregnancy[2]	34 (41.0%)	20.5%	34 (50%)	8.8%	0.73
HG in previous pregnancy requiring hospital admission[2]	18 (21.7%)	0.0%	19 (27.9%)	20.6%	0.20
Pregnancy characteristics					
Primigravida	43 (34.1%)	0.0%	21 (23.6%)	0.0%	0.13
Twin pregnancy	3 (2.4%)	0.0%	2 (2.2%)	0.0%	1.00
Gestational age of onset of HG symptoms (weeks)	6.00 (5.00-7.00)	21.4%	6.00 (5.50-7.00)	25.8%	0.63
Gestational age at baseline	9.00 (8.00-11.00)	0.0%	8.00 (6.00-10.00)	0.0%	<0.01
First admission at study entry	117 (92.9%)	0.0%	74 (83.1%)	0.0%	0.03
Outcomes					
HG severity at baseline					
-Weight change (kg)	-2.86±3.64	4.0%	-3.00±4.61	1.1%	0.80
-PUQE-24	9.98±3.19	33.3%	10.06±3.51	42.7%	0.89
-NVPQoL	175.20±18.67	33.3%	170.80±29.16	37.1%	0.28
-HIS	28.50±3.61	30.2%	27.32±4.24	40.4%	0.29
Clinical course of HG					
-PUQE-24 one week after inclusion	8.50 (6.00-11.25)	44.4%	9.00 (6.00-11.00)	47.2%	0.48
-NVPQoL one week after inclusion	76.00 (61.25-95.00)	47.6%	80.00 (60.00-117.00)	51.7%	0.62
-HIS one week after inclusion	25.53±3.66	47.6%	25.98±4.09	51.7%	0.55
-Duration first admission (days)	4.00 (3.00-5.00)	0.0%	4.00 (3.00-5.00)	0.0%	0.98
-Total days of hospital admission for HG	5.00 (4.00-7.00)	0.0%	5.00 (3.00-8.50)	0.0%	0.80
-Readmitted	41 (32.5%)	0.0%	30 (33.7%)	0.0%	0.86
-Readmitted ≥ two times	17 (13.5%)	0.0%	12 (13.5%)	0.0%	1.00

A *P*-value<0.05 is considered significant. Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). [1] Mental health disorder consists of an eating disorder, anxiety disorder or a depressive disorder. [2] Percentage shown is frequency divided by number of multigravidas. BMI: body mass index. HG: Hyperemesis Gravidarum. TSH: Thyroid Stimulating Hormone. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HIS: Hyperemesis Impact of Symptoms. MoM: Multiple of the Median. NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life. A higher PUQE-24, HIS- or NVPQoL-score indicates more severe symptoms or lower quality of life.

Supplementary Table S4. Differences in baseline characteristics and outcome variables between women with HG with and without **FT4** available

	Included for **FT4** analysis	Missing (%)	Excluded for **FT4** analysis	Missing (%)	***P***
n	106		109		
Demographics					
Age (years)	28.69±4.87	0.0%	28.97±4.81	0.0%	0.67
Prepregnancy weight (kg)	71.24±15.34	2.8%	70.99±14.78	1.8%	0.90
Prepregnancy BMI (kg/m^2)	25.46±5.19	2.8%	24.78±4.57	4.6%	0.32
Ethnic origin		15.1%		21.1%	0.20
- Western	59 (55.7%)		64 (58.7%)		
- Non-western	31 (29.2%)		22 (20.2%)		
Education level		27.4%		40.4%	0.54
- Primary or secondary	44 (41.5%)		42 (38.5%)		
- Higher	33 (31.1%)		23 (21.1%)		
Mental health disorder in medical history[1]	19 (17.9%)	0.0%	22 (20.2%)	0.0%	0.67
HG in previous pregnancy[2]	28 (40.0%)	17.1%	40 (49.4%)	13.6%	0.32
HG in previous pregnancy requiring hospital admission[2]	14 (20.0%)	0.0%	23 (28.4)	17.5%	0.12
Pregnancy characteristics					
Primigravida	36 (34.0%)	0.0%	28 (25.7%)	0.0%	0.19
Twin pregnancy	3 (2.8%)	0.0%	2 (1.8%)	0.0%	0.68
Gestational age of onset of HG symptoms (weeks)	6.00 (5.00-7.00)	25.7%	6.00 (5.75-7.00)	20.8%	0.13
Gestational age at baseline	8.00 (7.00-10.00)	0.0%	9.00 (7.00-12.00)	0.0%	0.26
First admission at study entry	97 (91.5%)	0.0%	94 (86.2%)	0.0%	0.33
Outcomes					
HG severity at baseline					
-Weight change (kg)	-3.14±3.91	2.8%	-2.71±4.23	2.8%	0.45
-PUQE-24	10.28±3.15	32.1%	9.70±3.45	42.2%	0.31
-NVPQoL	173.71±22.11	28.3%	173.13±25.09	41.3%	0.88
-HIS	27.87±3.79	26.4%	27.63±3.97	42.2%	0.72
Clinical course of HG					
-PUQE-24 one week after inclusion	9.00 (6.00-11.25)	41.5%	8.00 (6.00-11.00)	49.5%	0.34
-NVPQoL one week after inclusion	79.00 (59.00-98.00)	47.2%	73.00 (62.00-111.00)	51.4%	0.68
-HIS one week after inclusion	25.46±3.53	47.2%	25.96±4.13	51.4%	0.50
-Duration first admission (days)	4.00 (3.00-5.00)	0.0%	4.00 (3.00-5.00)	0.0%	0.45
-Total days of hospital admission for HG	5.00 (4.00-8.00)	0.0%	5.00 (3.00-8.00)	0.0%	0.30
-Readmitted	41 (38.7%)	0.0%	30 (27.5%)	0.0%	0.08
-Readmitted ≥ two times	15 (14.2%)	0.0%	14 (12.8%)	0.0%	0.78

A *P*-value<0.05 is considered significant. Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). [1] Mental health disorder consists of an eating disorder, anxiety disorder or a depressive disorder. [2] Percentage shown is frequency divided by number of multigravidas. BMI: body mass index. FT4: free thyroxine. HG: Hyperemesis Gravidarum. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gain weight. HIS: Hyperemesis Impact of Symptoms. NVPQoL: Nausea and Vomiting in Pregnancy Quality of Life. A higher PUQE-24, HIS- or NVPQoL-score indicates more severe symptoms or lower quality of life.

CHAPTER 4

Recurrence, postponing pregnancy, and termination rates after hyperemesis gravidarum: follow up of the MOTHER study

Kelly Nijsten, Caitlin Dean, Loïs M. van der Minnen, Joke M.J. Bais, Carrie Ris-Stalpers, Rik van Eekelen, Henk A. Bremer, David P. van der Ham, Wieteke M. Heidema, Anjoke Huisjes, Gunilla Kleiverda, Simone M. Kuppens, Judith O.E.H. van Laar, Josje Langenveld, Flip van der Made, Dimitri Papatsonis, Marie-José Pelinck, Paula J. Pernet, Leonie van Rheenen-Flach, Robbert J. Rijnders, Hubertina C.J. Scheepers, Tatjana Vogelvang, Prof. Ben W. Mol, Prof. Tessa J. Roseboom, Marjette H. Koot, Iris J. Grooten, Rebecca C. Painter

Acta Obstetricia et Gynecologica Scandinavica, 2021; 100: 1636-1643.

ABSTRACT

Introduction

Hyperemesis gravidarum (HG) complicates 1% of pregnancies and has a major impact on maternal quality of life and wellbeing. We know very little about HG's long-term impact after an affected pregnancy, including recurrence rates in future pregnancies, which is essential information for women considering subsequent pregnancies. In this study, we aimed to prospectively measure the recurrence rate of HG and the number of postponed and terminated subsequent pregnancies due to HG. We also aimed to evaluate if there were predictive factors that could identify women at increased risk for HG recurrence, postponing and terminating subsequent pregnancies.

Material and methods

We conducted a prospective cohort study. A total of 215 women admitted for HG to public hospitals in the Netherlands were enrolled in the original MOTHER randomized controlled trial and associated observational cohort. Seventy-three women were included in this follow-up study. Data were collected via an online questionnaire. Recurrent HG was defined as vomiting symptoms accompanied by any of the following: multiple medication use, weight loss, admission, tube feeding or if nausea and vomiting symptoms were severe enough to affect life and/or work. Outcome measures were recurrence, postponing and termination rates due to HG. Univariable logistic regression analysis was used to identify predictive factors associated with HG recurrence, postponing and terminating subsequent pregnancies.

Results

Thirty-five women (48%) became pregnant again of whom 40% had postponed their pregnancy due to HG. HG recurred in 89% of pregnancies. One woman terminated and eight women (23%) considered terminating their pregnancy because of recurrent HG. Twenty-four out of 38 women did not get pregnant again because of HG in the past. Univariable logistic regression analysis identifying possible predictive factors found that having a western background was associated with having weight loss due to recurrent HG in subsequent pregnancies (OR 12.9, 95% CI: 1.3-130.5, P=0.03).

Conclusions

High rates of HG recurrence and a high number of postponed pregnancies due to HG were observed. Women can be informed of a high chance of recurrence to enable informed family planning.

INTRODUCTION

Hyperemesis gravidarum (HG) is a severe complication of pregnancy affecting around 1% of pregnancies globally.[1] HG can cause significant physical and psychological morbidity for mothers.[2] A lack of effective treatment makes HG a challenging condition to manage and therapeutic termination is commonly reported.[3, 4]

A history of HG is the single most important risk factor for developing HG.[5] The recurrence rate has been reported to be well above baseline risk, but literature shows a wide range varying from 15 to 81%.[5] A recent systematic review failed to produce an aggregate recurrence rate due to the contributing studies' methodological shortcomings, including poor external validity and significant heterogeneity.[6]

Such a wide risk prediction bracket for a condition with substantial biopsychosocial impacts, makes informed decision-making regarding subsequent pregnancies difficult. Patients have expressed a desire for research to provide a definitive recurrence risk and recently this was also recognized as a priority research question by a priority setting partnership.[6, 7] Furthermore, there is evidence suggesting that early treatment and lifestyle preparation strategies may reduce the overall severity of the condition.[8, 9] For such interventions to be appropriately implemented, the recurrence rate must be understood.[6]

Both over- and underestimating the recurrence rate can have substantial impacts on people's lives. There are reports of families curtailing future pregnancies believing HG is unavoidable as well as reports of women deciding to terminate on the assumption that their risk in a future pregnancy is that of the general population.[3, 10, 11]

In this study, we aimed to prospectively measure the self-reported recurrence rate of HG, the postponement of pregnancy because of previous HG, and pregnancies terminated due to recurrent HG. Additionally, we aimed to identify predictive factors associated with an increased risk of HG recurrence, postponing and terminating subsequent pregnancies.

METHODS

This study is a prospective cohort follow-up study of the Maternal and Offspring outcomes after Treatment for HyperEmesis by Refeeding (MOTHER) randomized control trial (RCT) and associated observational cohort.[12]

The original MOTHER RCT assessed whether early enteral tube feeding in addition to standard care for women admitted with HG improved neonatal and maternal outcomes.[12] Women admitted for HG, between 5- and 20-weeks' gestation in 19 different hospitals in the Netherlands between 2013 and 2016 were recruited. In total, 115 women were randomized and 100 women, who declined randomization, were recruited to an associated observational cohort. Since early enteral tube feeding did not affect maternal and perinatal outcomes,[12] we combined the RCT and cohort into one study population for this follow-up study. Detailed information about data collection can be found in the original study protocol and earlier published results of the MOTHER RCT.[12, 13]

The MOTHER follow-up study consisted of a single, self-reported, online questionnaire that assessed health and reproductive outcomes after participating in the MOTHER study. Participants who gave consent to be approached for follow-up studies were emailed with a link to the online questionnaire. Both Dutch and English language options were available. In case of no response, a reminder was sent after one, three and six weeks. Individual informed consent had been obtained during both the MOTHER and follow-up study.

For the full questionnaire please see **Appendix S1**. Women self-reported whether they had conceived again since participating in the MOTHER study. Those who had not had a further pregnancy were asked whether they had curtailed or postponed any future pregnancies due to fear of recurrent HG. For those who had subsequent pregnancies, nausea and vomiting symptoms were assessed with a series of questions regarding the onset of symptoms, hospital admission including duration and frequency, anti-emetics use and tube feeding. We considered that HG had recurred if vomiting symptoms were reported with either: multiple HG medication use (≥2, including anti-emetics and corticosteroids, see full list in **Appendix S1**), weight loss during pregnancy, admission for HG, requiring tube feeding or symptoms severe enough to affect their life and/or work. The HG definition we used was based on the recently internationally developed WINDSOR HG definition (unpublished results, manuscript currently submitted for publication). Weight loss was reported as lowest weight during pregnancy compared to pre-pregnancy weight and reported as any weight loss and >5% weight loss.

We also assessed whether pregnancies had ended as miscarriages or ectopic pregnancies and if women had considered terminating or terminated their pregnancy due to recurrent HG. Because of ethical considerations, we were unable to verify answers to the questionnaire with medical records. The follow-up questionnaire also included questions about depression, anxiety and post-traumatic stress disorder symptoms after suffering from HG in the index pregnancy. These results will be discussed in a different manuscript that is currently submitted for publication.

We also assessed if we could identify factors that could predict which women were at increased risk to have recurrent HG, to postpone subsequent pregnancies or to terminate or consider terminating subsequent pregnancies because of severe recurrent HG with use of univariable regression analysis. For this analysis, recurrent HG in subsequent pregnancies was broken down into the following outcome measures: being admitted to the hospital, having weight loss and receiving tube feeding due to recurrent HG in subsequent pregnancies.

We assessed the following possible predictive factors: maternal age, ethnicity and several measures of HG severity in the index pregnancy, when participating in the MOTHER study. Measures of HG severity in the index pregnancy as predictor variables were: higher symptom severity (measured by the self-reported, validated 24-hour Pregnancy Unique Quantification of Emesis (PUQE) score at baseline),[14, 15] lower weight gain at inclusion of the MOTHER study compared to pre-pregnancy weight, higher total duration of hospital admissions and admission after the first trimester. The PUQE-24 score can vary from 3 to 15 with a higher score indicating more severe symptoms.

Statistical analyses

Continuous variables were presented as means with standard deviations (SDs) if they were normally distributed, or otherwise presented as medians with interquartile ranges (IQRs). Dichotomous and categorical variables were presented as frequencies with percentages. A sensitivity analysis was performed to assess differences in demographic and clinical characteristics of the index pregnancy between participating and non-participating women in this follow-up study. Independent Student's *t* test, Mann-Whitney U test and chi-square test were used for analyses.

Univariable logistic regression analysis was used to identify possible predictive factors for an increased risk of HG recurrence, postponing or terminating subsequent pregnancies. Due to the

low number of events, we were not able to perform multivariable logistic regression analysis and adjust for confounders.[16]

As described earlier, we deemed it appropriate to combine the MOTHER RCT and associated observational cohort into one combined study population. However, since this study is a follow-up of an RCT, we felt it was necessary, for ethical reasons, to also assess whether there were differences in recurrence, postponement and termination rates between the RCT arms. Methods and results of these analyses can be found in **Appendix S2**. *P*-values <0.05 were considered statistically significant and SPSS Statistics 26.0 for Windows (IBM Corp., Armonk, NY, USA) was used for analyses.

Patient and Public Involvement (PPI)

Patients have been involved in this research from the inception of the MOTHER study when patients expressed a desire for the research question. The Dutch HG patient charity, Zwangerschapsmisselijkheid en Hyperemesis Gravidarum (ZEHG), was consulted at various points, including piloting the survey questions. Desire for a prospective study to address the recurrence rate of HG is well documented by one of the authors, who is a patient representative (CD),[6] and who has given patient perspective on the results and interpretation of this study.

Ethical approval

The MOTHER trial was registered at www.trialregister.nl (NTR4197) and was approved by the research ethics committee of the Amsterdam UMC, location AMC on the third of April 2013. Ethics approval was not required for the follow-up study under the Medical Research Involving Human Subjects Act (reference number W20_066 #20.094).

RESULTS

One hundred and ninety out of 215 MOTHER participants, who had given consent to be contacted for follow-up studies, were approached. Seventy-five participants completed the follow-up survey between March and May 2020. About half of the respondents completed the questionnaire after receiving the initial email invitation (40/75, 53%). Respectively 11% (8/75), 20% (15/75) and 16% (12/75) of the participants responded after the first, second and third reminder email. Two women were excluded because they reported on pregnancies prior to the index pregnancy and not on subsequent ones in a distinguishable way. Therefore 73 participants were included for analysis as shown in **Figure 1**.

Table 1 details baseline characteristics of women included in our study. **Supporting information Table S1** shows a sensitivity analysis between follow-up participants and those who did not participated in the follow-up study. Participants were more highly educated (P <0.01) and had had higher vomiting scores at inclusion during the index pregnancy (P=0.02), than those who did not participated.

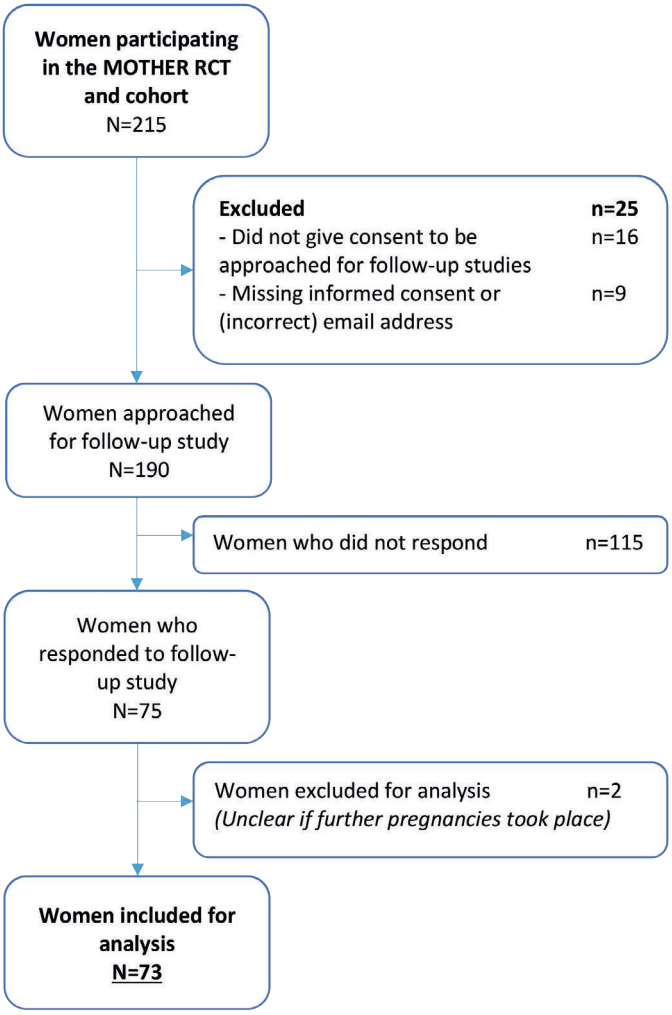


Figure 1. Flow of participants included and excluded for follow-up study

Table 1. Baseline characteristics of women included in this follow-up study.

Baseline characteristics	N=73	
Age (years), mean ± SD	29.2	± 4.6
Education level, n (%)		
- Primary or secondary	27	37.0%
- Higher	29	39.7%
Ethnicity, n (%)		
- Western	52	71.2%
- Non-Western	12	16.4%
Primigravida at time of MOTHER inclusion, n (%)	27	37.0%
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	22	47.8%

Data presented with mean ± SD, median (IQR) or frequency (%).
[a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at inclusion minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. **Abbreviations**: HG: hyperemesis gravidarum, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms.

Following the index pregnancy, 38 women (52%) did not get pregnant again. Of them, two-thirds (24/38) stated that this was because of HG, while 14 women stated other reasons (**Table 2**). Thirty-five women conceived one or more subsequent pregnancies. Of those women, 40% (14/35) had postponed their pregnancy due to HG.

HG recurred in 88.6% of subsequent pregnancies. **(Table 2)** Of the women with recurrent HG, 54% required two or more medications to manage symptoms, 60% were admitted to hospital for HG and 63% experienced weight loss with an average of -4.5 ± 4.4 kg. One woman terminated and eight women (23%) considered terminating their subsequent pregnancy because of recurrent HG.

In six out of 45 subsequent pregnancies no vomiting symptoms were reported. Of these six pregnancies, one was an ectopic pregnancy and four ended in a miscarriage. Four of these five women went on to have another, successful pregnancy in which they did experience vomiting symptoms. The fifth woman had three miscarriages in which she suffered from HG each time.

In univariable regression analysis, we assessed if there were factors that could predict HG recurrence, postponing and (consideration of) terminating subsequent pregnancies. Baseline characteristics of women that became pregnant again and were included in regression analysis are shown in **Supporting information Table S2.** Univariable logistic regression analysis showed that having a western background was associated with having weight loss due to

recurrent HG in subsequent pregnancies (OR 12.86, 95% CI: 1.27-130.54, *P*=0.03). No associations were found between maternal age and several measures of HG severity in the index pregnancy, and HG recurrence, postponing and (consideration of) terminating subsequent pregnancies. **(Table 3)**

Table 2. Recurrence, postponing and termination rates in subsequent pregnancies

Subsequent pregnancies	N=73	
Women who experienced a subsequent pregnancy, n (%)	35	47.9%
Number of pregnancies after MOTHER study, median (IQR)	1.0	1.0-2.0
1 pregnancy, n (%)	26	74.3%
2 pregnancies, n (%)	8	22.9%
3 or more pregnancies, n (%)	1	2.9%
Women who did not become pregnant again due to fear for recurrent HG, n (%)	24	32.9%
Time interval between MOTHER and follow-up study participation (years), median (IQR)	4.5	4.1-5.0
Recurrence rate of HG [a]	N=35	
Recurrence of vomiting symptoms, n (%)	34	97.1%
Gestational age when vomiting started (weeks), median (IQR)	6.0	4.0-7.0
Used multiple (2 or more) HG related medications, n (%)	19	54.3%
Was admitted to hospital, n (%)	21	60.0%
Duration of hospital admissions (days), mean ± SD	6.5	± 4.0
Had weight loss, n (%) [b]	22	62.9%
Had >5% weight loss, n (%)	16	45.7%
Average weight change (kg), mean ± SD [b]	-4.2	± 4.3
Received tube feeding, n (%)	5	14.3%
NVP affected her job, n (%)	20	57.1%
NVP affected her life, n (%)	26	74.3%
Had HG, n (%) [c]	31	88.6%
Postponed or terminated pregnancies		
Postponed a pregnancy due to HG in the past, n (%)	14	40.0%
Considered terminating a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	8	22.9%
Terminated a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	1	2.9%

Data presented with mean ± SD, median (IQR) or frequency (%).
[a] HG recurrence rate in any subsequent pregnancy. [b] Lowest weight during pregnancy minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. [c] HG defined as: vomiting symptoms with either multiple medication use, hospital admission, weight loss during pregnancy, tube feeding or NVP affecting her job and/or life.
Abbreviations: HG: hyperemesis gravidarum, NVP: nausea and vomiting in pregnancy.

Table 3. Univariable logistic regression analysis to assess possible predictive factors, including measures of HG severity of the index pregnancy, for HG recurrence, postponing and terminating subsequent pregnancies

	Was admitted for HG in any subsequent pregnancy		**Had weight loss in any subsequent pregnancy [b]**		**Had tube feeding in any subsequent pregnancy**		**Postponed subsequent pregnancies due to HG in the past**		**Terminated or considered terminating a subsequent pregnancy**	
Possible predictive factors:	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
- Maternal age at inclusion MOTHER study	1.00	0.85-1.17	0.95	0.81-1.12	1.08	0.87-1.33	1.00	0.85-1.17	0.94	0.77-1.15
- Ethnicity (western or non-western)	1.78	0.30-10.72	**12.86**	**1.27-130.54**	- [a]	- [a]	0.79	0.13-4.68	1.25	0.12-13.24
Measures of HG severity in the index pregnancy										
- Weight change (kg) at inclusion MOTHER study [b]	1.10	0.94-1.29	1.03	0.89-1.20	1.12	0.87-1.44	0.97	0.84-1.12	0.97	0.82-1.14
- PUQE-24 at inclusion MOTHER study	0.88	0.67-1.16	1.29	0.96-1.73	1.21	0.80-1.84	0.99	0.76-1.29	0.96	0.70-1.30
- Total duration of hospital admission(s) in the index pregnancy (days)	1.01	0.85-1.19	1.09	0.90-1.32	1.16	0.94-1.42	1.01	0.85-1.19	1.05	0.87-1.26
- Admission in the second or third trimester in the index pregnancy	8.00	0.87-73.40	1.25	0.25-6.16	0.69	0.07-7.11	0.68	0.14-3.34	2.10	0.39-11.43

A P-value<0.05 is considered significant and marked in bold. OR is the odds ratio. 95% CI is the 95% confidence interval.

[a] All women that received tube feeding in a subsequent pregnancy had a western background. [b] Weight change during index pregnancy is weight at baseline minus prepregnancy weight. Weight loss during subsequent pregnancies is prepregnancy weight minus lowest weight during pregnancy. Both can be < 0 if women lost weight and > 0 if women gained weight. Abbreviations: HG: Hyperemesis gravidarum. PUQE: 24-hour Pregnancy Unique Quantification of Emesis and nausea. A higher PUQE-24 score indicates more severe symptoms.

DISCUSSION

In a well-defined prospective cohort, we found a high HG recurrence rate of 89%. Furthermore, we found high proportions of women who avoided a subsequent pregnancy (33%), postponed their pregnancy (40%), or considered terminating their pregnancy (23%) because of HG. Additionally, we found that having a western background was associated with having weight loss due to recurrent HG in subsequent pregnancies.

Our study found an 89% recurrence rate of HG. A recent systematic review identified five previously published, prospective studies assessing the HG recurrence rate.[6] Four were population database cohorts which used birth registry data and ICD-10 codes to identify HG patients and which reported relatively low recurrence rates between 15% and 26%.[5, 17-19] Whilst large populations were included with the number of HG cases varying from 447 to 33 214, Dean *et al.*[6] concluded that methods lacked both external validity and internal reliability. While ICD codes may seem an effective method for pregnancy data collection, attempts to validate them for identifying HG have proved unsuccessful.[6] Norwegian researchers found that the Medical Birth Registry and ICD codes were valid for mild, but not for severe pregnancy sickness or HG.[20] In our study, only 60% of women suffering from HG were admitted in subsequent pregnancies, which would suggest that ICD codes are missing for around 30% of women with recurrent HG in the Dutch system. The fifth study, by Fejzo *et al.*,[21] reported a substantially higher recurrence rate of 81% (46/57 women), but consisted of a self-reported follow-up from an online survey of self-selected participants, making it prone to selection bias. While our follow-up survey was also self-reported, the initial population was recruited with robust inclusion criteria for HG, which provided our study with a greater degree of external validity.

To our knowledge, Fejzo *et al.*[21] is the only study assessing HG severity in subsequent pregnancies. They reported higher rates of tube feeding than our study (20% vs 14.3%), but similar admission rates (48% vs 60%), which is likely a reflection of healthcare system differences between the United States of America (USA) and European countries.[22]

Literature regarding women curtailing pregnancies after suffering from HG is scarce and heterogenic. Fiaschi *et al.*[5] found no evidence of HG sufferers curtailing any future pregnancies compared to non-HG sufferers in their population-based cohort study that included 33 214 women with HG. We consider this a surprising finding considering that Heitmann *et al.*[23] found that 75.7% (159/210 women) of those with severe nausea and vomiting symptoms considered never getting pregnant again. Furthermore, Fejzo *et al.*[21] reported that 37% (37/100 women) had

avoided any further pregnancies due to HG. Our study found that 33% curtailed pregnancies due to HG. We also found that 40% of the women who got pregnant again after their index pregnancy postponed their pregnancy due to HG in the past. Poursharif *et al.*[24] described that 76% (614/808) of participants in their large self-selected online-survey cohort reported a change in personal attitude to future childbearing following an HG pregnancy, including increased spacing of pregnancies or fewer children than previously desired. This phenomenon is also described in a review from Dean *et al.*[25] reporting on HG's effect on women's lives and is recognized by our PPI representatives.

Previous surveys have identified the termination rate for HG between 3% and 15%.[3, 11, 26] Poursharif *et al.*[3] reported that 6% (49/808 women) underwent multiple terminations for HG and an additional 13% "almost" terminated their pregnancy due to HG. While in our study only one woman terminated a pregnancy due to suffering from HG again, we found that 23% considered terminating a pregnancy, which is consistent with a rate of 26.7% (56/210 women) reported in a Scandinavian population.[23] Variation in rates reported may reflect differences in access to treatment and social support around the world. For example, women included in Poursharif *et al.*[3] were predominantly from the USA, where sick pay and employment rights are not statutory and treatment for HG is expensive. Our study participants are from the Netherlands where treatment is covered by universal health care insurance and employees can make use of extended paid sick leave and are protected from termination of contract due to illness.[27] However, 23% of women considering termination of pregnancy due to HG is worryingly high and highlights the importance of early recognition and treatment of symptoms and supportive care.

Our study has several strengths. All participants had well documented HG during their index pregnancy, which is of benefit over previous studies which relied on hospital admission records, usually only including pregnancies that had led to a delivery. Including patient representatives in the conception and design of the research and interpretation of the results is also a strength, since this has been earlier recognized to improve quality and relevance of research.[28] Our study also assessed if subsequent pregnancies were viable and whether measures of HG severity of the index pregnancy can be used as predictive factors of recurrent HG, postponing and termination rates. These are both recognized as important matters in clinical practice by our patient representatives, especially since in our study, most women *without* vomiting symptoms had a non-viable pregnancy.

Our results are limited by the small sample size. Of the cohort of 215, 75 women (35%) responded, despite our recruitment efforts, and only 35 women (16%) had become pregnant again. Selective participation led to women with more severe symptoms in the index pregnancy being overrepresented in the current study, probably leading to selection bias. It is conceivable that those who did not participate in this follow-up study had a lower HG recurrence in subsequent pregnancies, which means that the recurrence rate presented here may be overestimated. Additionally, RCT analyses were hampered by an even smaller sample size, including 24 women who had become pregnant again. Also the association between having a western background and having weight loss in subsequent pregnancies is likely to be affected by the size of our study, with only 6/35 women that became pregnancy again having a non-western background. External validity may therefore be limited.

In this follow-up study, data was gathered through a self-reported, non-validated questionnaire, since there is no validated questionnaire available. The nature of some of the included questions could be considered as subjective. Additionally, there was potential for recall bias on pregnancies experienced up to seven years before participants completed the questionnaire as subsequent pregnancies could have taken place from 2013 onwards. This may have led to both under- and over reporting of HG symptoms, although previous studies have shown that self-reporting questionnaires are well validated for reporting on pregnancy.[29]

Conclusion

Our study found a high recurrence rate for HG of 89%. Although it seems plausible that selective attrition has occurred and led to an overestimation, our study suggests that the recurrence rate is more likely to be at the high end of the current available range of 15-81%. While such findings may be distressing for women who were hoping future pregnancies would be better, it is important information to give during preconception consultations, so that people are able to make informed decisions about their family planning. Knowing that HG has a very high chance of recurrence allows families to plan in advance for childcare and finances, but also to discuss available treatments and the possibility of early interventions, which may make the burden of the condition easier to bear. Finally, it is important that healthcare professionals do not give false hope regarding the chance of recurrence and to recognize the severe burden of the condition which leads so many women to consider, or actually terminate their otherwise wanted pregnancies.

Acknowledgements

We thank all women who participated in this follow-up study.

Funding

The MOTHER study was supported by a research grant from North West Hospital Group, Alkmaar, the Netherlands (Grant number: 2013T085). The follow-up study was funded by a research grant from the Amsterdam Reproduction and Development (AR&D) research institute, Amsterdam UMC, the Netherlands (Project number: 23346). Neither funders had any role in designing or conducting the studies or in interpreting the results.

Contribution to authorship

IJG, TJR and RCP conceived and designed the MOTHER study. KN and RCP conceived and designed the follow-up study. LMvdM helped develop the online survey tool. KN performed all statistical analyses supervised by RCP and RvE. KN and CD drafted the manuscript. The following authors were responsible for recruitment of, and data collection from the original MOTHER RCT and cohort: JMJB, CR-S, HAB, DPvdH, WMH, AH, GK, SK, JOEHvL, JL, FvdM, DP, M-JP, PJP, LvRF, RJR, HCJS, TV, BWM, MHK, IJG and RCP. All authors critically reviewed the manuscript and approved the final draft.

REFERENCES

1. Einarson TR, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: a meta analysis. Journal of population therapeutics and clinical pharmacology = Journal de la therapeutique des populations et de la pharamcologie clinique. 2013;20(2):e171-83.
2. Fiaschi L, Nelson-Piercy C, Gibson J, Szatkowski L, Tata LJ. Adverse Maternal and Birth Outcomes in Women Admitted to Hospital for Hyperemesis Gravidarum: a Population-Based Cohort Study. Paediatr Perinat Epidemiol. 2018;32(1):40-51.
3. Poursharif B, Korst L, MacGibbon K, Fejzo M, Romero R, Goodwin T. Elective pregnancy termination in a large cohort of women with hyperemesis gravidarum. Contraception. 2007;76(6):451-5.
4. Boelig RC, Barton SJ, Saccone G, Kelly AJ, Edwards SJ, Berghella V. Interventions for treating hyperemesis gravidarum. Cochrane Database Syst Rev. 2016;5:DOI: 10.1002/14651858.CD010607.pub2.
5. Fiaschi L, Nelson-Piercy C, Tata LJ. Hospital admission for hyperemesis gravidarum: a nationwide study of occurrence, reoccurrence and risk factors among 8.2 million pregnancies. Hum Reprod. 2016.
6. Dean CR, Bruin CM, O'Hara ME, Roseboom TJ, Leeflang MM, Spijker R, et al. The chance of recurrence of hyperemesis gravidarum: A systematic review. Eur J Obstet Gynecol Reprod Biol X. 2020;5:100105.
7. Dean C., Bierma H., Clarke R., Cleary B., Ellis P., Gadsby R., et al. A Patient-Clinician James Lind Alliance Partnership to Identify Research Priorities for Hyperemesis Gravidarum. BMJ Open. 2021;11(1):e041254.
8. Dean C. Helping women prepare for hyperemesis gravidarum. British Journal of Midwifery. 2014;22(12):847-52
9. Koren G, Maltepe C. Pre-emptive therapy for severe nausea and vomiting of pregnancy and hyperemesis gravidarum. J Obstet Gynaecol. 2004;24(5):530-3.
10. Fejzo M, Jalil S, MacGibbon K, Opper N, Romero R, Goodwin T, et al. Recurrence Risk of Hyperemesis Gravidarum. Reproductive Sciences. 2010;17(3):191A-2A.
11. Nana M, Tydeman F, Bevan G, Boulding H, Kavanagh K, Dean C, et al. Hyperemesis gravidarum is associated with increased rates of termination of pregnancy and suicidal ideation: results from a survey completed by >5000 participants. Am J Obstet Gynecol. 2021.
12. Iris J Grooten, Marjette H Koot, Joris AM van der Post, Joke MJ Bais, Carrie Ris-Stalpers, Christiana Naaktgeboren, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. The American journal of clinical nutrition. 2017;106(3):812-20.
13. Grooten I, Mol B, van der Post J, Ris-Stalpers C, Kok M, Bais J, et al. Early nasogastric tube feeding in optimising treatment for hyperemesis gravidarum: the MOTHER randomised controlled trial (Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding). BMC Pregnancy Childbirth. 2016;16(1):22.
14. Ebrahimi N, Maltepe C, Bournissen F, Koren G. Nausea and vomiting of pregnancy: using the 24-hour Pregnancy-Unique Quantification of Emesis (PUQE-24) scale. Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC. 2009;31(9):803-7.
15. Lacasse A, Rey E, Ferreira E, Morin C, Berard A. Validity of a modified Pregnancy-Unique Quantification of Emesis and Nausea (PUQE) scoring index to assess severity of nausea and vomiting of pregnancy. American Journal of Obstetrics and Gynecology. 2008;198(1).
16. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology. 1996;49(12):1373-9.
17. Trogstad L, Stoltenberg C, Magnus P, Skjaerven R, Irgens L. Recurrance risk in hyperemesis gravidarum. Bjog-an International Journal of Obstetrics and Gynaecology. 2005;112(12):1641-5.

18. Fell DB, Dodds L, Joseph KS, Allen VM, Butler B. Risk factors for hyperemesis gravidarum requiring hospital admission during pregnancy. Obstetrics and Gynecology. 2006;107(2):277-84.

19. Nurmi M, Rautava P, Gissler M, Vahlberg T, Polo-Kantola P. Recurrence patterns of hyperemesis gravidarum. Am J Obstet Gynecol. 2018;219(5):469 e1- e10.

20. Vikanes Å, Magnus P, Vangen S, Lomsdal S, Grjibovski A. Hyperemesis Gravidarum in the Medical Birth Registry of Norway - a validity study. BMC Pregnancy Childbirth. 2012;12:1-6.

21. Fejzo MS, MacGibbon KW, Romero R, Goodwin TM, Mullin PM. Recurrence Risk of Hyperemesis Gravidarum. Journal of Midwifery & Womens Health. 2011;56(2):132-6.

22. Ridic G, Gleason S, Ridic O. Comparisons of health care systems in the United States, Germany and Canada. Mater Sociomed. 2012;24(2):112-20.

23. Heitmann K, Nordeng H, Havnen GC, Solheimsnes A, Holst L. The burden of nausea and vomiting during pregnancy: severe impacts on quality of life, daily life functioning and willingness to become pregnant again - results from a cross-sectional study. BMC Pregnancy Childbirth. 2017;17(1):75.

24. Poursharif B, Korst L, Fejzo M, MacGibbon K, Romero R, Goodwin T. The psychosocial burden of hyperemesis gravidarum. Journal of Perinatology. 2008;28(3):176-81.

25. Dean C, Bannigan K, Marsden J. Reviewing the effect of hyperemesis gravidarum on women's lives and mental health. British Journal of Midwifery. 2018;26(2):109-19.

26. Mazzotta P, Stewart DE, Koren G, Magee LA. Factors associated with elective termination of pregnancy among Canadian and American women with nausea and vomiting of pregnancy. Journal of Psychosomatic Obstetrics and Gynecology. 2001;22(1):7-12.

27. Rijksinstituut voor Volksgezondheid en Milieu. Gelijke behandeling bij zwangerschap op het werk (Equal treatment at work during pregnancy). 2020 [Available from: https://www.rijksoverheid.nl/onderwerpen/gelijke-behandeling-op-het-werk/gelijke-behandeling-bij-zwangerschap-op-het-werk.

28. Lalani M, Baines R, Bryce M, Marshall M, Mead S, Barasi S, et al. Patient and public involvement in medical performance processes: A systematic review. Health Expect. 2019;22(2):149-61.

29. Falkegard M, Schirmer H, Lochen ML, Oian P, Acharya G. The validity of self-reported information about hypertensive disorders of pregnancy in a population-based survey: the Tromso Study. Acta Obstet Gynecol Scand. 2015;94(1):28-34.

Appendix S1. Follow-up questionnaire

Part A: Subsequent pregnancies after participating in the MOTHER study

1. After the pregnancy in which you participated in the MOTHER study, were you pregnant again? (A miscarriage, ectopic pregnancy or preterm birth also count!)
    a. Yes, continue with question 3.
    b. No, continue with question 22.
2. How many times have you been pregnant since participating in the MOTHER study?
    a. ….. *(fill in a number)*

If question 2 is answered with 2 or more then the following questions will be answered forevery pregnancy for a maximum of 5 pregnancies

3. In which year was this pregnancy?
    a. ….. *(For example '2015')*
4. Did you postpone this pregnancy because of the severity of the nausea and vomiting symptoms in your previous pregnancy?
    a. Yes
    b. No
5. Did this pregnancy end in a miscarriage or was this an ectopic pregnancy?
    a. No
    b. Yes, this pregnancy ended in a miscarriage
    c. Yes, this was an ectopic pregnancy
6. Did you experience any symptoms of nausea in this pregnancy?
    a. Yes, continue with question 7
    b. No, not at all. Continue with question 23 or go to '**adding a new pregnancy'**
7. How many weeks were you pregnant when you first felt nauseous?
    a. ….. *(in weeks)*
8. How many weeks were you pregnant when you first started vomiting?
    a. ….. *(in weeks)*
    b. I did not have complains of vomiting
9. How many weeks were you pregnant when the nausea and vomiting symptoms had practically disappeared?
    a. ….. *(in weeks)*

10. Were you admitted in the hospital with severe nausea and vomiting in this pregnancy?
    a. Yes, continue with question 11
    b. No, continue with question 13
11. How many times were you admitted in the hospital in this pregnancy?
    a. Once
    b. Twice
    c. 3 times
    d. More than 3 times, namely …. Times
12. How many days were you in total admitted in the hospital in this pregnancy? (the day of admission and the day of discharge both count as 1 day)
    a. ….. (Answer in days)
    b. I don't remember
13. Did you use any medication for the nausea and vomiting symptoms in this pregnancy?
    a. Yes, continue with question 14
    b. No, continue with question 15
14. Which medication did you use? (multiple options possible)
    a. Suprimal
    b. Emesafene
    c. Primperan (Metoclopramide)
    d. Zofran (Ondansetron)
    e. Potassium solution (potassium drink ('kaliumdrank') or potassium intravenous)
    f. Corticosteroids (methylprednisolon or hydrocortison)
    g. Omeprazole or Ranitidine
    h. Other, namely …. (free text)
    i. I don't remember which medication I used
15. Did you receive nasogastric tube feeding in this pregnancy?
    a. Yes
    b. No
16. Was this a singleton or a multiple pregnancy?
    a. Singleton pregnancy
    b. Twin pregnancy
    c. Multiple pregnancy of three or more babies (triplets or quadruplets)?

17. Did the severity of nausea and vomiting affect your ability to work?
    a. No, I was able to go to my work and did not have to call in sick at all
    b. Partly: I was not able to go to my work for some days
    c. Partly: I was not able to work for prolonged periods (eg weeks)
    d. I was not able to work at all
    e. I did not had a job at the time
18. Did the severity of nausea and vomiting affect your everyday life?
    a. Yes, the nausea and vomiting had an enormous effect on my everyday life
    b. The nausea and vomiting affected my everyday life to some degree
    c. No, The nausea and vomiting symptoms did not affect my everyday life at all
19. What was your weight before this pregnancy?
    a. …. (in kilograms)
    b. I don't remember
20. What was your lowest weight during this pregnancy?
    a. …. (in kilograms)
    b. I don't remember
21. Did you consider terminating this pregnancy because of the severity of the nausea and vomiting symptoms in this pregnancy or your previous pregnancy?
    a. Yes, I terminated this pregnancy because of the severity of nausea and vomiting
    b. I considered terminating this pregnancy, but in the end I continued this pregnancy
    c. No, I did not consider terminating this pregnancy
    d. Yes, I terminated this pregnancy, but due to other reasons than HG (f.e. congenital abnormalities or unwanted pregnancy)

- ***If answered 'no' to question 1 (and thus not have become pregnant again after participating in the MOTHER Study), continue with question 22***

- ***If answered 'yes' to question 1 (and thus finished question 18), continue with question 23***

22. After the pregnancy in which you participated in the MOTHER Study, did you're not becoming pregnant again have to do with the severity of the nausea and vomiting during that pregnancy, or fear of having hyperemesis gravidarum again?
    a. Yes
    b. No, there were other reasons

23. Do you have family members who also had severe nausea and vomiting or hyperemesis gravidarum in pregnancy?
    a. No
    b. Yes: (multiple options possible)
        i. Mother
        ii. Aunt
        iii. Sister
        iv. Grandmother
24. Do you have migraines?
    a. Yes
    b. No
25. Do you get motion sickness (eg car sick)?
    a. Yes
    b. No
26. What is your current height?
    a. ….. (in centimeters)
27. What is your current weight?
    b. ….. (in kilograms)

Part B: Depression and anxiety symptoms (HADS questionnaire)

Emotions play an important part in most illnesses. This questionnaire is designed to find out how you feel. Read each item below and tick the answer that comes closest to how you have been feeling in the past week.

28. I feel tense or 'wound up':
    a. Most of the time
    b. A lot of the time
    c. From time to time, occasionally
    d. Not at all
29. I still enjoy the things I used to enjoy:
    a. Definitely as much
    b. Not quite so much
    c. Only a little
    d. Hardly at all

30. I get a sort of frightened feeling as if something awful is about to happen:
    a. Very definitely and quite badly
    b. Yes, but not too badly
    c. A little, but it doesn't worry me
    d. Not at all
31. I can laugh and see the funny side of things:
    a. As much as I always could
    b. Not quite so much now
    c. Definitely not so much now
    d. Not at all
    32. Worrying thoughts go through my mind:

a. A great deal of the time
b. A lot of the time
c. From time to time, but not too often
d. Only occasionally
    33. I feel cheerful:

a. Not at all
b. Not often
c. Sometimes
d. Most of the time
    34. I can sit at ease and feel relaxed:
    a. Definitely
    b. Usually
    c. Not often
    d. Not at all
35. I feel as if I am slowed down:
    a. Nearly all the time
    b. Very often
    c. Sometimes
    d. Not at all
36. I get a sort of frightened feeling like 'butterflies' in the stomach:
    a. Not at all
    b. Occasionally
    c. Quite often
    d. Very often

37. I have lost interest in my appearance:
    a. Definitely
    b. I don't take as much care as I should
    c. I may not take quite as much care
    d. I take just as much care as ever
38. I feel restless as I have to be on the move
    a. Very much indeed
    b. Quite a lot
    c. Not very much
    d. Not at all
39. I look forward with enjoyment to things:
    a. As much as I ever did
    b. Rather less than I used to
    c. Definitely less than I used to
    d. Hardly at all
40. I get sudden feelings of panic
    a. Very often indeed
    b. Quite often
    c. Not very often
    d. Not at all
41. I can enjoy a good book or radio or TV program:
    a. Often
    b. Sometimes
    c. Not often
    d. Very seldom

Part C: Post-traumatic stress symptoms (PCL-5 questionnaire)

Below is a list of problems that people sometimes have in response to a very stressful experience. For the next questions, keep your pregnancy complicated by hyperemesis gravidarum in mind, please read each problem carefully and then circle one of the numbers to the right to indicate how much you have been bothered by that problem in the past month.

42. Repeated, disturbing, and unwanted memories of the stressful experience?
    a. Not at all

b. A little bit
c. Moderately
d. Quite a bit
e. Extremely

43. Repeated, disturbing dreams of the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
44. Suddenly feeling or acting as if the stressful experience were actually happening again (as if you were actually back there reliving it)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
45. Feeling very upset when something reminded you of the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
46. Having strong physical reactions when something reminded you of the stressful experience (for example, heart pounding, trouble breathing, sweating)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
47. Avoiding memories, thoughts, or feelings related to the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely

48. Avoiding external reminders of the stressful experience (for example, people, places, conversations, activities, objects, or situations)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
49. Trouble remembering important parts of the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
50. Having strong negative beliefs about yourself, other people, or the world (for example, having thoughts such as: I am bad, there is something seriously wrong with me, no one can be trusted, the world is completely dangerous)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
51. Blaming yourself or someone else for the stressful experience or what happened after it?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
52. Having strong negative feelings such as fear, horror, anger, guilt, or shame?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
53 Loss of interest in activities that you used to enjoy?
    a. Not at all
    b. A little bit

c. Moderately
d. Quite a bit
e. Extremely

54. Feeling distant or cut off from other people?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
55. Trouble experiencing positive feelings (for example, being unable to feel happiness or have loving feelings for people close to you)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
56. Irritable behavior, angry outbursts, or acting aggressively?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
57. Taking too many risks or doing things that could cause you harm?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
58. Being "super alert" or watchful or on guard?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
59. Feeling jumpy or easily startled?
    a. Not at all

b. A little bit
c. Moderately
d. Quite a bit
e. Extremely

60. Having difficulty concentrating?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
61. Trouble falling or staying asleep?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit

e. Extremely62. Have u experienced another very stressful or traumatic event?
    a. Yes, continue with question 63
    b. No, continue with question 65
    63. In which year did this very stressful or traumatic event happened?
    a. ….. (year)

64. What kind of very stressful or traumatic event did you experienced?
    a. Sexual assault
    b. Physical assault, violence or abuse
    c. Seeing someone be killed or seriously injured
    d. Dying of a loved one
    e. War
    f. Other, namely ….. (free text)
65. Do you wish to be informed about the results of this follow up of the MOTHER study?
    a. Yes
    b. No

Appendix S2. Follow-up MOTHER RCT – Methods and results

Methods

The MOTHER RCT was conducted between 2013 and 2016 and aimed to evaluate whether early enteral tube feeding in addition to standard care was beneficial to standard care alone. One hundred and fifteen women hospitalized for HG in 19 different hospitals in the Netherlands with a gestational age between 5 and 20 weeks were randomized. As published previously, the RCT found no differences in maternal and neonatal outcomes.[1] Since this study is a follow-up of an RCT, we also assessed whether there were differences in recurrence, postponing and termination rates between the RCT arms.

Statistical analysis

We assessed differences in recurrence, postponing and termination rates in subsequent pregnancies between the RCT arms according to the intention-to-treat, per protocol and as treated analyses. In the intention to treat analyses we compared differences in outcomes between RCT arms in how they were allocated. In the per-protocol analysis, early enteral tube feeding was defined as receiving nasogastric tube feeding within *three* days after randomization for at least seven days. In the as treated analysis, early enteral tube feeding was defined as receiving tube feeding within *seven* days after randomization for at least seven days.

A sensitivity analysis was performed to assess whether there were differences in baseline demographics and characteristics between RCT participants who were included in this follow-up study and RCT participants who were lost to follow-up. We used independent Student's t test, Mann-Whitney U test, chi-square test and Fisher's exact test for analyses and considered *P*-values <0.05 statistically significant. SPSS Statistics 26.0 for Windows (IBM Corp., Armonk, NY, USA) was used for analyses.

Results

RCT participants who received tube feeding were admitted to the hospital for longer periods in their subsequent pregnancies compared to RCT participants receiving standard care. This finding was found in the intention to treat (7.0 (5.0-8.0) vs 2.3 (1.1-3.8) days, *P*<0.01), as treated (6.0 (4.8-9.8) vs 2.3 (1.1-3.8) days, *P*=0.01) and per protocol analysis (6.0 (4.8-9.8) vs 2.3 (1.1-3.8) days, *P*=0.01) **(Supporting Information Tables S3-S5).** The number of hospitalizations for HG in subsequent pregnancies did not differ between the treatment arms.

In the intention to treat analysis, women allocated to early enteral tube feeding more often had weight loss and also lost more weight in subsequent pregnancies (86.7%, median weight loss -4.5 kg (-7.3 to -2.9)) than women allocated to standard care (33.3%, median weight loss 0.0 kg (-2.8 to 0.0), *P*=0.02 and 0.01). In both the intention to treat and per protocol analyses, women who received early enteral tube feeding also more often had >5% weight loss than women receiving standard care (*P*=0.01 and 0.03 respectively). A sensitivity analysis did not reveal any differences in measures of HG severity in the index pregnancy between RCT participants who participated in this follow-up study and those who did not. **(Supporting Information Table S6)**

References Appendix S2

1. Grooten IJ, Koot MH, van der Post JA, Bais JM, Ris-Stalpers C, Naaktgeboren C, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. The American journal of clinical nutrition. 2017;106(3):812-20.

Supporting Information Table S1. Sensitivity analysis between women included in the follow-up study and women who were lost to follow-up

	Women included in follow-up study N=73	**Women not included in follow-up study N=142**
Baseline characteristics		
Age (years), mean ± SD	29.2 ± 4.6	28.6 ± 5.0
Education level, n (%)		
- Primary or secondary	27 (37.0%)	60 (42.3%)
- Higher	29 (39.7%)	28 (19.7%)
Primigravida at time of MOTHER inclusion, n (%)	27 (37.0%)	37 (26.1%)
HG in pregnancy prior to MOTHER inclusion, n (%) [1]	22 (47.8%)	46 (43.8%)
HG severity during index pregnancy (MOTHER study)		
Weight change (kg), mean ± SD [2]	-3.4 ± 3.9	-2.7 ± 4.2
PUQE-24 at inclusion, median (IQR)	11.0 (9.0-13.0)	9.0 (7.0-12.0)
Mean PUQE-24 in the first 3 weeks after inclusion, mean ± SD	9.3 ± 2.4	8.8 ± 3.0
Total duration of hospital admissions (days), median (IQR)	5.0 (3.5-8.0)	5.0 (3.0-8.0)
Readmitted, n (%)	23 (31.5%)	48 (33.8%)
Admission in 2nd or 3rd trimester, n (%)	14 (19.2%)	29 (20.4%)

Data represented with mean±SD, median (IQR) or frequency (%). Significant P-values<0.05 are marked in bold.
* *P*-values using Independent Student's *t*-test, † *P*-values using Mann-Whitney U test, ‡ *P*-values using Chi-square test. [1] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [2] Weight change is weight at baseline minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. **Abbreviations:** HG: hyperemesis gravidarum, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms.

Supporting Information Table S2. Baseline characteristics of women included in this follow-up study and who became pregnancy again after participation in the MOTHER study

	N=35 N/Mean/ Median	%/SD/IQR	Missing
Baseline characteristics			
Age (years), mean ± SD	27.8	± 4.3	0.0%
Education level, n (%)			22.9%
- Primary or secondary	14	40.0%	
- Higher	13	37.1%	
Ethnicity, n (%)			11.4%
- Western	25	71.4%	
- Non-Western	6	17.1%	
Primigravida at time of MOTHER inclusion, n (%)	20	57.1%	0.0%
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	5	33.3%	6.7%
HG severity during index pregnancy (MOTHER study)			
Weight change (kg), mean ± SD [b]	-3.5	± 4.6	0.0%
PUQE-24 at inclusion, median (IQR)	12.0	10.0-14.0	17.1%
Total duration of hospital admissions (days), median (IQR)	5.0	3.0-8.0	0.0%
Admission in 2nd or 3rd trimester, n (%)	9	25.7%	0.0%

Data represented with mean ± SD, median (IQR) or frequency (%).
[a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at inclusion minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. **Abbreviations**: HG: hyperemesis gravidarum, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms.

Supporting Information Table S3. RCT analysis according to intention-to-treat

In total: N=49	Enteral tube feeding N=25	Standard care N=24	*P*-value
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years), mean ± SD	28.9 ± 5.0	30.1 ± 4.5	0.38*
Education level, n (%)			0.17‡
- Primary or secondary	9 (36.0%)	12 (50.0%)	
- Higher	11 (44.0%)	10 (41.7%)	
Primigravida at time of MOTHER inclusion, n (%)	11 (44.0%)	5 (20.8%)	0.08‡
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	7 (50.0%)	11 (57.9%)	0.65‡
Measures of HG severity			
Weight change (kg), mean ± SD [b]	-2.8 ± 4.1	-4.7 ± 4.4	0.12*
PUQE-24 at inclusion, median (IQR)	12.0 (10.5-14.0)	12.0 (10.0-13.0)	0.40†
Total duration of hospital admission(s) (days), median (IQR)	6.0 (3.0-9.0)	5.0 (3.3-9.5)	0.99†

Supporting Information Table S3. *Continued*

In total: N=49	Enteral tube feeding N=25	Standard care N=24	*P*-value
Admission in 2nd or 3rd trimester, n (%)	7 (28.0%)	3 (12.5%)	0.29§
Follow-up study			
Women who got pregnant again, n (%)	15 (60.0%)	9 (37.5%)	0.12‡
Number of pregnancies after MOTHER study, n (%)			0.46‡
- 1 pregnancy	12 (48.0%)	6 (25.0%)	
- 2 pregnancies	3 (12.0%)	2 (8.3%)	
- 3 or more pregnancies	-	1 (4.2%)	
Women who did not become pregnant again due to fear for recurrent HG, n (%)	6 (24.0%)	8 (33.3%)	1.00§
Time interval between MOTHER and follow-up study participation (years), median (IQR)	4.7 (3.9-5.1)	4.6 (4.1-4.9)	0.87†
Recurrence rate of HG			
Reoccurrence of vomiting symptoms, n (%)	14 (93.3%)	9 (100.0%)	1.00§
Gestation when vomiting started (weeks), median (IQR)	5.5 (4.8-6.3)	6.0 (4.3-7.0)	0.56†
Used multiple (2 or more) HG related medication, n (%)	11 (73.3%)	5 (55.6%)	0.41§
Was admitted to the hospital, n (%)	12 (80.0%)	4 (44.4%)	0.10§
Duration of hospital admissions (days), median (IQR)	7.0 (5.0-8.0)	2.3 (1.1-3.8)	**<0.01**†
- Had weight loss, n (%) [c]	13 (86.7%)	3 (33.3%)	**0.02**§
- Had >5% weight loss, n (%)	10 (66.7%)	1 (11.1%)	**0.01**§
Average weight change (kg), median (IQR) [c]	-4.5 (-7.3 to -2.9)	0.0 (-2.8-0.0)	**0.01**†
Received tube feeding, n (%)	3 (20.0%)	0 (0.0%)	0.27§
NVP affected her job, n (%)	12 (80.0%)	4 (44.4%)	0.10§
NVP affected her life, n (%)	13 (86.7%)	6 (66.7%)	0.33§
Had HG, n (%) [d]	14 (93.3%)	9 (100.0%)	1.00§
Postponed or terminated pregnancies			
Postponed a pregnancy due to HG in the past, n (%)	6 (24.0%)	3 (12.5%)	1.00§
Considered terminating a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	5 (33.3%)	2 (22.2%)	0.67§
Terminated a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	1 (6.7%)	0 (0.0%)	1.00§

Data represented with mean±SD, median (IQR) or frequency (%). Significant P-values<0.05 are marked in bold.
* *P*-values using Independent Student's t-test, † *P*-values using Mann-Whitney U test, ‡ *P*-values using Chi-square test, § *P*-values using Fisher's exact test. [a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at baseline minus prepregnancy weight. [c] Lowest weight during pregnancy minus prepregnancy weight. Both can be < 0 if women lost weight and can be > 0 if women gained weight. [d] HG defined as: vomiting symptoms with either multiple medication use, hospital admission, weight loss during pregnancy, tube feeding, NVP affecting her job and/or life. **Abbreviations**: HG: hyperemesis gravidarum, NVP: nausea and vomiting in pregnancy, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms, RCT: randomized controlled trial.

Supporting Information Table S4. RCT analysis according to as treated

In total: N=35	Enteral tube feeding N=15	Standard care N=20	*P*-value
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years), mean ± SD	27.7 ± 3.8	30.6 ± 4.3	**0.047***
Education level, n (%)			0.55‡
- Primary or secondary	5 (33.3%)	10 (50.0%)	
- Higher	7 (46.7%)	9 (45.0%)	
Primigravida at time of MOTHER inclusion, n (%)	7 (46.7%)	5 (25.0)	0.18‡
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	6 (75.0%)	9 (60.0%)	0.66§
Measures of HG severity			
Weight change (kg), mean ± SD [b]	-6.0 ± 5.6	-4.0 ± 2.3	0.22*
PUQE-24 at inclusion, median (IQR)	12.5 (11.0-14.0)	11.0 (9.5-13.0)	0.15†
Total duration of hospital admission(s) (days), median (IQR)	7.0 (3.0-10.0)	4.0 (3.0-5.8)	0.20†
Admission in 2nd or 3rd trimester, n (%)	2 (13.3%)	2 (10.0%)	1.00§
Follow-up study			
Women who got pregnant again, n (%)	8 (53.3%)	7 (35.0%)	0.28‡
Number of pregnancies after MOTHER study, n (%)			0.31‡
- 1 pregnancy	4 (50.0%)	5 (71.4%)	
- 2 pregnancies	3 (37.5%)	2 (28.6%)	
- 3 or more pregnancies	1 (12.5%)	0 (0.0%)	
Women who did not become pregnant again due to fear for recurrent HG, n (%)	6 (85.7%)	7 (53.8%)	0.33§
Time interval between MOTHER and follow-up study participation (years), median (IQR)	4.7 (3.9-5.0)	4.6 (4.1-5.0)	0.86†
Recurrence rate of HG			
Reoccurrence of vomiting symptoms, n (%)	8 (100.0%)	7 (100.0%)	-
Gestation when vomiting started (weeks), median (IQR)	6.0 (5.0-6.8)	6.0 (3.5-7.0)	1.00†
Used multiple (2 or more) HG related medication, n (%)	7 (87.5%)	4 (57.1%)	0.28§
Was admitted to the hospital, n (%)	6 (75.0%)	4 (57.1%)	0.61§
Duration of hospital admissions (days), median (IQR)	6.0 (4.8-9.8)	2.3 (1.1-3.8)	**0.01**†
Had weight loss, n (%) [c]	6 (75.0%)	2 (28.6%)	0.13§
Had >5% weight loss, n (%)	5 (62.5%)	1 (14.3%)	0.12§
Average weight change (kg), median (IQR) [c]	-4.0 (-8.0 to -2.0)	0.0 (-3.8-0.0)	0.23†
Received tube feeding, n (%)	2 (25.0%)	0 (0.0%)	0.47§
NVP affected her job, n (%)	6 (75.0%)	4 (57.1%)	0.61§
NVP affected her life, n (%)	7 (87.5%)	5 (71.4%)	0.57§
Had HG, n (%) [d]	8 (100.0%)	7 (100.0%)	-

Supporting Information Table S4. *Continued*

In total: N=35	Enteral tube feeding N=15	Standard care N=20	*P*-value
Postponed or terminated pregnancies			
Postponed a pregnancy due to HG in the past, n (%)	4 (26.7%)	3 (15.0%)	1.00§
Considered terminating a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	4 (50.0%)	1 (14.3%)	0.28§
Terminated a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	1 (12.5%)	0 (0.0%)	1.00§

Data represented with mean±SD, median (IQR) or frequency (%). Significant P-values<0.05 are marked in bold.
* *P*-values using Independent Student's t-test, † *P*-values using Mann-Whitney U test, ‡ *P*-values using Chi-square test, § *P*-values using Fisher's exact test. [a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at baseline minus prepregnancy weight. [c] Lowest weight during pregnancy minus prepregnancy weight. Both can be < 0 if women lost weight and can be > 0 if women gained weight. [d] HG defined as: vomiting symptoms with either multiple medication use, hospital admission, weight loss during pregnancy, tube feeding, NVP affecting her job and/or life. **Abbreviations**: HG: hyperemesis gravidarum, NVP: nausea and vomiting in pregnancy, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms, RCT: randomized controlled trial.

Supporting Information Table S5. RCT analysis according to per protocol

In total: N=32	Enteral tube feeding N=12	Standard care N=20	*P*-value
<u>Index pregnancy (during MOTHER study)</u>			
Baseline characteristics			
Age (years), mean ± SD	28.0 ± 3.4	30.6 ± 4.3	0.08*
Education level, n (%)			0.70§
- Primary or secondary	4 (33.3%)	10 (50.0%)	
- Higher	6 (50.0%)	9 (45.0%)	
Primigravida at time of MOTHER inclusion, n (%)	7 (58.3%)	5 (25.0%)	0.13§
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	4 (80.0%)	9 (60.0%)	0.61§
Measures of HG severity			
Weight change (kg), mean ± SD [b]	-4.9 ± 3.7	-4.0 ± 2.3	0.39*
PUQE-24 at inclusion, median (IQR)	13.0 (11.0-14.0)	11.0 (9.5-13.0)	0.13†
Total duration of hospital admission(s) (days), median (IQR)	5.0 (3.0-10.3)	4.0 (3.0-5.8)	0.50†
Admission in 2nd or 3rd trimester, n (%)	2 (16.7%)	2 (10.0%)	0.62§
<u>Follow-up study</u>			
Women who got pregnant again, n (%)	6 (50.0%)	7 (35.0%)	0.47§
Number of pregnancies after MOTHER study, n (%)			0.45‡
- 1 pregnancy	3 (50.0%)	5 (71.4%)	
- 2 pregnancies	3 (50.0%)	2 (28.6%)	
- 3 or more pregnancies	0 (0.0%)	0 (0.0%)	

Supporting Information Table S5. *Continued*

In total: N=32	Enteral tube feeding N=12	Standard care N=20	*P*-value
Women who did not become pregnant again due to fear for recurrent HG, n (%)	5 (83.3%)	7 (53.8%)	0.33§
Time interval between MOTHER and follow-up study participation (years), median (IQR)	4.6 (3.8-5.1)	4.6 (4.1-5.0)	0.72†
Recurrence rate of HG			
Reoccurrence of vomiting symptoms, n (%)	6 (100.0%)	7 (100.0%)	-
Gestation when vomiting started (weeks), median (IQR)	5.5 (4.8-7.0)	6.0 (3.5-7.0)	0.95†
Used multiple (2 or more) HG related medication, n (%)	6 (100.0%)	4 (57.1%)	0.19§
Was admitted to the hospital, n (%)	6 (100.0%)	4 (57.1%)	0.19§
Duration of hospital admissions (days), median (IQR)	6.0 (4.8-9.8)	2.3 (1.1-3.8)	**0.01**†
Had weight loss, n (%) [c]	5 (83.3%)	2 (28.6%)	0.10§
Had >5% weight loss, n (%)	5 (83.3%)	1 (14.3%)	**0.03**§
Average weight change (kg), median (IQR) [c]	-5.0 (-8.5 to -3.3)	0.0 (-3.8-0.0)	0.052†
Received tube feeding, n (%)	2 (33.3%)	0 (0.0%)	0.19§
NVP affected her job, n (%)	6 (100.0%)	4 (57.1%)	0.19§
NVP affected her life, n (%)	6 (100.0%)	5 (71.4%)	0.64§
Had HG, n (%) [d]	6 (100.0%)	7 (100.0%)	-
Postponed or terminated pregnancies			
Postponed a pregnancy due to HG in the past, n (%)	4 (33.3%)	3 (15.0%)	0.59§
Considered terminating a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	3 (50.0%)	1 (14.3%)	0.27§
Terminated a pregnancy due to suffering from HG in subsequent pregnancies, n (%)	1 (16.7%)	0 (0.0%)	0.46§

Data represented with mean±SD, median (IQR) or frequency (%). Significant P-values<0.05 are marked in bold.
* *P*-values using Independent Student's t-test, † *P*-values using Mann-Whitney U test, ‡ *P*-values using Chi-square test, § *P*-values using Fisher's exact test. [a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at baseline minus prepregnancy weight. [c] Lowest weight during pregnancy minus prepregnancy weight. Both can be < 0 if women lost weight and can be > 0 if women gained weight. [d] HG defined as: vomiting symptoms with either multiple medication use, hospital admission, weight loss during pregnancy, tube feeding, NVP affecting her job and/or life. **Abbreviations**: HG: hyperemesis gravidarum, NVP: nausea and vomiting in pregnancy, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms, RCT: randomized controlled trial.

Supporting Information Table S6. Sensitivity analysis between women of the RCT included in the follow-up study and women of the RCT who were lost to follow-up

	RCT participants included in follow-up study N=49	**RCT participants not included in follow-up study N=66**	***P*-value**
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years), mean ± SD	29.5 ± 4.8	27.7 ± 4.7	0.050*
Education level, n (%)			**0.03**‡
- Primary or secondary	21 (42.9%)	34 (51.5%)	
- Higher	21 (42.9%)	13 (19.7%)	
Primigravida at time of MOTHER inclusion, n (%)	16 (32.7%)	20 (30.3%)	0.79‡
HG in pregnancy prior to MOTHER inclusion, n (%) [a]	18 (54.5%)	23 (50.0%)	0.69‡
Measures of HG severity			
Weight change (kg), mean ± SD [b]	-3.7 ± 4.3	-3.1 ± 4.2	0.47*
PUQE-24 at inclusion, median (IQR)	12.0 (10.0-14.0)	11.0 (7.5-13.5)	0.06†
Total duration of hospital admission(s) (days), median (IQR)	5.0 (3.0-9.0)	5.0 (4.0-8.3)	0.41†
Admission in 2nd or 3rd trimester, n (%)	10 (20.4%)	18 (27.3%)	0.40‡

Data represented with mean±SD, median (IQR) or frequency (%). Significant P-values<0.05 are marked in bold.
* *P*-values using Independent Student's t-test, † *P*-values using Mann-Whitney U test, ‡ *P*-values using Chi-square test, [a] Percentage shown is frequency divided by multigravidas at time of MOTHER inclusion. [b] Weight change is weight at baseline minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. **Abbreviations:** HG: hyperemesis gravidarum, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms, RCT: randomized controlled trial.

CHAPTER 5

Depression, anxiety and post-traumatic stress disorder symptoms after hyperemesis gravidarum: a prospective cohort study

Kelly Nijsten, Loïs M. van der Minnen, Caitlin Dean, Joke M.J. Bais, Carrie Ris-Stalpers, Rik van Eekelen, Henk A. Bremer, David P. van der Ham, Wieteke M. Heidema, Anjoke Huisjes, Gunilla Kleiverda, Simone M. Kuppens, Judith O.E.H. van Laar, Josje Langenveld, Flip van der Made, Dimitri Papatsonis, Marie-José Pelinck, Paula J. Pernet, Leonie van Rheenen-Flach, Robbert J. Rijnders, Hubertina C.J. Scheepers, Tatjana Vogelvang, Ben W. Mol, Miranda Olff, Tessa J. Roseboom, Marjette H. Koot, Iris J. Grooten, Rebecca C. Painter

Journal of Maternal-Fetal & Neonatal Medicine. 2022:1-9.

ABSTRACT

Objective

To determine the prevalence of depression, anxiety and posttraumatic stress disorder (PTSD) years after hyperemesis gravidarum (HG) and its association with HG severity.

Material and methods

This prospective cohort study consisted of a follow-up of 215 women admitted for HG, who were eligible to participate in a randomized controlled trial and either declined or agreed to be randomized between 2013 and 2016 in 19 hospitals in the Netherlands. Participants completed the Hospital Anxiety and Depression Scale (HADS) six weeks postpartum and during follow-up and the PTSD checklist for DSM-5 (PCL-5) during follow-up. An anxiety or depression score ≥8 is indicative of an anxiety or depression disorder and a PCL-5 ≥31 indicative of PTSD. Measures of HG severity were symptom severity (PUQE-24: Pregnancy Unique Quantification of Emesis), weight change, duration of admissions, readmissions and admissions after the first trimester.

Results

54/215 participants completed the HADS six weeks postpartum and 73/215 participants completed the follow-up questionnaire, on average 4.5 years later. Six weeks postpartum, 13 participants (24.1%) had an anxiety score ≥8 and 11 participants (20.4%) a depression score ≥8. During follow-up, 29 participants (39.7%) had an anxiety score ≥8, 20 participants (27.4%) a depression score ≥8 and 16 participants (21.9%) a PCL-5 ≥31.

Multivariable logistic regression analysis showed that for every additional point of the mean PUQE-24 three weeks after inclusion, the likelihood of having an anxiety score ≥8 and PCL-5 ≥31 at follow-up increased with OR 1.41 (95% CI: 1.10;1.79) and OR 1.49 (95% CI: 1.06;2.10) respectively.

Conclusion

Depression, anxiety and PTSD symptoms are common years after HG occurred.

INTRODUCTION

Hyperemesis gravidarum (HG) is a severe form of nausea and vomiting in pregnancy. Dehydration, electrolyte disturbances or weight loss can necessitate hospital admission for intravenous rehydration or tube feeding. [1] A systematic review published in 2017 showed a higher incidence of depression and anxiety symptoms during pregnancy in women suffering from HG.[2] Some studies have suggested that psychiatric diagnoses predispose to HG,[3] whereas others have argued that HG *causes* depression, anxiety as well as posttraumatic stress disorder (PTSD) symptoms.[4-6] The fact that HG symptom improvement has been associated with a reduction in anxiety and depression symptoms, supports the latter of the two hypotheses.[7-9]

Increases in depression and PTSD symptoms after pregnancies complicated by HG have been reported.[4-6, 10, 11] There have been suggestions of a possible dose-response effect with higher depression and PTSD scores postpartum among women with increased HG symptoms or with a prolonged disease course, hinting at a causal relationship between HG and psychopathology that persists postpartum.[4, 5, 11]

Altogether, there is a limited body of evidence on the size and strength of the association between HG symptom severity and depression, anxiety and PTSD symptoms. Furthermore, recently the long term maternal mental health consequences of HG were identified by patients and clinicians as one of the top 10 HG priority research questions.[12] Therefore, in the present study, we aim to prospectively determine the association between HG symptom severity and depression, anxiety and PTSD symptoms in women years after HG diagnosis.

MATERIAL AND METHODS

Study design MOTHER

Our study is a prospective follow-up study of the MOTHER (Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding) study.[13] The MOTHER study included women admitted for HG between 5 and 20 weeks gestation and consisted of a randomized controlled trial (RCT) and associated observational cohort of women who were eligible for participation in the trial but who declined randomization. The RCT assessed whether early enteral tube feeding in addition to standard care for women admitted with HG improved maternal or perinatal outcomes. Since this was not the case, we were able to combine the RCT and cohort into one study population for this follow-up study. Between 2013 and 2016, the RCT included 115 women and the cohort another 100 women. The MOTHER study was approved by the research ethics committee of the Amsterdam UMC and registered at the Dutch trial register (NTR4197). More detailed information about the MOTHER study can be found in the previous published study protocol and earlier work.[14, 15]

Measures of HG severity of the index pregnancy

The MOTHER study prospectively collected detailed information about pregnancy and delivery, extracted by trained research staff from medical files. Data regarding medical history, including a history or ongoing disease of depressive, anxiety and/or eating disorders, were also collected from medical files. Participants self-reported their pre-pregnancy weight, ethnicity and highest completed level of education. If self-reported data was missing, these features were extracted from medical file where available.

In this study, the pregnancy in which women had participated in the MOTHER study was designated as the index pregnancy and measures of HG severity in this pregnancy were used as predictor variables in regression analysis. HG severity in the index pregnancy was assessed by symptom severity at inclusion, as measured by the self-reported 24-hour Pregnancy Unique Quantification of Emesis (PUQE-24) score, and the mean PUQE-24 three weeks after inclusion, weight change, total duration of hospital admissions, being readmitted and being admitted after the first trimester. The PUQE-24 score could vary from 3 to 15 with a higher PUQE-24 score indicating more severe symptoms.[16] The mean PUQE-24 score in the first three weeks after inclusion was calculated by summing up available weekly PUQE-24 scores and dividing by the total number of weekly PUQE-24 scores available. Weight change was calculated by comparing weight at inclusion to pre-pregnancy weight.

Follow-up study procedures

MOTHER study participants who gave consent to be approached for follow-up studies were invited for this follow-up study by email. Participants completed a single online questionnaire. Reminder emails were sent on three occasions to complete the questionnaire. Informed consent was obtained separately for the MOTHER study and the present follow-up study. Ethical approval for the follow-up study was not necessary, according to the Medical Research Involving Human Subjects Act (W20_066 #20.094).

The questionnaire consisted of questions about maternal mental health, as well as subsequent pregnancies after inclusion in the MOTHER. We asked whether participants had suffered from HG again after the MOTHER study, defining HG as vomiting symptoms which occurred with either: multiple medication use, weight loss, hospital admission for HG, tube feeding or in which nausea and vomiting symptoms affected their life and/or work. The full follow-up questionnaire is enclosed as **Appendix S1**. More detailed findings on HG recurrence rates in subsequent pregnancies of this follow-up study are previously published.[17]

Outcome measures: depression, anxiety and PTSD symptoms

Depression and anxiety symptoms were self-reported at inclusion of the MOTHER study and six weeks postpartum by use of the Hospital Anxiety and Depression Scale (HADS).[18] A higher HADS indicates more severe depression or anxiety symptoms: a depression or anxiety score greater than or equal to 8 is considered borderline and greater than or equal to 11 as abnormal.[18] Depression and anxiety symptoms were again assessed during the follow-up study using the HADS. PTSD symptoms were assessed during follow-up using the PTSD CheckList for DSM-5 (PCL-5).[19, 20] A higher PCL-5 score indicates more severe PTSD symptoms: a PCL-5 score greater or equal to 31 indicates PTSD. The follow-up questionnaire did not include questions about possible treatments for depression, anxiety or PTSD symptoms. The questionnaire was available in both Dutch and English. The HADS has been validated in both languages,[21] while the PCL-5 has only been validated in English.[22]

Statistical analyses

To assess selective attrition, baseline characteristics and measures of HG severity and course of the index pregnancy were compared between women participating in the follow-up study and those who were lost to follow-up or declined participation in the present study, using independent students *t* test, Mann-Whitney U test and Chi-square test.

Univariable logistic and linear regression analysis was used to assess possible risk factors for depression, anxiety and PTSD symptoms, including having an HG pregnancy between MOTHER and follow-up study participation. Subsequently, we performed univariable and multivariable logistic and linear regression analysis to assess the association between each of the measures of HG severity of the index pregnancy and dichotomous outcomes (HADS anxiety and depression score ≥8 and PCL-5 score ≥31) and continuous outcomes (total HADS and PCL-5 score). In multivariable *logistic* regression analysis, due to the small number of events, we were only able to adjust for a single confounder.[23] We opted to correct for the risk factor with the lowest *P*-value in univariable logistic regression analysis. In multivariable *linear* regression analysis we were able to adjust for more confounders and included risk factors which were significantly associated in univariable linear regression analysis based on a *P*-value <0.10 as confounders. Outcomes variables that were not normally distributed were log transformed, back transformed and expressed in percentages of differences.

Lastly, we assessed differences in depression, anxiety and PTSD symptoms between the treatment arms of the RCT. We performed an intention-to-treat, per protocol (receiving tube feeding within *3 days* after randomization and continued for at least 7 days) and as treated analysis (receiving tube feeding within *7 days* after randomization and continued for at least 7 days). Two-sided *P*-values <0.05 were considered statistically significant. We used SPSS Statistics 26.0 for Windows (IBM Corp., Armonk, NY, USA) for all analyses.

Patient and public involvement

The Dutch HG patient support group Zwangerschapsmisselijkheid en Hyperemesis Gravidarum was involved in setting up the MOTHER and the follow-up study. One author (CD) is a patient representative and gave perspective on the interpretation of the results.

RESULTS

Participants

From the 215 participating women in the MOTHER study, 54 women (25.1%) completed the HADS at six weeks postpartum between 2014 and 2016.**(Supplement Figure 1)** We approached 190 out of 215 women who gave consent to be contacted for follow-up studies and from whom we had an e-mail address available. 73 women (34.0%) completed the follow-up questionnaire between March and May 2020, on average 4.5 years later. Baseline characteristics are presented in **Table 1**.

As shown in **Supplement Table S1**, participants of the follow-up study were significantly more often of western ethnicity (71.2% vs. 50.0%, *P*=0.01), had higher educational attainment (39.7% vs. 19.7%, *P*=0.02) and had had higher PUQE-24 scores (11 (9-13) vs. 9 (7-12), *P*=0.01) than women who were lost to follow-up or declined participation in the present study.

Depression and anxiety symptoms

At inclusion of the MOTHER study, while suffering from HG, 31 out of 61 participants (50.8%) had an anxiety score ≥8 and 56 out of 61 participants (91.8%) had a depression score ≥8.**(Table 1)** Six weeks postpartum, 13 out of 54 participants (24.1%) had an anxiety score ≥8 and 11 participants (20.4%) had a depression score ≥8. At follow-up, 29 out of 73 participants (39.7%) had an anxiety score ≥8 and 20 women (27.4%) had a depression score ≥8.

We were not able to identify risk factors that were associated with a depression or anxiety score ≥8 at six weeks postpartum in univariable logistic regression.**(Supplement Table S2)** A history of any traumatic event (OR 3.63, 95% CI: 1.29;10.20) and having an HG pregnancy between MOTHER and follow-up participation (OR 2.74, 95% CI: 1.04;7.20) were associated with having an anxiety score ≥8 at follow-up.**(Supplement Table S2)** Younger maternal age (OR 0.84, 95% CI: 0.73;0.96), a higher HADS at inclusion of the MOTHER study (OR 1.14, 95% CI: 1.02;1.27) and a history of a traumatic event (OR 3.08, 95% CI: 1.05;9.03) were associated with a depression score ≥8 at follow-up.

HG severity of the index pregnancy and depression and anxiety symptoms

Multivariable logistic regression analysis showed that for every additional point of the mean PUQE-24 score in the first three weeks after inclusion of the index pregnancy, the likelihood of having an anxiety score ≥8 at time of the follow-up study increased with OR 1.41 (95% CI: 1.10;1.79).**(Table 2)** None of the HG severity measures of the index pregnancy were associated with depression and anxiety scores six weeks postpartum or depression score ≥8 and total HADS at follow-up in both logistic and linear regression analysis.**(Table 2 and Supplement Table S3)**

Depression and anxiety symptoms according to RCT treatment allocation

Among RCT participants, we did not find any differences in depression and anxiety symptoms at six weeks postpartum and at follow-up between women in the enteral tube feeding group and women in the standard care group in intention-to-treat, as treated and per protocol analysis.**(Supplement Tables S4-6)**

PTSD symptoms

At follow-up, 16 out of 73 participants (21.9%) had a PCL-5 score ≥31, indicating PTSD.**(Table 1)** A higher HADS at inclusion of the MOTHER study was associated with a PCL-5 score ≥31 in univariable logistic regression (OR 1.20, 95% CI: 1.04;1.29) and with higher PCL-5 scores in univariable linear regression (β 5.65, 95% CI: 1.31;10.08).**(Supplement Table S2)**.

Table 1. Baseline characteristics and outcome measures of women included in this follow-up study

	Index pregnancy (during MOTHER study)
Baseline characteristics	**N=73**
Age (years)	29.25 ± 4.62
Ethnicity	
- Western	52 (71.2%)
- Non-western	12 (16.4%)
Education level	
- Primary or secondary	27 (37.0%)
- Higher	29 (39.7%)
Primigravida at the time	26 (35.6%)
History of mental health disease [a]	12 (16.4%)
HG in previous pregnancy [b]	23 (48.9%)
HG in previous pregnancy requiring hospital admission [b]	12 (25.5%)
Maternal outcomes	
Weight change (kg) [c]	-3.40 ± 3.85
PUQE-24 at inclusion	11.00 (9.00-13.00)
Mean PUQE-24 in the first 3 weeks after admission	9.19 ± 2.50
Total duration of hospital admissions (days)	5.00 (4.00-8.00)
Readmitted	24 (32.9%)
Admission after the first trimester	14 (19.2%)
HADS at inclusion	21.28 ± 6.65
Anxiety score ≥8 [d]	31 (50.8%)
Depression score ≥8 [d]	56 (91.8%)
	Follow-up:
Depression, anxiety and PTSD symptoms	
HADS 6 weeks postpartum [e]	9.57 ± 6.24
- Anxiety score ≥8 [e]	13 (24.1%)
- Depression score ≥8 [e]	11 (20.4%)
HADS at follow-up study	10.00 (6.00-15.00)
- Anxiety score ≥8	29 (39.7%)
- Depression score ≥8	20 (27.4%)
PCL-5 score at follow-up study	13.00 (7.00-28.50)
- PCL-5 score ≥31	16 (21.9%)

Table 1. *Continued*

History of any traumatic event	23 (31.5%)
History of an obstetric traumatic event	3 (4.1%)
Women who experienced a subsequent pregnancy after MOTHER participation	35 (47.9%)
Suffered from HG again between participation of the MOTHER and follow-up study [f]	30 (41.1%)

Data represented with mean±SD**,** median (IQR) or frequency (%). [a] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [b] Percentage shown is frequency divided by multigravidas. [c] Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. [d] Percentage shown is frequency divided by number of HADS at inclusion available (n=61). [e] Different study group including 54 women. [f] HG defined as: if vomiting symptoms occurred with either multiple medication use, weight loss, hospital admission for HG, requiring tube feeding or whether nausea and vomiting symptoms affected their life and/or work. **Abbreviations:** HG: hyperemesis gravidarum, HADS: hospital anxiety and depression scale: a higher HADS indicates more severe anxiety or depression symptoms, PCL-5: PTSD checklist for the DSM 5, PTSD: posttraumatic stress disorder, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms and can vary from 3 to 15.

Table 2. Logistic regression analysis to assess the association between measures of HG severity in the index pregnancy and HADS at 6 weeks postpartum and HADS and PCL-5 score at follow-up study.

6 weeks postpartum	**HADS 6 weeks postpartum: Anxiety score ≥8**	
	OR	**95% CI**
Weight change (kg)	0.95	0.82;1.10
PUQE-24 at inclusion	0.83	0.63;1.08
Mean PUQE-24 in the first 3 weeks after admission	0.87	0.68;1.12
Total duration of hospital admission(s) (days)	1.08	0.94;1.24
Readmitted	2.02	0.57;7.14
Admission after the first trimester	2.22	0.58;8.49
	HADS 6 weeks postpartum: Depression score ≥8	
	OR	**95% CI**
Weight change (kg)	0.91	0.78;1.06
PUQE-24 at inclusion	0.76	0.57;1.02
Mean PUQE-24 in the first 3 weeks after admission	0.75	0.55;1.01
Total duration of hospital admission(s) (days)	1.12	0.97;1.29
Readmitted	1.28	0.34;4.84
Admission after the first trimester	1.89	0.46;7.78
Follow-up study	**HADS at follow-up: Anxiety score ≥8**	

Table 2. *Continued*

	Model 1		**Model 2***	
	OR	**95% CI**	**OR**	**95% CI**
Weight change (kg)	0.92	0.81;1.04	0.90	0.78;1.03
PUQE-24 at inclusion	1.03	0.86;1.23	1.07	0.88;1.30
Mean PUQE-24 in the first 3 weeks after admission	**1.32**	**1.06;1.64**	**1.41**	**1.10;1.79**
Total duration of hospital admission(s) (days)	0.95	0.86;1.05	0.95	0.86;1.06
Readmitted	0.58	0.22;1.52	0.63	0.23;1.73
Admission after the first trimester	0.81	0.24;2.72	0.74	0.21;2.63
	HADS at follow-up: Depression score ≥8			
	Model 1		**Model 2****	
	OR	**95% CI**	**OR**	**95% CI**
Weight change (kg)	0.89	0.77;1.03	0.92	0.80;1.07
PUQE-24 at inclusion	1.02	0.84;1.24	1.09	0.87;1.38
Mean PUQE-24 in the first 3 weeks after admission	1.12	0.90;1.38	1.15	0.89;1.48
Total duration of hospital admission(s) (days)	0.98	0.91;1.07	0.96	0.86;1.08
Readmitted	0.65	0.22;1.89	0.30	0.08;1.11
Admission after the first trimester	1.63	0.47;5.63	1.34	0.32;5.68
	PCL-5 score at follow-up ≥31			
	Model 1		**Model 2****	
	OR	**95% CI**	**OR**	**95% CI**
Weight change (kg)	0.95	0.83;1.10	0.99	0.82;1.18
PUQE-24 at inclusion	0.98	0.79;1.21	1.03	0.81;1.32
Mean PUQE-24 in the first 3 weeks after admission	**1.30**	**1.01;1.67**	**1.49**	**1.06;2.10**
Total duration of hospital admission(s) (days)	0.92	0.78;1.08	0.80	0.61;1.05
Readmitted	0.54	0.17;1.76	0.24	0.05;1.12
Admission after the first trimester	0.97	0.23;3.98	0.68	0.11;4.13

P-values<0.05 are considered significant and marked in bold. Weight change during index pregnancy is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and > 0 if women gained weight. **Logistic regression analysis:** we were only able to perform multivariable logistic regression analysis for HADS and PCL-5 score at follow-up. **Model 1**: univariable regression analysis. **Model 2**: multivariable regression analysis: *adjusted for history of any traumatic event, **adjusted for HADS at inclusion during MOTHER study. **Abbreviations:** HADS: Hospital Anxiety and Depression Scale. A higher HADS indicates more severe depression or anxiety symptoms whereas a depression or anxiety score ≥8 is considered borderline/abnormal. HG: hyperemesis gravidarum, PCL-5: post-traumatic stress disease (PTSD) checklist for the DSM-5. A higher PCL-5 score indicates more severe PTSD symptoms whereas a PCL-5 score ≥31 indicates PTSD, PUQE: 24-hour Pregnancy Unique Quantification of Emesis and nausea. A higher PUQE-24 score indicates more severe symptoms.

HG severity of the index pregnancy and PTSD symptoms

Multivariable logistic regression analysis showed that for every additional point of the mean PUQE-24 score in the first three weeks after inclusion in the index pregnancy, the likelihood

of having a PCL-5 score ≥31 increased with OR 1.49 (95% CI: 1.06;2.10).**(Table 2)** Higher PUQE-24 scores in the first three weeks after inclusion in the index pregnancy were also associated with a higher PCL-5 score in multivariable linear regression analysis as shown in **Supplement Table S3** (β 16.65, 95% CI: 5.34;29.05). None of the other measures of HG severity of the index pregnancy were associated with the PCL-5 score during follow-up.

PTSD symptoms according to RCT treatment allocation

We found higher PCL-5 scores at follow-up among those receiving enteral tube feeding compared to those receiving standard care (27.0 (18.0-43.0) vs. 12.0 (3.5-27.5), *P*=0.046) in the as treated analysis.(**Supplement Table S4-6)** There was no difference in the likelihood that women in the early enteral tube feeding group had PCL-5 scores ≥31 compared to women in the standard care group.

DISCUSSION

Main findings

We found that, on average 4.5 years after having been admitted for HG, women were commonly affected by depression, anxiety and PTSD symptoms. Depression (20%) and anxiety rates (24%) at six weeks postpartum and at follow-up (resp. 27% and 40%) were considerably higher than postpartum depression and anxiety rates reported in the general population (resp. 0.8-2.6% and 17%).[24, 25] Furthermore, we found that 22% of the women included in our study had PCL-5 scores indicative of probable PTSD, in line with earlier reports,[6] and again, considerably higher than PTSD rates reported in the general postpartum population (0.3-4.0%).[24, 26] Importantly, our study suggests that higher vomiting scores in the index pregnancy were associated with an increased chance of meeting the diagnostic criteria for anxiety disorder and PTSD at follow-up.

Strengths and limitations

Prospectively collected detailed information of the index pregnancy was one of the strengths of this study. The fact that we collected information on whether HG pregnancies had occurred between MOTHER and follow-up study participation was another strength. We used validated questionnaires to assess symptom severity during index pregnancy and to evaluate depression, anxiety and PTSD symptoms. Additionally, depression and anxiety symptoms were measured at three different moments.

The main limitation of this study was that it is a small sample study and therefore could have lacked sufficient power to detect differences. Only 34% (73/215) of the MOTHER participants

completed the follow-up questionnaire, which made our study prone to selection bias and may hamper generalizability of our findings. Sensitivity analysis revealed that participants of the follow-up study had higher PUQE-24 scores in the index pregnancy than women who were lost to follow-up. Selective participation of more severely affected participants could partially explain the high depression, anxiety or PTSD rates in our study, but cannot explain the dose-effect association between increased PUQE-24 scores and increased mental health symptoms. The time-interval between participation of the MOTHER and follow-up study, which could have taken up to 6 years, could also have affected our results. Depression, anxiety and PTSD symptoms may have improved over time, for example by receiving treatment, and therefore have led to an underestimation of symptoms.[27, 28] Since we performed multiple statistical comparisons, it could be that some of our findings were due to chance. Unfortunately, no information on possible treatments for mental health disorders were collected in this follow-up study. Finally, we were not able to compare depression, anxiety and PTSD rates between women with and without HG.

Interpretation

Depression rates in our study, at both six weeks postpartum and on average 4.5 years later, lie between the previously reported depression rates of 12% at six and twelve months postpartum by Kjeldgaard *et al.*[5] and 29% at six weeks postpartum by Mitchell-Jones *et al.*[29] These differences might be explained by the fact that Kjeldgaard *et al.*[5] included HG patients based on ICD codes instead of a clinical HG diagnosis, despite the fact that ICD codes have been demonstrated not to be reliably identify HG patients.[30] Misclassification of the disease could have led to underreporting of HG patients and in line depression rates. Conversely, due to self-selection participation in Mitchell-Jones *et al.*[29] and due to selective participation in our study, it could be that these study populations consisted of a more severe group of HG patients, which may have led to higher reported depression rates.

Two studies have reported that prolonged HG, persisting in or beyond the second trimester, increases the chance of depression symptoms postpartum, which is at odds with our study, since we did not find associations between measures of HG severity of the index pregnancy and depression symptoms.[4, 5] Since these two studies had larger study populations (respectively 4.308 and 92.947 women), it could be that our study was simply too small to detect any significant associations.

Twenty two percent of the included women in our study had PCL-5 scores ≥31, indicative of probable PTSD, which is similar to the previously reported 18% of Christodoulou-Smith

et al.[6] Their study did not specify which PTSD questionnaire was used and participants were retrieved through advertisement on a HG patient support group website with establishing HG pregnancies through self-reports, which could have led to an overestimation of PTSD rates. Both studies however provide evidence that PTSD symptoms are common in women who suffered from HG. There are several effective treatments available for PTSD, also during pregnancy.[31]

A more recent published study from Kjeldgaard *et al.*[11] evaluated PTSD symptoms in the Norwegian Mother and Child Cohort by use of the Impact of Event Score, and reported that women with HG had higher PTSD scores at 8 weeks postpartum compared to women with no, mild or severe nausea; an association that remained after adjusting for having a depression or anxiety disorder in their medical history. These findings are similar to ours. We found that higher vomiting scores were associated with higher PTSD and anxiety symptoms at follow-up. Unfortunately, our study design hampers our ability to draw any firm conclusions about the causal direction of this association. However, it is important to highlight the fact that our study, together with previous published studies, found a dose-effect response between HG symptom severity and depression, anxiety and PTSD symptoms. Together with the fact that only 16% of included patients in this study had a medical history of a mental health disease, these findings support the notion that HG itself leads to depression, anxiety or PTSD symptoms, instead of women with a pre-existent psychiatric illness being more predisposed to develop HG. This is consistent with qualitative studies stating that the burden of HG leads to developing psychological symptoms instead of being the cause of the disease.[32] This notion is further supported by the fact that in our study an additional pregnancy affected by HG between the index pregnancy and follow-up participation more than doubled the odds of having an anxiety disorder at follow-up, a further suggestion of a dose-response effect with increased HG burden negatively impacting future mental health.

Conclusion

Taken together, our study confirms that anxiety, depression and PTSD symptoms are common in women previously admitted for HG. Moreover, our study suggests that an increased burden of HG, either as evident from higher symptom scores or higher total number of HG affected pregnancies, are at an increased risk of developing an anxiety disorder or PTSD. Future studies should confirm whether better treatment of HG can prevent or improve depression, anxiety and PTSD symptoms.

Acknowledgements

We thank all participating women in both the MOTHER and follow-up study.

Funding

The MOTHER study was supported by a research grant from the North West Hospital Group, Alkmaar, The Netherlands under grant number 2013T085. The follow-up study was supported by the Amsterdam Reproduction and Development (AR&D) research institute, Amsterdam UMC, The Netherlands under project number 23346. Neither funders had any role in the planning, execution or interpretation of this study.

Contribution to authorsip

IJG, TJR and RCP conceived and conducted the MOTHER study. KN and RCP conceived and conducted the follow-up study. LMvdM helped develop the online survey tool. JMJB, CR-S, HAB, DPvdH, WMH, AH, GK, SK, JOEHvL, JL, FvdM, DP, M-JP, PJP, LvRF, RJR, HCJS, TV, BWM, MHK, IJG and RCP recruited participants for the original MOTHER study and collected data. KN performed the statistical analyses, under supervision of RCP and RvE, and drafted the manuscript. CD is a patient representative and gave perspective on the interpretation of the results. All authors (LvdM, CD, JMJB, CR-S, RvE, HAB, DPvdH, WMH, AH, GK, SK, JOEHvL, JL, FvdM, DP, M-JP, PJP, LvRF, RJR, HCJS, TV, BWM, MHK, IJG, RCP and MvO) critically reviewed and approved the final draft of the manuscript.

REFERENCES

1. Gazmararian JA, Petersen R, Jamieson DJ, Schild L, Adams MM, Deshpande AD, et al. Hospitalizations during pregnancy among managed care enrollees. Obstetrics and gynecology. 2002;100(1):94-100.
2. Mitchell-Jones N, Gallos I, Farren J, Tobias A, Bottomley C, Bourne T. Psychological morbidity associated with hyperemesis gravidarum: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2017;124(1):20-30.
3. Kjeldgaard HK, Eberhard-Gran M, Benth J, Nordeng H, Vikanes Å V. History of depression and risk of hyperemesis gravidarum: a population-based cohort study. Archives of women's mental health. 2017;20(3):397-404.
4. Iliadis SI, Axfors C, Johansson S, Skalkidou A, Mulic-Lutvica A. Women with prolonged nausea in pregnancy have increased risk for depressive symptoms postpartum. Scientific reports. 2018;8(1):15796.
5. Kjeldgaard HK, Eberhard-Gran M, Benth JS, Vikanes AV. Hyperemesis gravidarum and the risk of emotional distress during and after pregnancy. Archives of women's mental health. 2017;20(6):747-56.
6. Christodoulou-Smith J, Gold JI, Romero R, Goodwin TM, Macgibbon KW, Mullin PM, et al. Posttraumatic stress symptoms following pregnancy complicated by hyperemesis gravidarum. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet. 2011;24(11):1307-11.
7. McCarthy FP, Khashan AS, North RA, Moss-Morris R, Baker PN, Dekker G, et al. A Prospective Cohort Study Investigating Associations between Hyperemesis Gravidarum and Cognitive, Behavioural and Emotional Well-Being in Pregnancy. PloS one. 2011;6(11):e27678.
8. Annagür BB, Tazegül A, Gündüz S. Do psychiatric disorders continue during pregnancy in women with hyperemesis gravidarum: a prospective study. Gen Hosp Psychiatry. 2013;35(5):492-6.
9. Tan PC, Zaidi SN, Azmi N, Omar SZ, Khong SY. Depression, anxiety, stress and hyperemesis gravidarum: temporal and case controlled correlates. PloS one. 2014;9(3):e92036.
10. Senturk MB, Yildiz G, Yildiz P, Yorguner N, Cakmak Y. The relationship between hyperemesis gravidarum and maternal psychiatric well-being during and after pregnancy: controlled study. The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet. 2017;30(11):1314-9.
11. Kjeldgaard HK, Vikanes A, Benth JS, Junge C, Garthus-Niegel S, Eberhard-Gran M. The association between the degree of nausea in pregnancy and subsequent posttraumatic stress. Archives of women's mental health. 2019;22(4):493-501.
12. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient–clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ Open. 2021;11(1):e041254.
13. Grooten IJ, Koot MH, van der Post JA, Bais JM, Ris-Stalpers C, Naaktgeboren C, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. The American journal of clinical nutrition. 2017;106(3):812-20.
14. Grooten IJ, Mol BW, van der Post JA, Ris-Stalpers C, Kok M, Bais JM, et al. Early nasogastric tube feeding in optimising treatment for hyperemesis gravidarum: the MOTHER randomised controlled trial (Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding). BMC pregnancy and childbirth. 2016;16:22.

15. Nijsten K, Koot MH, van der Post JAM, Bais JMJ, Ris-Stalpers C, Naaktgeboren C, et al. Thyroid-stimulating hormone and free thyroxine fail to predict the severity and clinical course of hyperemesis gravidarum: A prospective cohort study. Acta obstetricia et gynecologica Scandinavica. 2021;100(8):1419-29.

16. Koren G, Piwko C, Ahn E, Boskovic R, Maltepe C, Einarson A, et al. Validation studies of the Pregnancy Unique-Quantification of Emesis (PUQE) scores. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2005;25(3):241-4.

17. Nijsten K, Dean C, van der Minnen LM, Bais JMJ, Ris-Stalpers C, van Eekelen R, et al. Recurrence, postponing pregnancy, and termination rates after hyperemesis gravidarum: Follow up of the MOTHER study. Acta obstetricia et gynecologica Scandinavica. 2021;100(9):1636-43.

18. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta psychiatrica Scandinavica. 1983;67(6):361-70.

19. Weathers FW LB, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD Checklist for DSM-5 (PCL-5). 2018 [Available from: https://www.ptsd.va.gov/professional/assessment/documents/PCL5_Standard_form.PDF.

20. van Herpen MM, Boeschoten MA, te Brake H, van der Aa N, Olff M. Mobile Insight in Risk, Resilience, and Online Referral (MIRROR): Psychometric Evaluation of an Online Self-Help Test. J Med Internet Res. 2020;22(9):e19716.

21. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res. 2002;52(2):69-77.

22. Bovin MJ, Marx BP, Weathers FW, Gallagher MW, Rodriguez P, Schnurr PP, et al. Psychometric properties of the PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (PCL-5) in veterans. Psychol Assess. 2016;28(11):1379-91.

23. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373-9.

24. Fawcett EJ, Fairbrother N, Cox ML, White IR, Fawcett JM. The Prevalence of Anxiety Disorders During Pregnancy and the Postpartum Period: A Multivariate Bayesian Meta-Analysis. J Clin Psychiatry. 2019;80(4).

25. Shorey S, Chee CYI, Ng ED, Chan YH, Tam WWS, Chong YS. Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. J Psychiatr Res. 2018;104:235-48.

26. Yildiz PD, Ayers S, Phillips L. The prevalence of posttraumatic stress disorder in pregnancy and after birth: A systematic review and meta-analysis. J Affect Disord. 2017;208:634-45.

27. Santiago PN, Ursano RJ, Gray CL, Pynoos RS, Spiegel D, Lewis-Fernandez R, et al. A systematic review of PTSD prevalence and trajectories in DSM-5 defined trauma exposed populations: intentional and non-intentional traumatic events. PloS one. 2013;8(4):e59236-e.

28. Stegenga BT, Kamphuis MH, King M, Nazareth I, Geerlings MI. The natural course and outcome of major depressive disorder in primary care: the PREDICT-NL study. Soc Psychiatry Psychiatr Epidemiol. 2012;47(1):87-95.

29. Mitchell-Jones N, Lawson K, Bobdiwala S, Farren JA, Tobias A, Bourne T, et al. Association between hyperemesis gravidarum and psychological symptoms, psychosocial outcomes and infant bonding: a two-point prospective case–control multicentre survey study in an inner city setting. BMJ Open. 2020;10(10):e039715.

30. Vikanes Å, Magnus P, Vangen S, Lomsdal S, Grjibovski AM. Hyperemesis gravidarum in the Medical Birth Registry of Norway – a validity study. BMC pregnancy and childbirth. 2012;12(1):115.

31. Bisson JI, Olff M. Prevention and treatment of PTSD: the current evidence base. European Journal of Psychotraumatology. 2021;12(1):1824381.

32. Dean C, Bannigan K, Marsden J. Reviewing the effect of hyperemesis gravidarum on women's lives and mental health. British Journal of Midwifery. 2018;26(2):109-19.

Appendix S1. Follow-up questionnaire

Part A: Subsequent pregnancies after participating in the MOTHER study

1. After the pregnancy in which you participated in the MOTHER study, were you pregnant again? (A miscarriage, ectopic pregnancy or preterm birth also count!)
    a. Yes, continue with question 3.
    b. No, continue with question 22.

2. How many times have you been pregnant since participating in the MOTHER study?
    a. ….. *(fill in a number)*

If question 2 is answered with 2 or more then the following questions will be answered for <u>every</u> pregnancy for a maximum of 5 pregnancies

3. In which year was this pregnancy?
    a. ….. *(For example '2015')*
4. Did you postpone this pregnancy because of the severity of the nausea and vomiting symptoms in your previous pregnancy?
    a. Yes
    b. No
5. Did this pregnancy end in a miscarriage or was this an ectopic pregnancy?
    a. No
    b. Yes, this pregnancy ended in a miscarriage
    c. Yes, this was an ectopic pregnancy
6. Did you experience any symptoms of nausea in this pregnancy?
    a. Yes, continue with question 7
    b. No, not at all. Continue with question 23 or go to '**adding a new pregnancy'**
7. How many weeks were you pregnant when you first felt nauseous?
    a. ….. *(in weeks)*
8. How many weeks were you pregnant when you first started vomiting?
    a. ….. *(in weeks)*
    b. I did not have complains of vomiting
9. How many weeks were you pregnant when the nausea and vomiting symptoms had practically disappeared?
    a. ….. *(in weeks)*

10. Were you admitted in the hospital with severe nausea and vomiting in this pregnancy?
    a. Yes, continue with question 11
    b. No, continue with question 13
11. How many times were you admitted in the hospital in this pregnancy?
    a. Once
    b. Twice
    c. 3 times
    d. More than 3 times, namely …. Times
12. How many days were you in total admitted in the hospital in this pregnancy? (the day of admission and the day of discharge both count as 1 day)
    a. ….. (Answer in days)
    b. I don't remember
13. Did you use any medication for the nausea and vomiting symptoms in this pregnancy?
    a. Yes, continue with question 14
    b. No, continue with question 15
14. Which medication did you use? (multiple options possible)
    a. Suprimal
    b. Emesafene
    c. Primperan (Metoclopramide)
    d. Zofran (Ondansetron)
    e. Potassium solution (potassium drink ('kaliumdrank') or potassium intravenous)
    f. Corticosteroids (methylprednisolon or hydrocortison)
    g. Omeprazole or Ranitidine
    h. Other, namely …. (free text)
    i. I don't remember which medication I used
15. Did you receive nasogastric tube feeding in this pregnancy?
    a. Yes
    b. No
16. Was this a singleton or a multiple pregnancy?
    a. Singleton pregnancy
    b. Twin pregnancy
    c. Multiple pregnancy of three or more babies (triplets or quadruplets)?
17. Did the severity of nausea and vomiting affect your ability to work?
    a. No, I was able to go to my work and did not have to call in sick at all
    b. Partly: I was not able to go to my work for some days
    c. Partly: I was not able to work for prolonged periods (eg weeks)

   d. I was not able to work at all
   e. I did not had a job at the time

18. Did the severity of nausea and vomiting affect your everyday life?
   a. Yes, the nausea and vomiting had an enormous effect on my everyday life
   b. The nausea and vomiting affected my everyday life to some degree
   c. No, The nausea and vomiting symptoms did not affect my everyday life at all
19. What was your weight before this pregnancy?
   a. …. (in kilograms)
   b. I don't remember
20. What was your lowest weight during this pregnancy?
   a. …. (in kilograms)
   b. I don't remember
21. Did you consider terminating this pregnancy because of the severity of the nausea and vomiting symptoms in this pregnancy or your previous pregnancy?
   a. Yes, I terminated this pregnancy because of the severity of nausea and vomiting
   b. I considered terminating this pregnancy, but in the end I continued this pregnancy
   c. No, I did not consider terminating this pregnancy
   d. Yes, I terminated this pregnancy, but due to other reasons than HG (f.e. congenital abnormalities or unwanted pregnancy)

- ***If answered 'no' to question 1 (and thus not have become pregnant again after participating in the MOTHER Study), continue with question 22***

- ***If answered 'yes' to question 1 (and thus finished question 18), continue with question 23***

22. After the pregnancy in which you participated in the MOTHER Study, did you're not becoming pregnant again have to do with the severity of the nausea and vomiting during that pregnancy, or fear of having hyperemesis gravidarum again?
   a. Yes
   b. No, there were other reasons
23. Do you have family members who also had severe nausea and vomiting or hyperemesis gravidarum in pregnancy?
   a. No
   b. Yes: (multiple options possible)

i. Mother
ii. Aunt
iii. Sister
iv. Grandmother

24. Do you have migraines?
    a. Yes
    b. No
25. Do you get motion sickness (eg car sick)?
    a. Yes
    b. No
26. What is your current height?
    a. ….. (in centimeters)
27. What is your current weight?
    b. ….. (in kilograms)

Part B: Depression and anxiety symptoms (HADS questionnaire)

Emotions play an important part in most illnesses. This questionnaire is designed to find out how you feel. Read each item below and tick the answer that comes closest to how you have been feeling in the past week.

28. I feel tense or 'wound up':
    a. Most of the time
    b. A lot of the time
    c. From time to time, occasionally
    d. Not at all
29. I still enjoy the things I used to enjoy:
    a. Definitely as much
    b. Not quite so much
    c. Only a little
    d. Hardly at all
30. I get a sort of frightened feeling as if something awful is about to happen:
    a. Very definitely and quite badly
    b. Yes, but not too badly
    c. A little, but it doesn't worry me
    d. Not at all

31. I can laugh and see the funny side of things:
    a. As much as I always could
    b. Not quite so much now
    c. Definitely not so much now
    d. Not at all
32. Worrying thoughts go through my mind:
    a. A great deal of the time
    b. A lot of the time
    c. From time to time, but not too often
    d. Only occasionally
33. I feel cheerful:
    a. Not at all
    b. Not often
    c. Sometimes
    d. Most of the time
34. I can sit at ease and feel relaxed:
    a. Definitely
    b. Usually
    c. Not often
    d. Not at all
35. I feel as if I am slowed down:
    a. Nearly all the time
    b. Very often
    c. Sometimes
    d. Not at all
36. I get a sort of frightened feeling like 'butterflies' in the stomach:
    a. Not at all
    b. Occasionally
    c. Quite often
    d. Very often
37. I have lost interest in my appearance:
    a. Definitely
    b. I don't take as much care as I should
    c. I may not take quite as much care
    d. I take just as much care as ever
38. I feel restless as I have to be on the move

a. Very much indeed
b. Quite a lot
c. Not very much
d. Not at all

39. I look forward with enjoyment to things:
    a. As much as I ever did
    b. Rather less than I used to
    c. Definitely less than I used to
    d. Hardly at all
40. I get sudden feelings of panic
    a. Very often indeed
    b. Quite often
    c. Not very often
    d. Not at all
41. I can enjoy a good book or radio or TV program:
    a. Often
    b. Sometimes
    c. Not often
    d. Very seldom

Part C: Post-traumatic stress symptoms (PCL-5 questionnaire)

Below is a list of problems that people sometimes have in response to a very stressful experience. For the next questions, keep your pregnancy complicated by hyperemesis gravidarum in mind, please read each problem carefully and then circle one of the numbers to the right to indicate how much you have been bothered by that problem in the past month.

42. Repeated, disturbing, and unwanted memories of the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
43. Repeated, disturbing dreams of the stressful experience?
    a. Not at all
    b. A little bit
    c. Moderately

   d. Quite a bit
   e. Extremely

44. Suddenly feeling or acting as if the stressful experience were actually happening again (as if you were actually back there reliving it)?
   a. Not at all
   b. A little bit
   c. Moderately
   d. Quite a bit
   e. Extremely
45. Feeling very upset when something reminded you of the stressful experience?
   a. Not at all
   b. A little bit
   c. Moderately
   d. Quite a bit
   e. Extremely
46. Having strong physical reactions when something reminded you of the stressful experience (for example, heart pounding, trouble breathing, sweating)?
   a. Not at all
   b. A little bit
   c. Moderately
   d. Quite a bit
   e. Extremely
47. Avoiding memories, thoughts, or feelings related to the stressful experience?
   a. Not at all
   b. A little bit
   c. Moderately
   d. Quite a bit
   e. Extremely
48. Avoiding external reminders of the stressful experience (for example, people, places, conversations, activities, objects, or situations)?
   a. Not at all
   b. A little bit
   c. Moderately
   d. Quite a bit
   e. Extremely
49. Trouble remembering important parts of the stressful experience?

a. Not at all
b. A little bit
c. Moderately
d. Quite a bit
e. Extremely

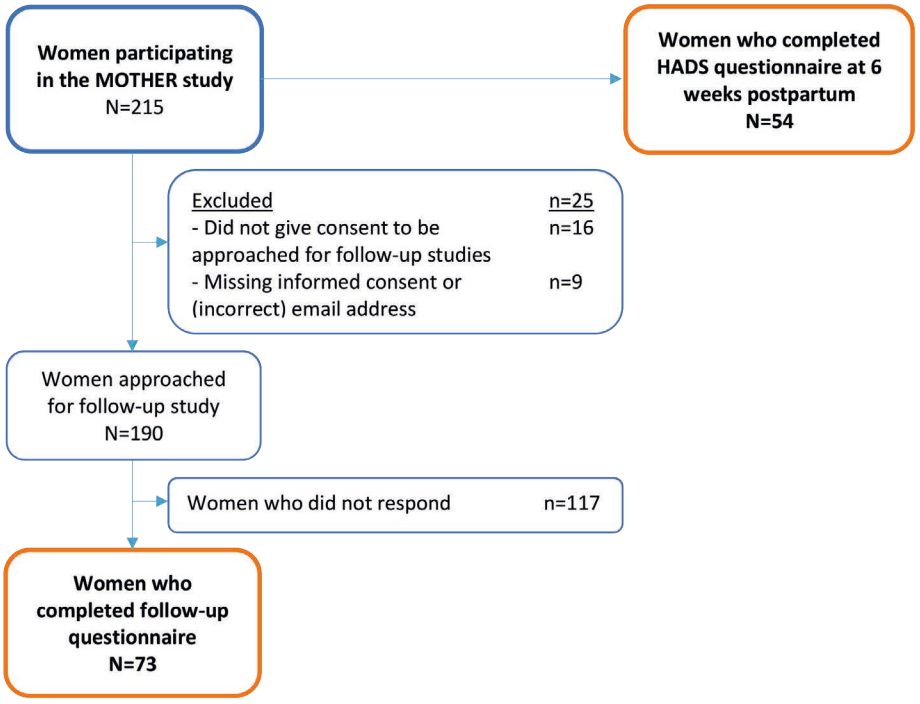
50. Having strong negative beliefs about yourself, other people, or the world (for example, having thoughts such as: I am bad, there is something seriously wrong with me, no one can be trusted, the world is completely dangerous)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
51. Blaming yourself or someone else for the stressful experience or what happened after it?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
52. Having strong negative feelings such as fear, horror, anger, guilt, or shame?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
53. Loss of interest in activities that you used to enjoy?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
54. Feeling distant or cut off from other people?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit

e. Extremely

55. Trouble experiencing positive feelings (for example, being unable to feel happiness or have loving feelings for people close to you)?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
56. Irritable behavior, angry outbursts, or acting aggressively?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
57. Taking too many risks or doing things that could cause you harm?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
58. Being "super alert" or watchful or on guard?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
59. Feeling jumpy or easily startled?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
60. Having difficulty concentrating?
    a. Not at all
    b. A little bit
    c. Moderately

d. Quite a bit
e. Extremely

61. Trouble falling or staying asleep?
    a. Not at all
    b. A little bit
    c. Moderately
    d. Quite a bit
    e. Extremely
62. Have u experienced another very stressful or traumatic event?
    a. Yes, continue with question 63
    b. No, continue with question 65
63. In which year did this very stressful or traumatic event happened?
    a. ….. (year)
64. What kind of very stressful or traumatic event did you experienced?
    a. Sexual assault
    b. Physical assault, violence or abuse
    c. Seeing someone be killed or seriously injured
    d. Dying of a loved one
    e. War
    f. Other, namely ….. (free text)
65. Do you wish to be informed about the results of this follow up of the MOTHER study?
    a. Yes
    b. No



Supplement Figure 1. Flowchart in- and exclusions follow-up study

Supplement Table S1. Sensitivity analysis between women included in the follow-up study and women who were lost to follow-up

	Included in follow-up N=73	**Not included in follow-up** N=142	***P*-value**
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years)	29.25 ± 4.62	28.62 ± 4.94	0.37
Ethnicity			**0.01**
- Western	52 (71.2%)	71 (50.0%)	
- Non-western	12 (16.4%)	41 (28.9%)	
Education level			**0.02**
- Primary or secondary	27 (37.0%)	60 (42.3%)	
- Higher	29 (39.7%)	28 (19.7%)	
Primigravida at the time	26 (35.6%)	38 (26.8%)	0.18
History of mental health disease [a]	12 (16.4%)	29 (20.4%)	0.48
HG in previous pregnancy [b]	23 (48.9%)	45 (43.3%)	0.52
HG in previous pregnancy requiring hospital admission[2]	12 (25.5%)	25 (24.0%)	0.84
Maternal outcomes			
Weight change (kg) [c]	-3.40 ± 3.85	-2.67 ± 4.17	0.22
PUQE-24 at inclusion	11.00 (9.00-13.00)	9.00 (7.00-12.00)	**0.01**
Mean PUQE-24 in the first 3 weeks after admission	9.19 ± 2.50	8.83 ± 2.93	0.40
Total duration of hospital admissions (days)	5.00 (4.00-8.00)	5.00 (3.00-8.00)	0.63
Readmitted	24 (32.9%)	47 (33.1%)	0.97
Admission after the first trimester	14 (19.2%)	29 (20.4%)	0.83
HADS at inclusion	21.55 ± 6.49	19.85 ± 7.56	0.18
- Anxiety score ≥8 [d]	31 (50.8%)	48 (49.5%)	0.87
- Depression score ≥8 [d]	56 (91.8%)	84 (86.6%)	0.32

Significant p-values <0.05 are marked in bold. Data represented with mean±SD**,** median (IQR) or frequency (%). [a] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [b] Percentage shown is frequency divided by multigravidas. [c] Weight change is weight at baseline minus prepregnancy weight: can be < 0 if women lost weight and can be > 0 if women gained weight. [d] Percentage shown is frequency divided by total HADS at inclusion available (included women: n=61, excluded women: n=97). **Abbreviations:** HG: hyperemesis gravidarum, HADS: hospital anxiety and depression scale: a higher HADS indicates more severe anxiety or depression symptoms, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms.

Supplement Table S2. Univariable logistic and linear regression to assess the association between possible risk factors and HADS, at 6 weeks postpartum and at follow-up, and PCL-5 score at follow-up

Univariable logistic regression	**HADS 6 weeks postpartum: Anxiety score ≥8**		**HADS 6 weeks postpartum: Depression score ≥8**		**HADS at follow-up: Anxiety score ≥8**		**HADS at follow-up: Depression score ≥8**		**PCL-5 score at follow-up ≥31**	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
-Maternal age	0.86*	0.74;1.00	0.90	0.78;1.05	0.91*	0.82;1.01	0.84^	0.73;0.96	0.91	0.80;1.03
-Educational attainment [a]	3.81*	0.86;16.94	3.94	0.70;22.15	0.96	0.33;2.84	3.30*	0.96;11.31	3.13*	0.83;11.75
-Ethnicity (western or not)	1.65	0.26;10.38	6.17*	1.00;37.98	0.46	0.11;1.88	1.36	0.35;5.22	2.10	0.53;8.38
-History of mental health disease [b]	4.11*	0.86;19.68	0.51	0.06;4.69	0.72	0.20;2.65	2.19	0.61;7.94	2.04	0.53;7.92
-HADS at inclusion MOTHER study	1.06	0.96;1.16	1.10	0.98;1.23	1.09*	1.00;1.19	1.14^	1.02;1.27	1.20^	1.04;1.39
-History of any traumatic event	-	-	-	-	3.63^	1.29;10.20	3.08^	1.05;9.03	1.99	0.64;6.26
-Had another HG pregnancy between MOTHER and follow-up participation[c]	-	-	-	-	2.74^	1.04;7.20	1.41	0.50;3.96	1.00	0.33;3.04

Univariable linear regression	**Total HADS 6 weeks postpartum**		**Total HADS at follow-up ∫**		**PCL-5 score at follow-up ∫**	
	β	**95% CI**	***β***	**95% CI**	***β***	**95% CI**
-Maternal age	-0.38^	-0.73;-0.02	-4.11^	-7.41;-0.70	-2.66	-8.06;3.15
-Educational attainment [a]	4.38^	0.65;8.11	20.93	-16.81;75.77	55.12	-15.04;182.92
-Ethnicity (western or not)	6.16^	0.82;11.50	-1.78	-36.43;51.59	-11.93	-57.13;80.76
-History of mental health disease [b]	3.29	-1.47;8.05	-6.39	-40.19;46.52	3.67	-49.39;112.55
-HADS at inclusion MOTHER study	0.27^	0.07;0.48	3.67^	1.21;6.08	5.65^	1.31;10.08
-History of any traumatic event	-	-	55.12^	10.08;118.37	71.94*	-0.50;197.13
-Had another HG pregnancy between MOTHER and follow-up participation[c]	-	-	45.94^	5.65;101.58	4.08	-38.37;75.77

* *P*-value <0.1; ^ *P*-value <0.05. OR is the odds ratio. *β* unstandardized regression coefficient. 95% CI: 95% confidence interval. ∫ log transformed, back transformed and expressed as percentages of differences. [a] Educational attainment: finished primary or secondary school or finished higher education. [b] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [c] HG defined as: if vomiting symptoms occurred with either multiple medication use, weight loss, hospital admission for HG, requiring tube feeding or whether nausea and vomiting symptoms affected their life and/or work. Abbreviations: HADS: Hospital Anxiety and Depression Scale. A higher HADS indicates more severe depression or anxiety symptoms whereas a depression or anxiety score ≥8 is considered borderline/abnormal. HG: hyperemesis gravidarum, PCL-5: post-traumatic stress disease (PTSD) checklist for the DSM-5. A higher PCL-5 score indicates more severe PTSD symptoms whereas a PCL-5 score ≥31 indicates PTSD

Supplement Table S3. Linear regression analysis to assess the association between the severity of HG in the index pregnancy and the HADS 6 weeks postpartum and HADS and PCL-5 score at follow-up study.

	HADS 6 weeks postpartum				HADS at follow-up study ∫				PCL-5 score at follow-up study ∫			
	Model 1		Model 2*		Model 1		Model 2**		Model 1		Model 2***	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI	β	95% CI
Weight change (kg)	-0.06	-0.48;0.36	0.00	-0.40;0.40	-0.80	-4.97;3.67	-0.80	-4.40;3.05	0.00	-6.57;7.04	1.11	-5.73;8.33
PUQE-24 at inclusion	-0.59	-1.28;0.11	-0.24	-0.97;0.49	-1.59	-7.13;4.29	1.21	-4.21;6.93	-2.18	-10.95;7.47	2.22	-7.32;12.86
Mean PUQE-24 in the first 3 weeks after admission	-0.33	-0.93;0.28	-0.16	-0.78;0.45	5.65	-1.29;13.09	5.23	-0.50;11.40	9.64	-1.49;22.02	**16.65**	**5.34;29.05**
Total duration of hospital admissions (days)	0.13	-0.28;0.54	0.07	-0.38;0.52	-0.50	-2.57;1.61	-0.70	-2.47;1.01	-0.50	-31.61;2.74	-0.50	-3.63;2.84
Readmitted	0.41	-3.09;3.91	0.44	-3.00;3.88	-6.95	-33.50;30.08	-18.13	-38.61;9.20	-28.61	-57.64;20.20	-34.95	-61.67;10.41
Admission after the first trimester	1.73	-2.17;5.63	0.05	-3.61;3.70	29.05	-15.04;96.01	2.33	-28.75;46.96	-16.47	-56.48;60.32	-30.02	-63.94;35.66

P-values<0.05 are considered significant and marked in bold. β is the unstandardized regression coefficient. 95% CI: 95% confidence interval. ∫ outcome variables are log transformed, back transformed and expressed as percentages of differences. Weight change during index pregnancy is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and > 0 if women gained weight. **Abbreviations:** HADS: Hospital Anxiety and Depression Scale. A higher HADS indicates more severe anxiety or depression symptoms. HG: Hyperemesis gravidarum, PCL-5: post-traumatic stress disease (PTSD) checklist for the DSM-5. A higher PCL-5 score indicates more severe PTSD symptoms whereas a PCL-5 score ≥31 indicates PTSD. PUQE: 24-hour Pregnancy Unique Quantification of Emesis and nausea. A higher PUQE-24 score indicates more severe symptoms.

Adjustments were made based on univariable linear regression analysis: risk factors with a p-value <0.10 in as shown in Table 2 were included as confounder in multivariable regression model.

* adjusted for maternal age, educational attainment (primary/secondary school or higher education), ethnicity (western or not) and HADS at inclusion during MOTHER study.

** adjusted for maternal age, HADS at inclusion during MOTHER study, history of any traumatic event and having another HG pregnancy between participation of the MOTHER and follow-up study.

*** adjusted for HADS at inclusion during MOTHER study and history of any traumatic event.

Supplement Table S4. RCT analysis according to intention-to-treat principle

In total: N=50	Early enteral tube feeding N=25	Standard care N=25	*P*-value
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years)	28.88 ± 5.04	30.00 ± 4.38	0.41
Ethnicity			0.08
- Western	20 (80.0%)	15 (60.0%)	
- Non-western	3 (12.0%)	8 (32.0%)	
Education level			0.64
- Primary or secondary	9 (36.0%)	12 (48.0%)	
- Higher	11 (44.0%)	11 (44.0%)	
Primigravida at the time	11 (44.0%)	5 (20.0%)	0.07
History of mental health disease [a]	4 (16.0%)	2 (8.0%)	0.67
HG in previous pregnancy prior to MOTHER study [b]	7 (50.0%)	12 (60.0%)	0.56
Maternal outcomes			
Weight change (kg) [c]	-2.76 ± 4.09	-4.56 ± 4.34	0.14
PUQE-24 at inclusion	12.00 (10.50-14.00)	11.50 (9.75-13.00)	0.31
Mean PUQE-24 in the first 3 weeks after admission	9.46 ± 2.89	9.59 ± 2.22	0.86
Total duration of hospital admissions (days)	6.00 (3.00-9.00)	5.00 (3.50-9.50)	0.89
Readmitted	9 (36.0%)	8 (32.0%)	0.77
Admission after the first trimester	7 (28.0%)	3 (12.0%)	0.16
HADS at inclusion	21.20 ± 6.98	21.77 ± 5.91	0.78
- Anxiety score ≥8 [d]	10 (40.0%)	12 (48.0%)	0.77
- Depression score ≥8 [d]	19 (76.0%)	20 (80.0%)	1.00
Follow-up study			
Depression, anxiety and PTSD symptoms			
HADS 6 weeks postpartum [e]	9.61 ± 6.49	9.68 ± 6.18	0.97
- Anxiety score ≥8 [e]	6 (21.4%)	7 (28.0%)	0.58
- Depression score ≥8 [e]	6 (21.4%)	5 (20.0%)	0.90
HADS at follow-up study	10.00 (6.00-15.00)	8.00 (5.00-15.00)	0.62
- Anxiety score ≥8	9 (36.0%)	7 (28.0%)	0.54
- Depression score ≥8	7 (28.0%)	7 (28.0%)	1.00
PCL-5 score at follow-up study	18.00 (8.50-33.50)	13.00 (5.50-30.00)	0.52
- PCL-5 score ≥31	6 (24.0%)	6 (24.0%)	1.00
History of any traumatic event	7 (28.0%)	6 (24.0%)	0.75
History of an obstetric traumatic event	0 (0.0%)	1 (4.0%)	0.46
Women who experienced a subsequent pregnancy after MOTHER participation	15 (60.0%)	8 (32.0%)	0.06
Suffered from HG again between participation of the MOTHER and follow-up study [f]	14 (56.0%)	10 (40.0%)	0.26

Data represented with mean±SD**,** median (IQR) or frequency (%). [a] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [b] Percentage shown is frequency divided by multigravidas. [c] Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. [d] Percentage shown is frequency divided by number of HADS at inclusion available (n=61). [e] Different study group including 42 women. [f] HG defined as: if vomiting symptoms occurred with either multiple medication use, weight loss, hospital admission for HG, requiring tube feeding or whether nausea and vomiting symptoms affected their life and/or work. **Abbreviations:** HG: hyperemesis gravidarum, HADS: hospital anxiety and depression scale: a higher HADS indicates more severe anxiety or depression symptoms, PCL-5: PTSD checklist for the DSM 5, PTSD: posttraumatic stress disorder, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms and can vary from 3 to 15, RCT: randomized controlled trial.

Supplement Table S5. RCT analysis according to as treated principle

In total: N=36	Early enteral tube feeding N=15	Standard care N=21	*P*-value
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years)	27.73 ± 3.75	30.48 ± 4.20	0.052
Ethnicity	10 (66.7%)	12 (57.1%)	0.25
- Western	2 (13.3%)	8 (38.1%)	
- Non-western			0.65
Education level	5 (33.3%)	10 (47.6%)	
- Primary or secondary	7 (46.7%)	10 (47.6%)	
- Higher	7 (46.7%)	5 (23.8%)	0.15
Primigravida at the time	3 (20.0%)	1 (4.8%)	0.29
History of mental health disease [a]	6 (75.0%)	10 (62.5%)	0.67
HG in previous pregnancy prior to MOTHER study [b]			
Maternal outcomes			
Weight change (kg) [c]	-5.95 ± 5.59	-3.88 ± 2.33	0.19
PUQE-24 at inclusion	12.50 (11.00-14.00)	11.00 (8.75-13.00)	0.10
Mean PUQE-24 in the first 3 weeks after admission	9.73 ± 2.57	9.44 ± 2.30	0.73
Total duration of hospital admissions (days)	7.00 (3.00-10.00)	4.00 (3.00-7.00)	0.25
Readmitted	7 (46.7%)	5 (23.8%)	0.15
Admission after the first trimester	2 (13.3%)	2 (9.5%)	1.00
HADS at inclusion	22.85 ± 5.23	21.67 ± 6.30	0.59
- Anxiety score ≥8 [d]	9 (60.0%)	9 (42.9%)	0.28
- Depression score ≥8 [d]	13 (86.7%)	16 (76.2%)	0.50
Follow-up study			
Depression, anxiety and PTSD symptoms			
HADS 6 weeks postpartum [e]	10.47 ± 4.99	9.33 ± 6.25	0.55
- Anxiety score ≥8 [e]	4 (23.5%)	5 (23.8%)	1.00
- Depression score ≥8 [e]	3 (17.6%)	4 (19.0%)	1.00
HADS at follow-up study	15.00 (7.00-21.00)	8.00 (5.00-15.00)	0.24
- Anxiety score ≥8	6 (40.0%)	6 (28.6%)	0.47
- Depression score ≥8	7 (46.7%)	6 (28.6%)	0.27
PCL-5 score at follow-up study	27.00 (18.00-43.00)	12.00 (3.50-27.50)	**0.046**
- PCL-5 score ≥31	5 (33.3%)	5 (23.8%)	0.71
History of any traumatic event	3 (20.0%)	5 (23.8%)	1.00
History of an obstetric traumatic event	0 (0.0%)	1 (4.8%)	1.00
Women who experienced a subsequent pregnancy after MOTHER participation	7 (46.7%)	7 (33.3%)	0.28
Suffered from HG again between participation of the MOTHER and follow-up study [f]	8 (53.3%)	8 (38.1%)	0.36

Data represented with mean±SD**,** median (IQR) or frequency (%). [a] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [b] Percentage shown is frequency divided by multigravidas. [c] Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. [d] Percentage shown is frequency divided by number of HADS at inclusion available (n=61). [e] Different study group including 38 women. [f] HG defined as: if vomiting symptoms occurred with either multiple medication use, weight loss, hospital admission for HG, requiring tube feeding or whether nausea and vomiting symptoms affected their life and/or work. **Abbreviations:** HG: hyperemesis gravidarum, HADS: hospital anxiety and depression scale: a higher HADS indicates more severe anxiety or depression symptoms, PCL-5: PTSD checklist for the DSM 5, PTSD: posttraumatic stress disorder, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms and can vary from 3 to 15, RCT: randomized controlled trial.

Supplement Table S6. RCT analysis according to per protocol principle

In total: N=33	Early enteral tube feeding N=12	Standard care N=21	*P*-value
Index pregnancy (during MOTHER study)			
Baseline characteristics			
Age (years)	28.00 ± 3.35	30.48 ± 4.20	0.09
			0.42
Ethnicity	8 (66.7%)	12 (57.1%)	
- Western	2 (16.7%)	8 (38.1%)	
- Non-western			0.71
Education level	4 (33.3%)	10 (47.6%)	
- Primary or secondary	6 (50.0%)	10 (47.6%)	
- Higher	7 (58.3%)	5 (23.8%)	0.07
Primigravida at the time	2 (16.7%)	1 (4.8%)	0.54
History of mental health disease [a]	4 (80.0%)	10 (62.5%)	0.62
HG in previous pregnancy prior to MOTHER study [b]			
Maternal outcomes			
Weight change (kg) [c]	-4.93 ± 3.69	-3.88 ± 2.33	0.33
PUQE-24 at inclusion	13.00 (11.00-14.00)	11.00 (8.75-13.00)	0.10
Mean PUQE-24 in the first 3 weeks after admission	9.65 ± 2.75	9.44 ± 2.30	0.82
Total duration of hospital admissions (days)	5.00 (3.00-10.25)	4.00 (3.00-7.00)	0.59
Readmitted	5 (41.7%)	5 (23.8%)	0.43
Admission after the first trimester	2 (16.7%)	2 (9.5%)	0.61
HADS at inclusion	23.30 ± 5.48	21.67 ± 6.30	0.50
- Anxiety score ≥8 [d]	7 (58.3%)	9 (42.9%)	0.43
- Depression score ≥8 [d]	10 (83.3%)	16 (76.2%)	0.52
Follow-up study			
Depression, anxiety and PTSD symptoms			
HADS 6 weeks postpartum [e]	10.79 ± 5.18	9.33 ± 6.26	0.48
- Anxiety score ≥8 [e]	3 (21.4%)	5 (23.8%)	1.00
- Depression score ≥8 [e]	3 (21.4%)	4 (19.0%)	1.00
HADS at follow-up study	15.00 (6.25-22.75)	8.00 (5.00-15.00)	0.27
- Anxiety score ≥8	5 (41.7%)	6 (28.6%)	0.41
- Depression score ≥8	6 (50.0%)	6 (28.6%)	0.27
PCL-5 score at follow-up study	25.00 (19.25-41.75)	12.00 (3.50-27.50)	0.10
- PCL-5 score ≥31	4 (33.3%)	5 (23.8%)	0.69
History of any traumatic event	2 (16.7%)	5 (23.8%)	1.00
History of an obstetric traumatic event	0 (0.0%)	1 (4.8%)	1.00
Women who experienced a subsequent pregnancy after MOTHER participation	6 (50.0%)	7 (33.3%)	0.45
Suffered from HG again between participation of the MOTHER and follow-up study [f]	6 (50.0%)	8 (38.1%)	0.51

Data represented with mean±SD**,** median (IQR) or frequency (%). [a] History of mental health disease can consist of a depressive, anxiety, PTSD or eating disorder. [b] Percentage shown is frequency divided by multigravidas. [c] Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. [d] Percentage shown is frequency divided by number of HADS at inclusion available (n=61). [e] Different study group including 35 women. [f] HG defined as: if vomiting symptoms occurred with either multiple medication use, weight loss, hospital admission for HG, requiring tube feeding or whether nausea and vomiting symptoms affected their life and/or work. **Abbreviations:** HG: hyperemesis gravidarum, HADS: hospital anxiety and depression scale: a higher HADS indicates more severe anxiety or depression symptoms, PCL-5: PTSD checklist for the DSM 5, PTSD: posttraumatic stress disorder, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms and can vary from 3 to 15, RCT: randomized controlled trial.

CHAPTER 6

Hyperemesis gravidarum and vitamin K deficiency: a systematic review

Kelly Nijsten,* Loïs M. van der Minnen,* Hanke M.G. Wiegers, Marjette H. Koot, Saskia Middeldorp, Tessa J. Roseboom, Iris J. Grooten, Rebecca C. Painter

**Both authors contributed equally to this work*

British Journal of Nutrition, 2021; 1-13.

ABSTRACT

Hyperemesis gravidarum (HG), severe nausea and vomiting in pregnancy, can lead to vitamin deficiencies. Little is known about HG-related vitamin K deficiency. We aim to summarize available evidence on the occurrence of HG-related vitamin K deficiency and corresponding maternal and neonatal complications.

A systematic review was conducted, searching Medline and EMBASE from inception to November 12th, 2020.

We identified 1564 articles, of which we included 15 in this study: 14 case reports (n=21 women) and one retrospective cohort study (n=109 women). Nine out of 21 women reported in case reports had a prolonged prothrombin time (PT). The cohort study measured PT in 39/109 women with HG, of whom 10/39 women (26%) had prolonged PT. In total, 30-50% women received vitamin K supplementation after vitamin K deficiency had been diagnosed. Four case reports (n=4 women) reported corresponding maternal complications, all consisting of coagulopathy-related haemorrhage. Nine case reports (n= 16 neonates) reported corresponding neonatal complications including intracranial haemorrhage (n=2 neonates) and embryopathy (n=14 neonates), which consisted of Binder phenotype (n=14 neonates), chondrodysplasia punctata (n=9 neonates) and grey matter heterotopia (n=3 neonates).

In conclusion, vitamin K deficiency and related complications occur among women with HG. In our systematic review, we were unable to assess the incidence rate.

INTRODUCTION

Hyperemesis gravidarum (HG) is severe nausea and vomiting in pregnancy. HG can be complicated by dehydration, electrolyte disturbances, poor nutritional intake and weight loss.[1] Vitamin deficiencies, including vitamin B1 deficiency, can further complicate HG, although little is known about the incidence and consequences of such deficiencies.[2]

The fact that vitamin K deficiency has been frequently described in chronic malnutrition makes it of possible interest in the context of HG.[3, 4] Vitamin K is primarily obtained through dietary intake, but is also synthesized by bacteria in the large intestine.[5] Although vitamin K is a fat soluble vitamin, the body's stores of vitamin K are limited, and vitamin K can be depleted after metabolic surgery and in fat malabsorption syndromes.[3, 4, 6] Vitamin K is important for coagulation, serving as a cofactor in the synthesis of multiple vitamin K-dependent proteins (Factors II, VII, IX, X and protein C and S) in the intrinsic pathway.[7] Besides its effects on coagulation, vitamin K deficiency can also lead to abnormal calcium depositions and growth of cartilage.[6]

Vitamin K deficiency can cause a range of maternal and fetal complications. Maternal and neonatal coagulopathy-related haemorrhage has been described[8, 9] as well as neonatal vitamin K deficiency embryopathy and grey matter heterotopia, most commonly described in the context of maternal vitamin K antagonist medication use.[10, 11] Vitamin K deficiency embryopathy includes Binder phenotype and chondrodysplasia punctata. Binder phenotype is the result of maxillonasal hypoplasia and causes a flat facial profile with a short nose and flat nasal bridge.[12] Chondrodysplasia punctata is a skeletal abnormality classified by stippled calcifications of certain bones, most commonly toes, ankles or fingers.[13] Short or misshapen bones can also be present, for example short distal phalanges, also known as brachytelephalangy.[11] Vitamin K deficiency-related chondrodysplasia punctata should not be mistaken for the genetic form of chondrodysplasia punctata, which is caused by mutations in the X-linked arylsulfatase E (ARSE) gene and can be ruled out by genetic testing.[13] Grey matter heterotopia is a neurological disorder classified by common malformations of cortical development, possibly caused by depletions in the vitamin K dependent growth arrest specific 6 protein which is widely expressed in the nervous system.[14-16]

The fact that HG has a profound impact on nutritional intake, sometimes necessitating enteral or parenteral nutrition, has raised concerns about the possibility that vitamin K deficiency can also occur in pregnancies complicated by HG.[1, 17, 18] Recently, the identification of the immediate

and long term effects of HG for pregnant women and their offspring were selected as urgent research questions by patients and health care professionals, which triggered the current work.[19] In this systematic review, we aim to summarize the available literature on HG-related maternal and neonatal vitamin K deficiency and determine the relevance of measuring vitamin K-related coagulopathy factors or prothrombin time (PT) in routine work-up for women with HG.

METHODS

The study protocol was registered at the website of Prospero, an international prospective register of systematic reviews, on August 17th, 2020 (CRD42020199501). This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Search strategy

We performed a search to identify all available studies reporting on vitamin K deficiency in women suffering from HG and their offspring. We searched Medline and Embase from inception to November 12th, 2020. Our search included the following terms: 'hyperemesis gravidarum', 'pregnancy sickness', 'vitamin K deficiency', 'embryopathy', 'haemorrhage' and their synonyms, as shown in Appendix A. De-duplication of database search results were conducted using Endnote software.[20] We also searched citation lists of eligible primary studies and reviews.

Study selection

Two reviewers (KN and LM) independently screened titles and abstracts. Conflicts were resolved by discussion until consensus was reached, or by consultation of a third reviewer (RP). All potentially relevant articles were retrieved as full text and assessed on the following inclusion and exclusion criteria. Inclusion criteria were: 1. women diagnosed with or admitted for HG with either 2. Maternal vitamin K deficiency or signs/symptoms of vitamin K deficiency (for example: prolonged PT or signs of any type of haemorrhage) and/or 3. Offspring of women with HG with vitamin K deficiency embryopathy or any type of vitamin K deficiency-related haemorrhage. Exclusion criteria were: 1. Non-human subjects, 2. Women with vitamin K deficiency due to any other cause than HG. We included observational studies, case reports, case series and research letters. Conference abstracts were included, if they provided sufficient information. We did not apply any language restrictions.

Data extraction

Data extraction was performed independently by two reviewers (KN and LM). We extracted data on study characteristics, demographics, details about pregnancy and specifically about the severity and clinical course of HG (if available), laboratory results (including prothrombin time, coagulation factors and vitamin K measurements) and both maternal and neonatal outcomes (vitamin K deficiency-related haemorrhage or embryopathy).

Quality assessment

We assessed the risk of bias of included case reports using the Joanna Briggs Institute checklist for case reports and the Newcastle-Ottawa Scale (NOS) for included cohort studies.[21, 22] The NOS assigns up to a total maximum score of 9 based on eight items: a score ≥7 was considered as good quality, a score ≥5 as fair quality and a score ≤4 as poor quality.[22] All included articles were critically appraised and were included, despite of their quality assessment.

Statistical analysis

Data of included case reports were combined by entering available information on baseline characteristics and outcome measures of each reported case of women with HG or their offspring into a SPSS database (SPSS Statistics, version 26.0 for Windows, IBM Corp., Armonk, NY, USA). If a case report included multiple HG patients or multiple HG-exposed offspring, all of the cases were entered separately. Continuous data were presented as means with standard deviations (SD) if they were normally distributed. Not normally distributed continuous data were presented as medians with interquartile ranges (IQR). Dichotomous and categorical data were displayed as frequencies with percentages.

RESULTS

Search results

We identified 1741 articles and one additional article through searching citation lists as shown in **Figure 1.** After removing duplicates, 1564 articles remained for title and abstract screening, of which 36 were deemed possibly eligible. Upon further eligibility screening after full-texts for possibly eligible papers had been retrieved, we included 15 articles reporting on HG and vitamin K deficiency.[23-37] Fourteen of the included studies were case reports[23-27, 29-37] and we included one retrospective cohort study.[28] Two of the included studies were conference abstracts [33, 35] and two additional included studies were written in French.[28, 29]

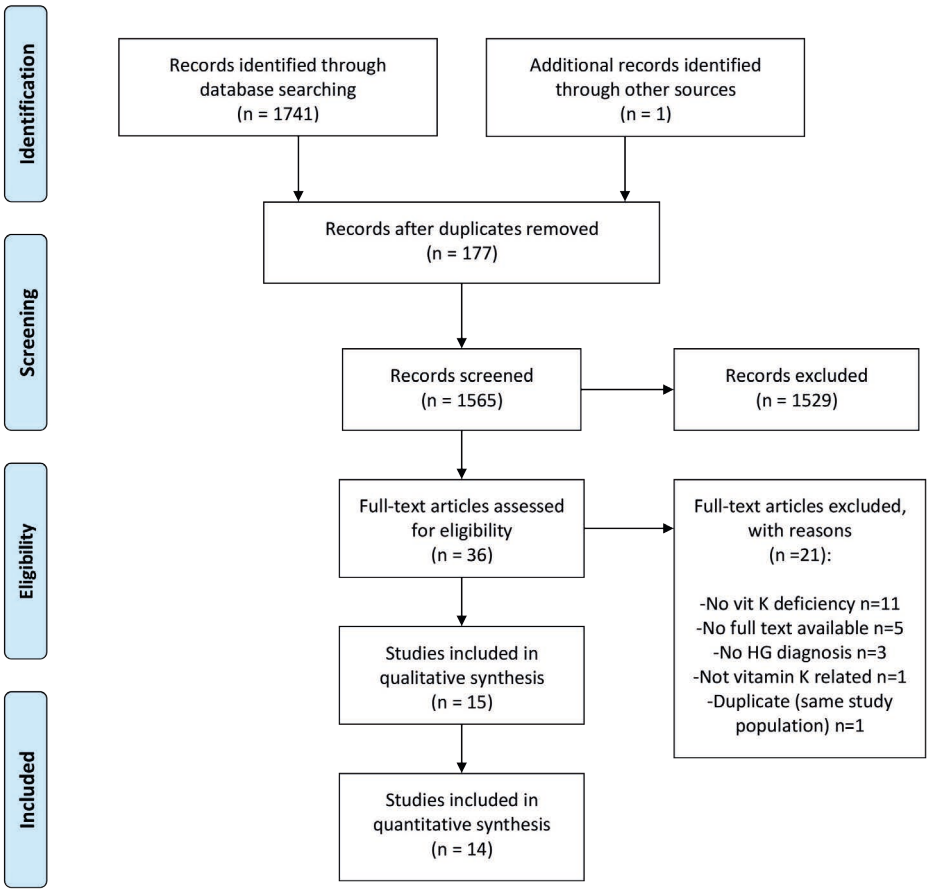


Figure 1. PRISMA diagram selection process of articles

Two case reports included multiple cases: Miller *et al.*[33] included three cases and Toriello *et al.*[37] included eight cases. From the eight cases of Toriello *et al.*[37], case 8 was excluded for this review since vitamin K deficiency was caused by Crohn's disease instead of HG. Case 1 of Toriello *et al.*[37] was identical to the included case report of Robinson *et al.*[34], but contained follow-up information of the neonate, so we combined data of these two case reports.

Risk of bias assessment

The risk of bias assessment of case reports is showed in **Figure 2.** For most domains, case reports were assessed as low risk of bias. However, in half of the studies a patient's medical history was not or poorly described. In addition, in almost half of the studies which reported

a treatment, the treatment was not clearly described in terms of dosage or frequency and therefore was rated as having a high risk of bias. The cohort study was rated to be of fair quality, as shown in **Table 1**.

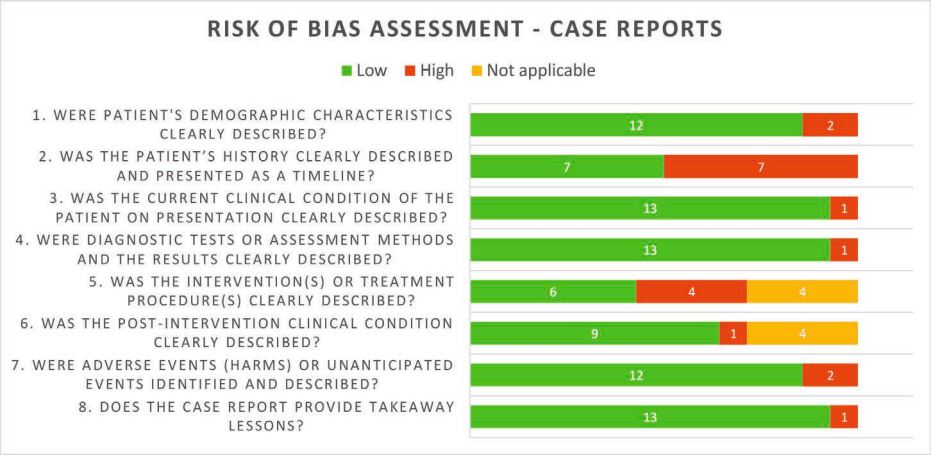


Figure 2. Risk of bias assessment of included case reports

6

Table 1. Risk of bias assessment of the included cohort study using the Newcastle-Ottawa Quality Assessment Scale (NOS)

	Selection	Comparability	Outcome	Total score	Quality score
Chraïbi *et al*, 2014	**		***	5	Fair quality

The NOS risk consisted of 8 items with a total maximum score of 9. A score ≥7 was considered as good quality, a score ≥5 as fair quality and a score ≤4 as poor quality.

Baseline characteristics

Baseline characteristics of all included studies are shown in **Table 2** and data of the included women from case reports was combined and shown in **Table 3**. In 17 women the gestational age of onset of symptoms was reported; in the vast majority (16/17) symptoms of HG had started in the first trimester (mean 8.47±3.16 weeks) **(Table 2 and 3)**. Ten articles (n=14 women) reported whether weight loss during pregnancy due to HG had occurred: 13 out of 14 women had some degree of weight loss, ranging from 5 to 28 kg with an average weight loss of 13.64±8.03 kg compared to 5.6±3.1 kg weight loss reported in the cohort study.[23, 24, 27-32, 34, 35, 37] Nine out of twenty-one women of included case reports had more than 10 kg weight loss.[23, 27, 29, 32, 34, 37]

In all three cases of Miller *et al.*[33] and in case 4 of Toriello *et al.*[37] treatment for HG was not described **(Table 2)**. All other 17 included women of remaining case reports received some form of treatment for HG[23-27, 29-32, 34-37], varying from receiving anti-emetics (9/17) [23, 27, 29, 30, 32, 34, 35, 37], intravenous rehydration (13/17)[23, 24, 27, 29-32, 34, 36, 37] to receiving tube feeding (6/17)[23, 25-27, 32, 37]. Chraïbi *et al.*[28] described a cohort of women admitted for HG: all 109 included women (100%) received intravenous treatment and 106 women (98.1%) received at least one anti-emetic **(Table 2)**. From the 21 women included from case reports, 11 women had been admitted for HG **(Table 3)**.[23-26, 29, 31, 32, 34-37]

Vitamin K deficiency diagnosis

In half of the case reports, a vitamin K deficiency diagnosis was made retrospectively based on neonatal clinical signs of embryopathy.[26, 27, 32, 33, 37] The other half performed laboratory measurements to confirm vitamin K deficiency. PT was most commonly used and prolonged PT was reported as prolonged PT in seconds or as decreased prothrombin levels. PT was measured in 9 out of 21 women included in case reports: 8/21 women (38.1%) had a prolonged PT **(Table 4 and 5)**.[23-25, 29, 31, 34, 36, 37] In 4 out of 9 women PT was measured secondary to maternal signs of haemorrhage.[24, 25, 29, 34] In the other 5 cases PT was included in routine laboratory measurements, without the presence of clinical signs of maternal or fetal haemorrhage or embryopathy.[23, 30, 31, 36, 37]

Four case reports performed additional coagulopathy laboratory measurements. Three case reports measured activated Partial Thromboplastin Time (aPTT).[24, 29, 34] Two of them found a prolonged aPTT, but also found a decreased factor II, VII, IX, X and Protein C and S, which are vitamin K *dependent* coagulation factors.[29, 34] **(Table 4)** The fourth study was the only study that measured vitamin K concentrations in addition to PT and that found vitamin K deficiency (below 0.05ng/mL).[36] Selvarajah *et al.*[35] mentioned that the woman included had a deranged clotting profile, but did not further specify which laboratory measurements were performed **(Table 4)**.

In one neonate coagulation factors were measured postpartum because of low Apgar scores together with signs of haemorrhage: first a haematoma in the hand palm and later intracranial haemorrhage. A prolonged PT together with a decreased Factor II, VII, IX and X was found.[30]

In the cohort study from Chraïbi *et al.*[28], PT was measured in 39 out of 109 women (35.8%) admitted for HG: 10 out of these 39 women (25.6%) had a prolonged PT with a level below 70% and 2 out of these 10 women (5.1%) had a PT level below 50% **(Table 5)**. The cohort

study did not describe why PT was initially measured or whether other coagulation factors were measured.[28]

Vitamin K supplementation

Vitamin K was supplemented in all case reports reporting a prolonged PT (n=8 women and n=1 neonate) and in one women described to had a 'deranged clotting profile' **(Table 4)**.[23-25, 29-31, 34, 36, 37] One additional woman received vitamin K as part of parenteral nutrition, so in total 10 out of 21 (47.6%) women and one neonate received vitamin K supplementation as shown in **Table 5**.[23-25, 29-32, 34, 36, 37] Vitamin K was administered by different routes, but most women (60.0%) and the described neonate received intravenous vitamin K supplementation **(Table 5)**.[23, 24, 29-31, 35, 37] In all of them, PT normalized after vitamin K supplementation.[23-25, 29-31, 34, 36, 37]

In the cohort study of Chraïbi *et al.*[28] 3 out of 10 women with a prolonged PT (level below 70%) received vitamin K, which was not further specified in route of administration, dosage or frequency **(Table 4)**.

Liver function measurements

Liver transaminases tests were performed in 7 out of 21 women included in case reports of whom 4 women (19.0%) had elevated liver transaminases **(Table 4 and 5)**.[23, 29, 31, 36] Three out of these 4 women also had elevated total bilirubin levels and 2 women had elevated gamma glutamyl transferase (GGT) levels.

As shown in **Table 4**, Chraïbi *et al.*[28] reported elevated alanine transaminase (ALAT) and aspartate aminotransferase (ASAT) in respectively 20.7 and 25.7%. PT levels were significantly lower in women with an increased ALAT than in women with normal ALAT levels (68±14% versus 78±9%).

Table 2. Baseline characteristics of included studies

	General			*Demographic characteristics*					*HG severity and course*				*HG treatment*			*Other pregnancy characteristics*			
Study	***Year***	***Country***	***Study design***	***Age (year)***	***Ethnicity***	***G..P..***	***Pre-pregnancy weight (kg)***	***Pre-pregnancy BMI***	***Gestation at onset of HG symptoms***	***Total weight loss (kg)***	***Admitted for HG (duration)***	***Re-admitted for HG***	***IV***	***Anti-emetics***	***TPN***	***Gestation at delivery***	***Sex***	***Birth weight (gram)***	***Medical history or complications***
Alessandri	2010	France	Case report	20	Western	G1P0	70	26.7	7 weeks	15	Yes (4 wks)	No	Yes	Yes	Yes	37 weeks	Girl	2780	Gallbladder lithiasis
Baba	2016	Japan	Case report	36	Asian	G1P0	62	25.8	10 weeks	8	Yes (6 wks)	No	Yes	No	No	-	-	-	Large myoma with intestinal obstruction
Bailey	1964	UK	Case report	21	-	G1P1	-	-	12 weeks	-	Yes (5 wks)	No	-	-	Yes	-	Girl	3000	-
Bhoj	2013	USA	Case report	-	-	G2P2	-	-	6 weeks	-	-	-	-	-	Yes	37 weeks	Girl	2190	-
Brunetti-Pierri	2007	USA	Letter	-	Western	G3P1	-	-	8 weeks	18	-	-	Yes	Yes	Yes	34 weeks	Boy	2540	-
Chraibi [a]	2015	France	Cohort (n=109)	28±5.7	46.5% French	56.4% Nullipara	64.3±13.7	23.9±4.5	46±15 (days)	5.6±3.1	109 (100%)	12.8%	100%	98.1%	-	274±16 (days)	57% Girl	3283 ±527	-
Devignes	2009	France	Case report	23	-	G1	-	-	14 weeks	18	Yes (-)	No	Yes	Yes	No	-	-	-	-
Eventov-Friedman	2009	Israel	Case report	41	-	G8P4	50	19.5	16 weeks	0	-	-	Yes	Yes	No	32 weeks	Boy	2200	-
Kawamura	2007	Japan	Letter	33	Asian	G2P0	45	20.0	9 weeks	5	Yes (5 wks)	No	Yes	No	No	20 weeks	-	-	-
Lane	2015	USA	Case report	21	African-American	G1P0	94.4	-	10 weeks	17	Yes (-)	No	Yes	Yes	Yes	-	Boy	-	
Miller ***CASE 1***	2018	USA	Case report	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CASE 2	"	"	"	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CASE 3	"	"	"	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 2. *Continued*

	General			*Demographic characteristics*					*HG severity and course*				*HG treatment*			*Other pregnancy characteristics*			
Robinson [b]	1998	USA	Case report	22	African-American	G2P0	-	-	4 weeks	14	Yes (-)	Yes (1 time)	Yes	Yes	No	40 weeks	Girl	2800	Anaemia
Selvarajah	2014	UK	Case report	33	Western	G1P0	-	-	7 weeks	6	Yes (-)	-	-	Yes	-				-
Shigemi	2015	Japan	Case report	39	Asian	G1P0	64.1	25.0	8 weeks	-	Yes (1 wk)	Yes (5 times)	Yes	-	-	38 weeks	Girl	2640	Oesophageal hiatus hernia diagnosed at 32 weeks
Toriello ***CASE 2***	2012	USA	Case report	-	Western	-	-	-	-	9	-	-	Yes	No	No	40 weeks	Girl	3600	Syncope
CASE 3	"	"	"	27	Asian	G5P3	50	-	8 weeks	13	-	-	-		Yes	38 weeks	Girl	2520	-
CASE 4	"	"	"	25	Western	-	64.5	-	6 weeks	28	Yes(12wks)	No	-	-	-	33 weeks	Boy	-	-
CASE 5	"	"	"	-	Western	-	-	-	5 weeks	-	-	-	Yes	-	-	40 weeks	Girl	3540	-
CASE 6	"	"	"	-	Western	-	-	-	6 weeks	28	Yes (-)	Yes (3 times)	Yes	Yes	No	32 weeks	Girl	1280	Pre-eclampsia
CASE 7	"	"	"	-	African-American	-	-	-	8 weeks	12	-	-	Yes	Yes	-	33 weeks	Girl	-	-

Abbreviations: G..P..: Gravidity..Parity.., BMI: Body Mass Index, IV: intravenous, TPN: total parenteral nutrition, UK: United Kingdom, USA: United States of America. [a] Cohort study: characteristics presented as mean+SD, median (IQR) or frequency (%). [b] Case of Robinson *et al.* is the same case as case 1 of Toriello *et al.*, so available data is combined.

Table 3. Combined baseline characteristics of included case reports in this systematic review

	N=21		% missing
Demographic characteristics			
Age (years) , median (IQR)	26.00	21.25-35.25	42.9%
Pre-pregnancy weight (kg), median (IQR)	63.05	50.00-68.63	61.9%
Pre-pregnancy BMI (kg/m^2), median (IQR)	25.00	19.75-26.25	76.2%
Ethnic origin, n (%)			33.3%
- Western	7	33.3%	
- Asian	4	19.0%	
- African American	3	14.3%	
Primigravida, n (%)	6	28.6%	38.1%
HG severity & course			
Gestational age at onset of symptoms of HG (weeks), mean±SD	8.47	3.16	19.0%
Total weight loss (kg), mean±SD	-13.64	8.03	33.3%
HG-related hospital admission, n (%)	11	52.4%	47.6%
Length of initial hospitalization (weeks), median (IQR)	5.00	3.25-7.50	71.4%
Readmission, n (%)	3	14.3%	52.4%
HG treatment			
Received treatment for HG, n (%)	17	81.0%	19.0%
Anti-emetics	9	42.9%	
IV Fluids	13	61.9%	
Parenteral nutrition	6	28.6%	
Other pregnancy characteristics			
Gestational age at delivery (weeks), median (IQR)	37.00	32.50-39.00	38.1%
Sex of neonate, n (%)			33.3%
- Female	10	47.6%	
- Male	4	19%	
Birth weight of neonate (grams), median (IQR)	2640.00	2200.00-3000.00	47.6%

Abbreviations: HG: hyperemesis gravidarum, IV: intravenous. Normally distributed continuous variables are presented as means with standard deviations (mean±SDs), skewed variables as medians with interquartile ranges (IQR) and dichotomous or categorical variables as frequencies with percentages (%).

Table 4. Maternal and neonatal outcomes of included studies

	Maternal:					**Neonatal:**	*Vitamin K embryopathy*					
Study	***PT prolonged*** (seconds or %; gestation)	***Vitamin K and/or other coagulation factors measured***	***Elevated Liver-enzymes***	***Vitamin K supplementation*** *(dosage; gestation)*	***Maternal complications***	***Neonatal Haemorrhage***	***Binder pheno type***	***Chrondro dysplasia punctata***	***Brachy tele-phalangy***	***Grey matter hetero topia***	***Anomaly first detected***	***Additional information***
Alessandri	Yes (42%; 11wks 25%; 12 wks)		Yes (ALAT 186 UI/l)	iv at 12 wks	-	-	Yes	Yes	Yes	-	US: 24 wks	-
Baba	Yes (14.2s; 16 wks)	Normal APTT	-	iv 10mg/day at 16 wks	Intraperitoneal haemorrhage	-	-	-	-	-	-	-
Bailey	Yes (63s)	-	-	im	Haematuria, vaginal bleeding	-	-	-	-	-	-	-
Bhoj	-	-	-	-	-	-	Yes	Yes	Yes	-	Postpartum	Epileptic seizures, chiari type II malformation
Brunetti-Pierri	-	-	-	-	-	-	Yes	Yes	Yes	Yes	US: 20 wks	Epileptic seizures, ventilary support, long term disability
Chraïbi [a]	Yes (10 out of 39 women (25.6%)) [b]	-	Yes (in 20.7 to 25.7%)	3 out of 10 (30%)	-	-	-	-	-	-	-	-
Devignes	Yes (11%; 18 wks)	aPTT↑, factor II, VII, X, proteïn C,S↓	Yes (ALAT 353 UI/l)	10 mg iv once at 18 wks	Haematuria, rectal bleeding	-	-	-	-	-	-	-
Eventov-Friedman	Normal		-	-	-	Intracranial haemorrhage	-	-	-	-	Postpartum	Neonatal lab: PT↑, Factor II, VII,IX,X↓. Treatment: 1mg iv
Kawamura	Yes (28%; 14 wks)		Yes	10mg iv & 2mg/ day at 14 wks	-	Intracranial haemorrhage	-	-	-	-	US: 17 wks	Induced abortion due to US anomalies, hydrocephalus
Lane	-	-	Normal	In TPN at 15 wks	-	-	yes	-	-	-	US: 14 wks	-
Miller CASE 1	-	-	-	-	-	-	Yes	Yes	Yes	-	Unclear	Neonate died at 3.5 months
CASE 2	-	-	-	-	-	-	Yes	Yes	Yes	Yes	Unclear	-
CASE 3	-	-	-	-	-	-	Yes	Yes	yes	-	Unclear	-

Table 4. *Continued*

	Maternal:					**Neonatal:**	*Vitamin K embryopathy*					
Robinson	Yes (36.5s; 15 wks)	aPTT↑, Factors II, VII, IX, X↓	Normal	10 mg sc/day for 3 days at 15 wks	Epistaxis with 1 liter blood loss	-	Yes	-	-	-	US: 17 wks	Calcaneal asymmetry
Selvarajah	-	Deranged clotting profile at 13 wks	Normal	Iv at 13 wks	-	-	-	-	-	-	-	-
Shigemi	Yes (15.2s; 9 wks 19.7s; 11 wks)	Vit K↓ (<0.05 ng/mL) & factor VII↓	Yes (ALAT 72 UI/l)	15 mg oral/day for 5 wks at 11 wks	-	-	-	-	-	-	-	-
Toriello CASE 2	-	-	-	-	-	-	Yes	-	Yes	-	Postpartum	-
CASE 3	Yes (22%; 8 wks)	-	-	Iv at 8 wks	-	-	Yes	-	Yes	-	US: 30 wks	-
CASE 4	-	-	-	-	-	-	Yes	Yes	Yes	-	Postpartum	Spastic quadriplegia and severe intellectual disability.
CASE 5	-	-	-	-	-	-	Yes	-	Yes	Yes	Postpartum	Normal development (1/2yr)
CASE 6	-	-	-	-	-	-	Yes	Yes	Yes	-	Postpartum	Normal development (3yr)
CASE 7	-	-	-	-	-	-	Yes	Yes	-	-	US	Trachystomy & gastrostomy

Abbreviations: PT: prothrombin time, US: ultrasound (perinatal), aPTT: activated Partial Thromboplastin Time, TPN: total parenteral nutrition. [a] Cohort study: data presented as frequencies/percentages. [b] PT measured in 39/109 women.

Table 5. Combined outcomes of included case reports in this systematic review

	N=21	
Maternal		
PT Prolonged, n (%)	8	38.1%
Vitamin K measured, n (%)	1	4.8%
Elevated liver transaminases, n (%)	4	19.0%
Vitamin K supplementation, n (%)	10	47.6%
- Oral [a]	2	20.0%
- Subcutaneous [a]	1	10.0%
- Intramuscular [a]	1	10.0%
- Intravenous [a]	6	60.0%
Gestational age when women received vitamin K supplementation (weeks), median (IQR)	14.00	11.50-15.50
Maternal haemorrhage occurred, n (%)	4	19.0%
Neonatal		
Neonatal haemorrhage occurred, n (%)	2	9.5%
Vitamin K embryopathy, n (%)	14	66.7%
- Binder phenotype	14	66.7%
- Chondrodysplasia punctata	9	42.9%
- Brachytelephalangy	11	52.4%
- Grey matter heterotopia	3	14.3%
Anomalies detected on foetal ultrasound, n (%)	7	33.3%
- Gestational age when anomalies were first detected, median (IQR)	18.50	16.25-25.50
Anomalies detected postpartum, n (%)	6	28.6%

Abbreviations: PT: Prothrombin time. Skewed variables are presented as medians with interquartile ranges (IQR) and dichotomous or categorical variables as frequencies with percentages (%). [a] Percentage shown is percentage of women who received vitamin K supplementation.

Maternal complications due to HG-related vitamin K deficiency

We identified four studies, including four women, that reported on maternal complications due to HG-related vitamin K deficiency. All four studies reported coagulopathy-related haemorrhage **(Table 4 and 5)**. Two women had mild haemorrhage symptoms, not in the context of their delivery, consisting of haematuria, bruising and/or vaginal or rectal bleeding.[25, 29] Two other studies reported more severe cases of haemorrhage. Robinson *et al.*[34] described a case of severe epistaxis with one litre blood loss, which was initially treated with topical silver nitrate and after the diagnosis of vitamin K deficiency was made vitamin K was supplemented. Baba *et al.*[24] described a case of a woman with HG who developed intraperitoneal haemorrhage due to a pedunculated myoma, which was operatively resected at 16 weeks gestation. In total,

perioperative blood loss contained 290 ml of which 110 ml intraperitoneal blood loss was noted at the start of the operation. Postoperative laboratory results revealed coagulopathy based on a prolonged PT with a normal aPTT and international normalized ratio (INR). Coagulopathy was strongly suspected to be secondary to vitamin K deficiency, since PT normalized after intravenous supplementation of vitamin K, and the amount of blood loss was thought to be insufficient to induce secondary coagulopathy.

Neonatal complications due to HG-related vitamin K deficiency

Nine studies reported neonatal complications due to HG-related vitamin K deficiency.[23, 26, 27, 30-34, 37] Two case reports, including two neonates, reported neonatal intracranial haemorrhage[30, 31] and seven case reports, including 14 neonates, reported neonatal embryopathy as shown in **Table 4 and 5**.[23, 26, 27, 32-34, 37]

Neonatal intracranial haemorrhage

Two studies reported neonatal intracranial haemorrhage **(Table 5)**. Kawamura *et al.*[31] described a case where fetal intracranial haemorrhage accompanied by hydrocephalus was detected during the midtrimester ultrasound at 17 weeks gestation. Due to these fetal anomalies the woman decided to terminate her pregnancy. Autopsy showed a subarachnoid haemorrhage with hemosiderin deposits to the choroid plexus near the foramen of Luschka and on the surface of the brainstem which blocked the pathway of cerebrospinal fluid absorption and subsequently lead to a non-obstructive hydrocephalus. No evidence of chromosomal abnormalities was found and a diagnosis of a Dandy-Walker syndrome was rejected because of the presence of a non-obstructive hydrocephalus.

Eventov-Friedman *et al.*[30] also reported a case of neonatal intracranial haemorrhage, which was diagnosed postpartum **(Table 4)**. An emergency caesarean was performed at 32 weeks gestation due to suspected fetal distress. The neonate had an Apgar score of 1, 1 and 3, after respectively 1, 5 and 10 minutes. A cranial ultrasound revealed extensive intracranial haemorrhage and neonatal coagulopathy laboratory results confirmed a vitamin K deficiency. A cranial computed tomography on day two postpartum showed no midline shift and therefore the infant was managed conservatively. The neonate developed recurrent seizures which was treated with phenobarbital. No further neonatal long term outcomes were described.

Neonatal vitamin K related-embryopathy

From the 14 neonates diagnosed with vitamin K related-embryopathy in the studies included in our review, all neonates had Binder phenotype, 9 neonates also had chondrodysplasia punctata

of whom three also suffered from grey matter heterotopia as shown in **Table 5**.[23, 26, 27, 32-34, 37] Also brachytelephalangy was noted in 11 out of 14 neonates with vitamin K-related embryopathy.[23, 26, 27, 33, 37] Genetic testing was performed in 9 out of 14 neonates, none of which found genetic abnormalities.[23, 26, 27, 32, 34, 37] Three studies specifically described that no mutations in the ARSE gene were found.[26, 27, 37]

Anomalies detected and timing of vitamin K supplementation

Of the 10 women who received vitamin K supplementation, 5 cases had neonatal complications.[23, 31, 32, 34, 37] Four cases had neonatal vitamin K deficiency-related embryopathy[23, 32, 34, 37] and one case had intracranial haemorrhage.[31] As shown in **Table 4**, in Alessandri *et al.*[23], Kawamura *et al.*[31], Robinson *et al.*[34] and case 3 of Toriello *et al.*[37] vitamin K supplementation was started *before* fetal anomalies were detected on perinatal ultrasound. Here, PT was measured on maternal indication or during routine maternal laboratory measurements and subsequently vitamin K was supplemented at respectively 12, 14, 15 and 8 weeks gestation. In Lane *et al.*[32] vitamin K was administered *after* fetal anomalies were detected on perinatal ultrasound. Vitamin K was included in parenteral nutrition which was started at 15 weeks gestation. The median gestational age when vitamin K supplementation was commenced was 14 weeks (IQR 12-16) compared to the median gestational age of 19 weeks (IQR 16-26) when fetal anomalies were detected on perinatal ultrasound **(Table 5)**.

Neonatal prognosis

Eleven out of 21 neonates had been given a good prognosis by the paediatrician during follow-up visits.[23, 24, 26, 29, 32, 36, 37] One neonate described in Miller *et al.*[33] died at 3.5 months: she had a severe nasal aperture stenosis, critical cervical spinal stenosis and myelomalacia of the upper cervical cord **(Table 4)**. Two neonates were described as having a poor prognosis.[27, 37] One of these neonates suffered from long term disability due to ventilatory support dependence and severe neurodevelopmental delay.[27] While the other neonate described in case 4 of Toriello *et al.*[37] suffered from severe intellectual disability and spastic quadriplegia following spinal surgery because of severe cervical spinal stenosis **(Table 4)**. [27] Two neonates described in Bhoj *et al.*[26] and case 6 of Toriello *et al.*[37] had a mild delay in neurodevelopment.

DISCUSSION

Principal findings

In this systematic review, which identified 15 articles, we found evidence that vitamin K deficiency secondary to HG can lead to severe adverse maternal and neonatal outcomes. Our review highlights the fact that HG, usually considered a benign and self-limiting condition of early pregnancy, can lead to irreversible morbidity and mortality, and therefore deserves the prompt attention of clinicians to avoid these sequelae. Although selective reporting likely has affected our findings, two thirds of the neonates included in the case reports suffered from vitamin K embryopathy, making it the most commonly reported vitamin K deficiency-related complication among women with HG, followed by maternal haemorrhage (19%) and neonatal haemorrhage (10%). A further 26-38% of cases showed evidence of disturbed maternal coagulation due to vitamin K deficiency, with 30-48% receiving vitamin K supplementation.

Strengths and limitations of the study

One of the main strengths of this study is that it presents an overview of a rare complication, and summarizes the evidence on vitamin K deficiency in women with HG and their offspring. Besides case reports, research letters and conference abstracts, we were also able to include one cohort study. We did not apply a date or language restriction, which avoided selective inclusion of English language literature. Lastly, all articles included were critically appraised and were rated as low to moderate bias.

Our study also has some limitations. Although we were able to include one cohort study, the remainder of the included studies were case reports. Case reports are subject to publication bias, and could result in a bias towards the increased reporting of more unfavourable outcomes. The fact that our review only recovered case reports and one cohort study hampers estimation of the incidence of vitamin K deficiency among women with HG. Furthermore, the case reports suffered from incomplete reporting of data essential to our review, which compromised our ability to link indicators of the severity or course of HG to maternal, fetal and neonatal outcomes in many studies; some articles focused primarily on the course of HG and maternal complications, while other case reports focused more on neonatal complications and did not report extensive details of HG. In addition, direct measures of vitamin K deficiency, for example PT, were only reported in 43% of included women, which hampered our ability to determine timing of maternal vitamin K depletion and its relation to fetal and neonatal outcomes in many cases.

Interpretation

Due to the fact that our review included mostly case reports, we are not able to estimate the incidence of vitamin K deficiency among women with HG. In the included cohort study however, 10 out of 39 women (26%) had a prolonged PT, suggesting that the presence of vitamin K deficiency may be more common among women suffering from HG than currently recognised.[28] However, the fact that PT was only measured in 39 out of the 109 women in the cohort, raises the possibility of this percentage only being representative of a selected group of more severely affected patients. Unfortunately, we are uninformed about the severity of HG in these specific 39 cases. Unlike the included case reports, the cohort study reported no further vitamin K deficiency complications, suggesting that only a small proportion of cases of vitamin K deficiency lead to complications including haemorrhage and embryopathy. A larger prospective cohort study measuring vitamin K deficiency in women with HG could determine the true incidence of both phenomena. The fact that this systematic review found mostly neonatal complications (9 studies) instead of maternal complications (4 studies) could largely be explained by the given that only very little vitamin K crosses the placenta from mother to fetus. This would suggest that the fetus is more at risk to develop a more severe vitamin K deficiency than mother.[38, 39]

It is hypothesised that in women with HG vitamin K deficiency is caused by poor nutritional intake, as is evident from marked weight loss. Most women in the included case reports had severe weight loss, with a mean weight loss of 13.6 kg. In examining the association between the severity of weight loss and presence of vitamin K deficiency induced complications, we found that in three cases reporting maternal haemorrhage, the maternal weight loss varied from 8 to 18kg.[24, 29, 34] In two included cases where neonates had long term disabilities, the maternal weight loss due to HG was respectively 18 and 28 kg.[27, 37] The woman who lost 28 kg, was also admitted to the hospital for 12 weeks in total.[37] The mean weight loss of 5.6±3.1 kg in women with HG included in Chraïbi *et al.*[28], but also in other HG cohort studies[40, 41], was considerably lower and they did not report any vitamin K deficiency-related complications This may suggest that a more severe clinical course of HG causes more severe malnutrition which can in line lead to an increased risk of developing vitamin K deficiency and related complications.

Embryopathy is also described in neonates born to women using warfarin, a vitamin K antagonist, during pregnancy, better known as the fetal warfarin syndrome.[10] Studies assessing the fetal warfarin syndrome showed that mainly first trimester deficiency of vitamin K results in embryopathy[42, 43] and that warfarin use throughout every trimester of pregnancy can result

in neonatal central nervous system (CNS) abnormalities.[10, 44] This corresponds to the onset, duration and severity of HG and its relation to neonatal complications reported in included case reports. In all cases reporting embryopathy the onset of HG lay in the first trimester and the two neonates described to have long term disabilities were born to mothers with a severe HG with a prolonged disease course. [27, 37]

The optimal timing of measuring vitamin K deficiency though is difficult to define. When maternal haemorrhage complications occurred, laboratory tests were performed at the time and vitamin K deficiency was then diagnosed and subsequently supplemented.[24, 25, 29, 34] In case reports describing neonatal embryopathy however, in the majority of the cases fetal anomalies were found on antenatal ultrasonography, despite earlier treatment with vitamin K. Since the origin of neonatal embryopathy lays in the first trimester and vitamin K supplementation took place primarily in the *second* trimester, the most likely explanation for this would be that vitamin K was supplemented too late and that fetal anomalies were already present at time of vitamin K treatment. Bearing this in mind, a solution would be to prophylactically administer vitamin K in women with HG, which has been proposed in previous studies.[23, 24, 27, 30, 31, 33-36] Most of these studies suggested that prophylactic treatment should be given in women with severe HG, undernutrition or severe weight loss but do not further specify this.[23, 24, 35] On the contrary, the Royal College of Obstetricians and Gynaecologists advices that women admitted with HG should be offered thromboprophylaxis because of an increased risk of venous thromboembolism. This might make care givers reluctant to follow that advice, although it is important to clarify that vitamin K supplementation does not increase the risk of venous thromboembolic complications.[45]

HG is known to be associated with raised transaminases, and can lead to liver dysfunction. Nonetheless, we think it is unlikely that liver dysfunction due to HG led to increased PT described in a number of articles. This is illustrated by the fact that the 4 case reports to measure liver transaminases found universally raised PT, which promptly resolved after vitamin K supplementation.

Conclusion

In this systematic review, we have demonstrated that women with HG can develop vitamin K deficiency and the corresponding maternal and neonatal complications. We were not able to derive the incidence among women with HG from the studies we retrieved, but found evidence vitamin K deficiency could affect up to 26% of HG patients. Which aspects of HG severity or disease course increase the risk of vitamin K deficiency remains unclear; severe

weight loss and prolonged disease did appear to be common factors in affected HG patients, and may therefore present risk factors. Larger prospective cohort studies of women with HG are needed to assess the incidence of vitamin K deficiency. It remains to be established whether early prophylactic vitamin K supplementation is safe and effective in preventing complications including embryopathy. Meanwhile, in women with HG and severe malnutrition or weight loss, measuring and supplementing vitamin K should be considered in order to prevent maternal or neonatal complications.

Acknowledgements

None.

Funding

Drs. K. Nijsten is funded by the Amsterdam Reproduction & Development (AR&D), project number 23346, and the Dutch Heart Foundation, grant number 2013T085.

Contribution to authorship

KN and RCP conceived the study. KN and LvdM performed the search, screened for eligible studies and performed data extraction. KN and LvdM performed all statistical analyses, supervised by RCP. KN and LvdM drafted the manuscript. HMGW, SM, MHK, IJG, TJR and RCP contributed in interpreting the results and revising the manuscript. All authors approved the final draft of the manuscript.

REFERENCES

1. Niebyl JR. Clinical practice. Nausea and vomiting in pregnancy. N Engl J Med. 2010;363(16):1544-50.
2. Oudman E, Wijnia JW, Oey M, van Dam M, Painter RC, Postma A. Wernicke's encephalopathy in hyperemesis gravidarum: A systematic review. Eur J Obstet Gynecol Reprod Biol. 2019;236:84-93.
3. Krzyżanowska P, Książyk J, Kocielińska-Kłos M, Banaś E, Kaleta M, Popińska K, et al. Vitamin K status in patients with short bowel syndrome. Clin Nutr. 2012;31(6):1015-7.
4. Sherf-Dagan S, Goldenshluger A, Azran C, Sakran N, Sinai T, Ben-Porat T. Vitamin K-what is known regarding bariatric surgery patients: a systematic review. Surg Obes Relat Dis. 2019;15(8):1402-13.
5. Walther B, Karl JP, Booth SL, Boyaval P. Menaquinones, bacteria, and the food supply: the relevance of dairy and fermented food products to vitamin K requirements. Adv Nutr. 2013;4(4):463-73.
6. Shearer MJ, Newman P. Metabolism and cell biology of vitamin K. Thromb Haemost. 2008;100(4):530-47.
7. Shearer MJ. Vitamin K in parenteral nutrition. Gastroenterology. 2009;137(5 Suppl):S105-18.
8. Menger H, Lin AE, Toriello HV, Bernert G, Spranger JW. Vitamin K deficiency embryopathy: a phenocopy of the warfarin embryopathy due to a disorder of embryonic vitamin K metabolism. Am J Med Genet. 1997;72(2):129-34.
9. Shearer MJ. Vitamin K deficiency bleeding (VKDB) in early infancy. Blood Rev. 2009;23(2):49-59.
10. Hall JG, Pauli RM, Wilson KM. Maternal and fetal sequelae of anticoagulation during pregnancy. Am J Med. 1980;68(1):122-40.
11. Howe AM, Webster WS, Lipson AH, Halliday JL, Sheffield LJ. Binder's syndrome due to prenatal vitamin K deficiency: a theory of pathogenesis. Aust Dent J. 1992;37(6):453-60.
12. Binder K. Dystosis maxillo-nasalis: ein arhinen-cephaler. Unabhängige Monatshefte für die politische und geistige Gestaltung der Gegenwart. 1962;6:438-44.
13. Irving MD, Chitty LS, Mansour S, Hall CM. Chondrodysplasia punctata: a clinical diagnostic and radiological review. Clin Dysmorphol. 2008;17(4):229-41.
14. Alisi L, Cao R, De Angelis C, Cafolla A, Caramia F, Cartocci G, et al. The Relationships Between Vitamin K and Cognition: A Review of Current Evidence. Frontiers in Neurology. 2019;10(239).
15. Ferland G. Vitamin K and the nervous system: an overview of its actions. Advances in nutrition (Bethesda, Md). 2012;3(2):204-12.
16. Barkovich AJ, Kuzniecky RI. Gray matter heterotopia. Neurology. 2000;55(11):1603-8.
17. Birkeland E, Stokke G, Tangvik RJ, Torkildsen EA, Boateng J, Wollen AL, et al. Norwegian PUQE (Pregnancy-Unique Quantification of Emesis and nausea) identifies patients with hyperemesis gravidarum and poor nutritional intake: a prospective cohort validation study. PLoS One. 2015;10(4):e0119962-e.
18. van Stuijvenberg ME, Schabort I, Labadarios D, Nel JT. The nutritional status and treatment of patients with hyperemesis gravidarum. Am J Obstet Gynecol. 1995;172(5):1585-91.
19. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient-clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ Open. 2021;11(1):e041254.
20. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. 2016;104(3):240-3.
21. Moola S, Munn, Z., Tufanaru, C., Aromataris, E., Sears, K., Sfetcu, R., Currie, M., Qureshi, R., Mattis, P., Lisy, K., Mu, P-F. Joanna Briggs Institute Reviewer's Manual. 2017 [Chapter 7: Systematic reviews of etiology and risk.]. Available from: https://reviewersmanual.joannabriggs.org/.

22. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. 2014 [Available from: www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

23. Alessandri JL, Ramful D, Cuillier F. Binder phenotype and brachytelephalangic chondrodysplasia punctata secondary to maternal vitamin K deficiency. Clin Dysmorphol. 2010;19(2):85-7.

24. Baba Y, Morisawa H, Saito K, Takahashi H, Rifu K, Matsubara S. Intraperitoneal Hemorrhage in a Pregnant Woman with Hyperemesis Gravidarum: Vitamin K Deficiency as a Possible Cause. Case Rep Obstet Gynecol. 2016;2016:5384943.

25. Bailey P. VITAMIN-K DEFICIENCY IN EARLY PREGNANCY. Br Med J. 1964;2(5418):1199.

26. Bhoj E, Dubbs H, McDonald-McGinn D, Zackai E. Late-onset partial complex seizures secondary to cortical dysplasia in a patient with maternal vitamin K deficient embryopathy: Comments on the article by Toriello et al. [2013] and first report of the natural history. American Journal of Medical Genetics Part A. 2013;161(9):2396-8.

27. Brunetti-Pierri N, Hunter JV, Boerkoel CF. Gray matter heterotopias and brachytelephalangic chondrodysplasia punctata: a complication of hyperemesis gravidarum induced vitamin K deficiency? Am J Med Genet A. 2007;143a(2):200-4.

28. Chraïbi Z, Ouldamer L, Body G, Bacq Y. [Hyperemesis gravidarum: a ten-year French retrospective study of 109 patients]. Presse Med. 2015;44(1):e13-22.

29. Devignes J, Grare M, Raft J, Vial F, Hacquard M, Bouaziz H, et al. [A case of cutaneous and mucous haemorrhage secondary to vitamin K deficiency in hyperemesis gravidarum]. Ann Fr Anesth Reanim. 2009;28(7-8):697-700.

30. Eventov-Friedman S, Klinger G, Shinwell ES. Third trimester fetal intracranial hemorrhage owing to vitamin K deficiency associated with hyperemesis gravidarum. J Pediatr Hematol Oncol. 2009;31(12):985-8.

31. Kawamura Y, Kawamata K, Shinya M, Higashi M, Niiro M, Douchi T. Vitamin K deficiency in hyperemesis gravidarum as a potential cause of fetal intracranial hemorrhage and hydrocephalus. Prenatal Diagnosis. 2008;28(1):59-61.

32. Lane AS, Stallworth JL, Eichelberger KY, Trofatter KF. Vitamin K Deficiency Embryopathy from Hyperemesis Gravidarum. Case Rep Obstet Gynecol. 2015;2015:324173.

33. Miller SF, Mostafavi R, Mroczkowski H, Khalid A, Ward J, Pivnick EK. Chondrodysplasia punctata associated with maternal hyperemesis gravidarum: A variable and potentially preventable phenotype. American Journal of Obstetrics and Gynecology. 2018;218 (1 Supplement 1):S171-S2.

34. Robinson JN, Banerjee R, Thiet MP. Coagulopathy secondary to vitamin K deficiency in hyperemesis gravidarum. Obstet Gynecol. 1998;92(4 Pt 2):673-5.

35. Selvarajah SL, Gupta M, Deo ND. Vitamin K deficiency: An under-reported phenomenon in Hyperemesis Gravidarum? Archives of Disease in Childhood: Fetal and Neonatal Edition. 2014;99:A145-A6.

36. Shigemi D, Nakanishi K, Miyazaki M, Shibata Y, Suzuki S. A case of maternal vitamin K deficiency associated with hyperemesis gravidarum: its potential impact on fetal blood coagulability. J Nippon Med Sch. 2015;82(1):54-8.

37. Toriello HV, Erick M, Alessandri JL, Bailey D, Brunetti-Pierri N, Cox H, et al. Maternal vitamin K deficient embryopathy: association with hyperemesis gravidarum and Crohn disease. Am J Med Genet A. 2013;161a(3):417-29.

38. Shearer MJ, Rahim S, Barkhan P, Stimmler L. Plasma vitamin K1 in mothers and their newborn babies. Lancet. 1982;2(8296):460-3.

39. Greer FR. Vitamin K the basics--what's new? Early Hum Dev. 2010;86 Suppl 1:43-7.

40. Grooten IJ, Koot MH, van der Post JA, Bais JM, Ris-Stalpers C, Naaktgeboren C, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. Am J Clin Nutr. 2017;106(3):812-20.

41. Stokke G, Gjelsvik BL, Flaatten KT, Birkeland E, Flaatten H, Trovik J. Hyperemesis gravidarum, nutritional treatment by nasogastric tube feeding: a 10-year retrospective cohort study. Acta Obstet Gynecol Scand. 2015;94(4):359-67.

42. Howe AM, Webster WS. Vitamin K--its essential role in craniofacial development. A review of the literature regarding vitamin K and craniofacial development. Aust Dent J. 1994;39(2):88-92.

43. Chan WS, Anand S, Ginsberg JS. Anticoagulation of Pregnant Women With Mechanical Heart Valves: A Systematic Review of the Literature. Archives of Internal Medicine. 2000;160(2):191-6.

44. Duhl AJ, Paidas MJ, Ural SH, Branch W, Casele H, Cox-Gill J, et al. Antithrombotic therapy and pregnancy: consensus report and recommendations for prevention and treatment of venous thromboembolism and adverse pregnancy outcomes. Am J Obstet Gynecol. 2007;197(5):457.e1-21.

45. Royal College of Obstetricians and Gynaecologists. The Management of Nausea and Vomiting of Pregnancy and Hyperemesis Gravidarum. Green-top Guideline No. 69. 2016.

APPENDIX A

Systematic review search strategies

Medline	
#1. ((pregnan*[Title/Abstract]) OR (gestation*[Title/Abstract])) OR (pregnancy[MeSH Terms])	1.079.028
#2. (nausea[Title/Abstract]) OR (vomit*[Title/Abstract])	94.807
#3. (#1) AND (#2)	6.337
#4. morning sickness[MeSH Terms]	1.796
#5. hyperemesis gravid*[Title/Abstract]	1.549
#6. ((#3) OR (#4)) OR (#5)	7.679
#7. ((((((vitamin k deficiency[MeSH Terms]) OR (vitamin k def*[Title/Abstract])) OR (embryopathy[MeSH Terms])) OR (embryopath*[Title/Abstract])) OR (hemorrhage[MeSH Terms])) OR (hemorrh*[Title/Abstract])) OR (bleeding[Title/Abstract])	633.608
#8. (#6) AND (#7)	1.137

Embase	
#1. exp hyperemesis gravidarum/ or 'hyperemesis gravidarum'.ti,ab.	3.311
#2. ((pregnan* or gravidar* or gravidit* or gestat* or antenat* or prenat* or ante-nat* or pre-nat*) adj6 (nausea* or antinause* or vomit* or antivomit* or emes* or hypereme* or antiemet* or emetic*)).tw,kw.	4.267
#3. 1 or 2	5.293
#4. exp vitamin k deficiency/ or exp bleeding/ or exp embryopathy/ or embryopat*.ab,ti. or vitamin k def*.ab,ti. or hemorrh*.ab,ti. or bleeding.ab,ti.	1.200.837
#5. 3 AND 4	604

PART III
OFFSPRINGS FUTURE HEALTH

CHAPTER 7

Perinatal outcomes of offspring of women suffering from hyperemesis gravidarum: a systematic review and meta-analysis

Larissa A.W. Jansen, Kelly Nijsten, Jacqueline Limpens, Rik van Eekelen, Marjette H. Koot, Iris J. Grooten, Tessa J. Roseboom, Rebecca C. Painter

Submitted

AJOG AT GLANCE

A. Why was this study conducted?

In 2012 a systematic review was published describing the effects of HG on perinatal outcomes. Since then, several studies have been conducted to evaluate perinatal outcomes after maternal HG.

B. What are the key findings?

Meta-analysis shows that HG is significantly associated with:

- increased risk of:
  - preterm birth <34 weeks (OR 2.81)
  - birth weight <1500 grams (OR 1.43)
  - neonatal intensive care unit admission (OR 1.20)
  - placental abruption (OR 1.15)
  - neonatal resuscitation (OR 1.07)
- decreased risk of:
  - birthweight >4000 grams (OR 0.74)
  - stillbirth (OR 0.92)

C. What does this study add to what is already known?

Maternal HG is linked to an increased chance of several adverse perinatal outcomes, but may protect against macrosomia and stillbirth.

ABSTRACT

Background

Hyperemesis gravidarum (HG) is the severe form of nausea and vomiting during pregnancy and can lead to undernutrition and low maternal weight gain. Previous epidemiologic and animal studies have shown that undernutrition and low maternal weight gain in pregnancy can increase the risk of unfavorable perinatal outcomes, like shorter gestational age, small for gestational age, lower weight at birth.

Objective

To evaluate the effect of HG on perinatal outcomes.

Data sources

OVID Medline and Embase were searched from inception to February 9th, 2022

Study eligibility

Studies reporting on perinatal outcomes of infants born to mothers with HG or severe NVP were included. Case reports, case series, animal studies, reviews, editorials and conference abstract were excluded.

Study appraisal and synthesis methods

Two reviewers independently selected and extracted data. Risk of bias was assessed by the Newcastle-Ottawa Quality Assessment Scale. We conducted meta-analyses where possible.

Results

Our search yielded 1387 unique papers, of which 61 studies (n=20,532,671 participants) were included in our systematic review. Meta-analyses showed that HG was associated with preterm birth <34 weeks (2 studies n=2,882: OR 2.81, 95%CI:1.69-4.67), birth weight <1500 grams (2 studies, n=489,141: OR1.43, 95%CI:1.02-1.99), neonatal resuscitation (2 studies, n=4,289,344: OR 1.07, 95%CI:1.05-1.10), neonatal intensive care unit admission (7 studies, n=6,509,702: OR 1.20, 95%CI:1.14-1.26) and placental abruption (6 studies, n=9,368,360: OR 1.15, 95%CI:1.05-1.25). HG was associated with reductions in birthweight >4000 grams (2 studies, n=5,503,120: OR 0.74, 95%CI:0.72-0.76) and stillbirth (9 studies, n=3,973,154: OR 0.92, 95%CI:0.85-0.99). Meta-analyses revealed no association between HG and Apgar scores <7 at 1 and 5 minutes; fetal loss, perinatal deaths and neonatal deaths.

Conclusion

HG is associated with several adverse perinatal outcomes including low birth weight and preterm birth. We also found that pregnancies complicated with HG less often involved macrosomia and stillbirth. Although we were unable to investigate underlying mechanisms, poor nutritional status among mothers with HG is likely to have played a role our findings. It is unknown to which degree adequate and timely HG treatment could lead to improvements of these detrimental effects.

INTRODUCTION

The majority of women, up to 80%, experience nausea and vomiting during pregnancy (NVP).[1] Women falling at the severe end of this clinical spectrum may be diagnosed with hyperemesis gravidarum (HG). Recently, an international HG definition was developed, which defines HG as a condition that starts early in pregnancy, before a gestational age of 16 weeks, and is characterized by severe nausea and/or vomiting, inability to eat and/or drink normally and strongly limits daily activities.[2] HG can lead to dehydration, electrolyte disturbances and weight loss.[3,4] Currently there is no cure available for HG, patients receive symptomatic and supporting treatment.[5]

Previous epidemiologic and animal studies have shown that undernutrition and low maternal weight gain in pregnancy can increase the risk of unfavorable perinatal outcomes, including small for gestational age (SGA), low weight at birth or more preterm birth. [6-10] In 2012 a systematic review was published describing the effects of HG on infants, including perinatal outcomes.[11] Since then, a substantial number of studies describing the effects of HG on perinatal outcomes have been published, necessitating an updated estimate of aggregate effects.[12-14]

Recently, a top 10 of the most urgent priorities in HG research was developed by patients and clinicians, which prioritized the investigation of immediate and long term effects of HG on the developing fetus.[15] Therefore, we aimed to update the systematic summary of the available evidence on perinatal outcomes of infants born to mothers with HG.

METHODS

This systematic review was conducted in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and following the A MeaSurement Tool to Assess systematic Reviews (AMSTAR) methodology. The review protocol was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO, CRD42020209554).

Search strategy

A medical information specialist (JL) conducted a systematic literature search in OVID Medline and EMBASE from inception to February 9th, 2022. Search terms for HG or NVP were combined with search terms comprising infant outcomes (perinatal and long term outcomes). There was no language restriction. Editorials, reviews conference and case-reports were excluded (the complete search strategy can be found in **Appendix A)**. All references were imported in ENDNOTE (version X9.3.3) and duplicates were removed. Reference lists and citing references of relevant papers were checked using Web of Science.

Protocol deviations

The Risk Of Bias in Non-randomized Studies - of Exposure (ROBINS-E) tool was not used, due to the fact that the tool remained 'under development' during the review process. Furthermore, we did not search the Cochrane Central Register of Controlled Trials (CENTRAL), as we mainly expected evidence from published observational studies and any RCT would also be found in Medline and EMBASE.

Eligibility criteria

Studies reporting on perinatal outcomes of infants born to mothers with HG or severe NVP were included. Case reports, case series, animal studies, reviews, editorials and conference abstract were excluded. Of articles that reported on the same study population/data, only the parent study was included.

Study selection

Two reviewers (KN and LJ) independently screened the records on title and abstract, according to the eligibility criteria, using Rayyan.[16] Potential eligible studies were obtained in full text. A second eligibility check was performed independently by both reviewers (KN and LJ) for studies in full text. Hereafter a third reviewer (RP) was consulted in case of disagreement about the in- or exclusion of articles between the first two reviewers (KN and LJ).

Data collection process

One reviewer (LJ) extracted the data using a piloted data extraction form which was critically appraised by a second reviewer (KN). In case of missing or unclear data authors were contacted by email or by letter.

Assessment of risk of bias

The Newcastle-Ottawa Scale (NOS) was used to perform a quality assessment of included studies, and consists of 8 questions with a maximum score of 9.[17] Studies scoring ≥7 were considered good quality, ≥5 fair quality and ≤4 poor quality. We did not exclude studies with low quality. The quality of each study was independently assessed by two reviewers (KN and LJ), in case of disagreement a third reviewer was consulted (RP).

Data synthesis

Meta-analysis was conducted if sufficient data were available, otherwise data were described narratively. Data were presented as odds ratios (OR) with corresponding 95% confidence intervals (95% CI). We estimated Mantel-Haenszel odds ratio using a random effect model, as we anticipated substantial levels of heterogeneity. To assess heterogeneity, I^2 statistics were used: an I^2 value >75% was considered high heterogeneity. If at least 10 studies reported on the same outcome, publication bias was assessed by analyzing funnel plots, which is in line with the Cochrane Handbook for Systematic Reviews of Interventions.[18] Meta-analysis were performed using Review Manager (RevMan, version 5.4, The Cochrane Collaboration, 2020). A p-value below 0.05 was considered statistically significant.

Subgroup analyses

In case of high heterogeneity, we performed sensitivity analyses for European studies compared to studies conducted elsewhere, and case control versus cohort studies. To assess whether more recent improvements in HG treatment could have impacted HG's effects on perinatal outcomes, a subgroup analysis was performed including only studies performed after 2001 of which HG-exposed were compared with non-exposed (since from 2000 several patient organizations for HG were established, which advocated for HG awareness among policy makers and health professionals). To assess the possibility of HG severity affecting perinatal outcomes, subgroup analyses were performed for hospitalization and, if studies made a distinction between mild and severe HG, for mild and severe HG.

RESULTS

Study selection

Our search identified 1387 unique studies of which seventy studies were considered eligible based on full text screening, as is shown in **Figure 1**.[12, 19-86] Nine were excluded, four because of insufficient data[20, 22, 50, 58] and five because they reported on the same study population (**Appendix B**).[39, 57, 67, 69, 75] In total, 61 studies were included in this systematic review. Cited and citing reference searches did not reveal additional studies.

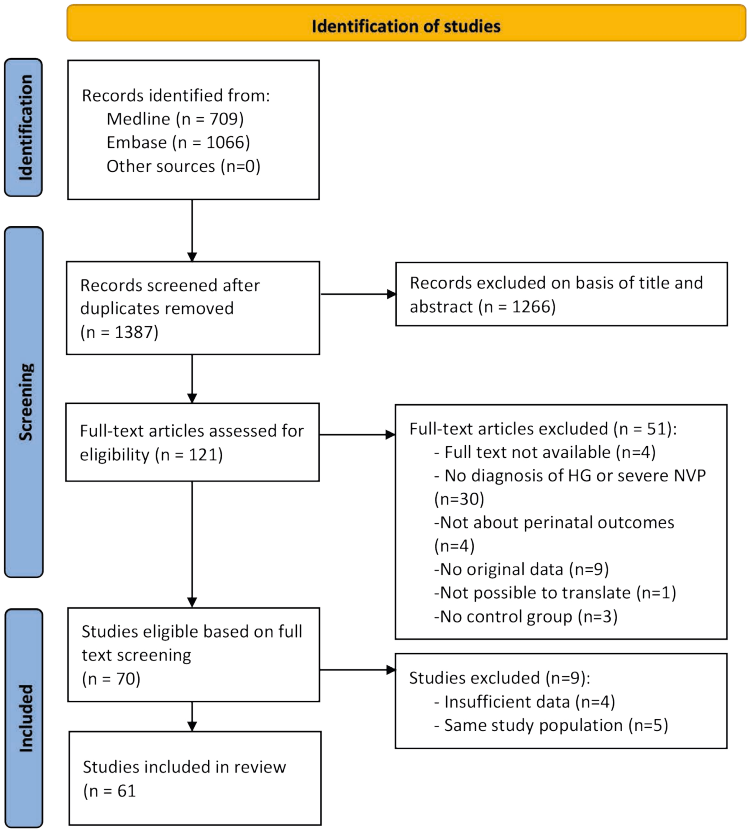


Figure 1. PRISMA flow diagram: selection process of articles

Study characteristics

The main characteristics of the included studies are summarized in **Table 1.** Two studies validated their results in another study population, these two populations were described separately.[37, 84] Forty-eight cohort studies and thirteen case-control studies were included in this systematic review.

Table 1. Study characteristics

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
1	Agmon *et al.* (2019)	Israel	Retrospective cohort	2013-2016	ICD-9	89/190	Apgar score, asphyxia and pH Birth weight Gestational age at delivery Infant deaths Placental abruption Preterm birth	
2	Askling *et al.* (1999)	Sweden	Retrospective cohort	1987-1995	ICD-9	8,186/1,019,027	Fetal sex	
3	Bailit *et al.* (2005)	USA	Retrospective cohort	1999	ICD-9	2,466/520,739	Birth weight Gestational age at delivery Infant deaths Multiple gestations	
4	Bashiri *et al.* (1995)	Israel	Retrospective case-control	1985-1988	*From medical records:* Diagnostic criteria Fairweather	164/373	Birth weight Congenital anomalies Fetal sex Gestational age at delivery Infant deaths Multiple gestations	The fetal loss rate in the control group was not reported and could therefore not be included in corresponding tables and meta-analysis.
5	Basso *et al.* (2001)	Denmark	Retrospective birth cohort	1980-1996	*From medical records:* Admitted with HG	6,084/82,888	Fetal sex Multiple gestations	
6	Bayraktar *et al.* (2021)	Turkey	Retrospective cohort	2016-2019	*From medical records:* nausea and vomiting combined with ketone in urine, >5% of body weight loss, or severe nausea and vomiting that could limit fluid intake	95/947	Fetal sex	
7	Boissière-O'Neill *et al.* (2021)	Canada	Retrospective cohort	1990-2016	*From medical records:* HG severe enough to require contact with care providers	17,512/ 2,115,581	Congenital anomalies	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
8	Bolin *et al.* (2013)	Sweden	Retrospective population based cohort	1997-2009	ICD-10	12,270/ 1,155,033	Birth weight Infants deaths Placental abruption	
9	Buyukkayaci *et al.* (2015)	Turkey	Prospective case-control	Not mentioned	*Self-reported:* Pregnant volunteers diagnosed with HG	50/100	Birth weight Congenital anomalies Infant deaths Preterm birth	No additional data received for preterm birth. So preterm birth not included in tables and meta-analysis.
10	Caltekin *et al.* (2021)	Turkey	Retrospective case-control	2017-2019	*From medical records* More than 2 severe vomiting episodes, the presence of ketonuria in a random urine sample, and weight loss of more than 5% of body weight	52/112	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery	
11	Chin *et al.* (1988)	China	Prospective cohort	1985-1986	*From medical records:* Diagnostic criteria Fairweather	72/8,874	Birth weight Gestational age at delivery	
12	Chin *et al.* (1989)	China	Prospective cohort	1988	*From medical records:* Diagnostic criteria Fairweather	5/1,453	Apgar score, asphyxia and pH Birth weight Gestational age at delivery Infant deaths Placental abruption Preterm birth	
13	Coetzee *et al.* (2011)	New Zealand	Matched cohort	2003-2005	ICD-10	75/217	Apgar score, asphyxia and pH Birth weight Congenital anomalies Fetal sex Gestational age at delivery New-born measurements Infant deaths Neonatal complications NICU admission Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
14	Coffey *et al.* (1957)	Ireland	Retrospective case-control	1953-1955	*Self-reported:* Interview, excessive vomiting	48/296	Congenital anomalies	
15	Czeizel *et al.* (2003)	Hungary	Population based matched case-control	1980-1996	ICD-9	217/2,742	Congenital anomalies	
16	Czeizel *et al.* (2005)	Hungary	Retrospective case-control	1980-1996	*From medical records:* Prenatal care logbooks and a structured questionnaire or a personal interview	125/30,281	Congenital anomalies	
17	Del Mar *et al.* (2000)	United Kingdom	Retrospective cohort	1992-1996	ICD-8	4,126/127,647	Fetal sex	
18	Depue *et al.* (1987)	USA	Study 1: retrospective case-control	1958-1965	*From medical records:* usual clinical practice	419/1,255	Congenital anomalies Infant deaths	
			Study 2: prospective case-control	1958-1965	*From medical records:* usual clinical practice	35/70	Birth weight	
19	Dodds *et al.* (2006)	Canada	Retrospective cohort	1988-2002	*From medical records:* Admitted with HG before 24 weeks of gestation	1,270/156,091	Apgar score, asphyxia and pH Birth weight Fetal sex Infant deaths Preterm birth	
20	Fejzo *et al.* (2013)	USA	Self-selected online survey cohort	2007-2011	*Self-reported:* HG diagnosis in their first pregnancy and treatment with IV fluids or TPN/nasogastric feeding tube	254/562	Birth weight Congenital anomalies Fetal sex Infant deaths Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
21	Fejzo *et al.* (2015)	USA	Self-selected online survey cohort	2007-2011	*Self-reported:* HG diagnosis with treatment with IV fluids or TPN/nasogastric feeding tube	312/481	Fetal sex Preterm birth	
22	Fiaschi *et al.* (2017)	United Kingdom	Population-based retrospective cohort study	1997-2011	ICD-10	118,197/ 8,211,850	Birth weight Fetal sex Infant deaths Multiple gestations Neonatal complications NICU admission Placental abruption Preterm birth	Gestational week is divided in categories and therefore not included in tables and meta-analysis
23	Getahun *et al.* (2019)	USA	Retrospective cohort	1991-2014	ICD-9	14,526/469,789	Fetal sex Preterm birth	
24	Grooten *et al.* (2017)	The Netherlands	Prospective population based cohort	2002-2006	*Self-reported:* Daily vomiting for more than 3 months	601/5,549	Birth weight Gestational age at delivery Multiple gestations Preterm birth	Because of overlap in data with Roseboom 2011 birth weight and preterm birth are not included in tables and meta-analysis.
25	Gu *et al.* (2021)	China	Prospective cohort	2011-2018	*Self-reported:* face-to-face interviews. People were asked if they experienced NVP and how severe this was (mild, moderate of severe)	232/779	Birth weight Fetal sex Gestational age at delivery Preterm birth	No additional data received for birth weight, gestational age at delivery and preterm birth. So these are not included in tables and meta-analysis.
26	Gunay *et al.* (2020)	Turkey	Retrospective cohort	2015-2018	*From medical records:* Hospitalized with HG in the first 20 weeks of pregnancy	186/386	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery Infant deaths Placental abruption Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
27	Hallak *et al.* (1993)	USA	Retrospective cohort	1984-1991	*From medical records:* Diagnostic criteria Fairweather	138/12,473	Apgar score, asphyxia and pH Birth weight Congenital anomalies Gestational age at delivery NICU admission Preterm birth	
28	Hastoy *et al.* (2013)	France	Retrospective cohort	2006-2009	*From medical records:* Hospitalized with HG in the first 22 weeks of pregnancy	197/589	Birth weight Fetal sex Infant deaths Preterm birth	
29	Hohlbein *et al.* (1961)	Germany	Retrospective cohort	1949-1959	*From medical records:* Severe hyperemesis early in pregnancy	120/28,120	Congenital anomalies	
30	Hsu *et al.* (1993)	USA	Retrospective cohort	1980-1990	*From medical records:* NICU admission for severe hyperemesis	66/20,864	Fetal sex	
31	Kidess *et al.* (1974)	Germany	Retrospective cohort	1967-1971	*From medical records:* NICU admission for hyperemesis	65/130	Apgar score, asphyxia and pH Birth weight Gestational age at delivery New-born measurements Preterm birth	Not defined which Apgar score was presented (i.e. at 1 or 5 minutes postpartum), so Apgar score was not included in tables and meta-analysis
32	Klebanoff *et al.* (1985)	USA	Prospective cohort	1959-1966	Not described	188/3,690	Birth weight Infant deaths Preterm birth	
33	Koot *et al.* (2017)	The Netherlands	Prospective cohort	1985-1986	ICD-8	62/8,953	Apgar score, asphyxia and pH Birth weight Fetal sex Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
34	Koudijs *et al.* (2016)	Indonesia	Prospective cohort	2012-2014	*From medical records:* Diagnosed with HG by midwives during routine visits in the first or second trimester	400/2,233	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery Infant deaths Multiple gestations	
35	Kruse *et al.* (1975)	Germany	Retrospective cohort	Missing	Not described	306/12,091	Fetal sex	
36	Kuru *et al.* (2012)	Turkey	Retrospective cohort	2003-2010	*From medical records:* One or more antepartum hospitalizations for HG, the first of which had to have occurred before 24 weeks of gestation	72/161	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery Infant deaths Preterm birth	
37	Mitsuda *et al.* (2018)	Japan	Retrospective cohort	2011-2014	*Self-reported:* Questionnaires, severe NVP those who experienced NVP and could not have meals	10,518/27,114	Gestational age at delivery Preterm birth	
38	Mitsuda *et al.* (2019)	Japan	Retrospective cohort	2011-2014	*Self-reported:* Questionnaires, severe NVP those who experienced NVP and could not have meals	10,159/25,997	Fetal sex Multiple gestations	
39	Morokuma *et al.* (2016)	Japan	Retrospective cohort	2011-2014	*Self-reported:* Questionnaires, HG those who experienced NVP and could not have meals and had lost >5% of prepregnancy weight	136/6,529	Birth weight	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
40	Muraoka *et al.* (2020)	Japan	Retrospective cohort	2009-2012	*From medical records:* Diagnosed with HG requiring hospital NICU admission	35/103	Birth weight Fetal sex Gestational age at delivery	Fetal sex was only provided as ratio. Since no additional data was received, fetal sex was not included in corresponding tables and meta-analysis
41	Nurmi *et al.* (2020)	Finland	Retrospective cohort	2005-2017	ICD-10	9,549/733,002	Fetal sex Multiple gestations	
42	Ong *et al.* (2021)	Malaysia	Prospective cohort	2009-2010	*Self-reported and medical records:* self-reported responses of retrospective recall alongside verification from hospitalization data in medical records	190/486	Birth weight Fetal sex Gestational age at delivery NICU admission Preterm birth	
43	Ozay *et al.* (2019)	Turkey	Retrospective cohort	2015-2017	*From medical records:* Diagnosed with HG in the first trimester of pregnancy: persistent vomiting and severe intolerance in terms of eating and drinking, dehydration development, electrolyte imbalance or metabolic disorder, or at least 5% or more weight loss.	46/100	Birth weight Fetal sex Gestational age at delivery Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
44	Paauw *et al.* (2005)	USA	Prospective cohort	1995-1998	*From medical records:* Diagnostic criteria Fairweather	45/351	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery NICU admission Preterm birth	Not defined which Apgar score was presented, so Apgar score not included in tables and meta-analysis
45	Peled *et al.* (2014)	Israel	Retrospective cohort	1997-2011	*From medical records:* Hospitalization for HG during the first trimester	599/2,396	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery Infant deaths Neonatal complications NICU admission Placental abruption Preterm birth	
46	Rashid *et al.* (2012)	United Kingdom	Retrospective cohort	2003-2007	*From medical records:* HG defined as a patient presenting with nausea and vomiting severe enough to cause ketonuria before 24 weeks gestation.	184/9,955	Fetal sex	
47	Roseboom *et al.* (2011)	The Netherlands	Population-based retrospective cohort	2000-2006	*From medical records:* Those pregnancies that were labeled during any time of pregnancy as complicated by HG by the caregiver	2,190/1,199,218	Apgar score, asphyxia and pH Birth weight Fetal sex Infant deaths NICU admission Preterm birth	
48	Salunkhe *et al.* (2021)	India	Prospective cohort	2013-2015	*From medical records:* not described further	76/1,035	Birth weight Gestational age at delivery Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
49	Schiff *et al.* (2004)	USA	Population-based retrospective case-control	1987-1996	ICD-9	2,110/11,893	Birth weight Fetal sex Preterm birth	
50	Sorenson *et al.* (2000)	Denmark	Population-based retrospective cohort	1991-1998	ICD-8 and ICD-10	650/47,931	Fetal sex	
51	Tan *et al.* (2006)	Malaysia	Retrospective case-control	2003-2005	*From medical records:* Hospitalisaton for HG	166/4,927	Fetal sex	
52	Tan *et al.* (2007)	Malaysia	Retrospective case-control	2004-2005	*From medical records:* Hospitalisaton for HG	166/664	Apgar score, asphyxia and pH Birth weight Gestational age at delivery Infant deaths Preterm birth	
53	Tsang *et al.* (1996)	USA	Retrospective cohort	1988-1994	*From medical records:* HG defined as excessive and persistent vomiting resulting in dehydration and requiring NICU admission or extensive medical therapy	193/13,053	Apgar score, asphyxia and pH Birth weight Congenital anomalies Fetal sex Infant deaths Gestational age at delivery Multiple gestations Preterm birth	
54	Ustun *et al.* (2004)	Turkey	Prospective cohort	2 year period	*From medical records:* HG was defined as weight loss of at least 2.25 kg or ketonuria >80 mg/dl or hypokalemia or hyponatremia requiring intravenous replacement or two or more visits to the obstetric emergency department for hyperemesis	35/74	Infant deaths	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
55	Vandraas *et al.* (2013)	Norway	Population based retrospective cohort	1967-2009	ICD-8 and ICD-10	20,004/ 2,266,345	Apgar score, asphyxia and pH Birth weight Gestational age at delivery Infant deaths Preterm birth	
56	Vikanes *et al.* (2013)	Norway	Prospective cohort	1998-2008	*From medical records:* HG was defined as long-lasting nausea and vomiting in pregnancy starting before 25 weeks of gestation and necessitating hospitalization	814/70,654	Apgar score, asphyxia and pH Birth weight Gestational age at delivery Infant deaths Preterm birth	
57	Vlachodimitropoulou *et al.* (2013)	United Kingdom	Retrospective case-control	2007-2010	*From medical records:* Diagnostic criteria Fairweather	208/416	Apgar score, asphyxia and pH Birth weight Fetal sex Gestational age at delivery Preterm birth	
58	Vilming *et al.* (2000)	Norway	Retrospective cohort	1993-1997	ICD-9	120/235	Birth weight Fetal sex Gestational age at delivery	
59	Wang et al. (2020)	USA Denmark	National based longitudinal cohort study	Not shown	*Self-reported:* Structured maternal interview. NVP with either weight loss or prolonged disease course (continued after gestation of 6 months)	1,496/10,710 21,282/ 2,092,897	Birth weight Fetal sex Preterm birth	
60	Zhang *et al.* (1991)	China	Retrospective case-control	1986-1987	*Self-reported:* Questionnaires, sever vomiting	201/1,867	Birth weight Fetal sex Preterm birth	

Table 1. *Continued*

	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Comments
61	Zhang *et al.* (2017)	China	Retrospective case-control	2002-2014	*Self-reported:* Questionnaires, severe NVP is NVP in the first trimester of pregnancy that continued or became aggravated, was not limited to the mornings, and necessitated bed rest or hospitalization	419/2,103	Congenital anomalies	

HG diagnosis was based on ICD codes in 15 studies,[19, 21, 23, 27, 32, 35, 36, 42, 43, 53, 63, 73, 74, 80, 82] on medical records in 31 studies[12, 24-26, 29-31, 34, 37, 38, 46-49, 51, 54, 56, 62, 65, 66, 68, 70-72, 76-79, 81, 83, 87] and self-reported in 12 studies.[28, 33, 40, 41, 44, 45, 59-61, 84-86] There was one study in which HG diagnosis was self-reported and based on medical records.[64] Two studies did not describe how HG was diagnosed.[52, 55] In total, 20,532,671 infants were included in this systematic review of which 270,873 were born to mothers with HG.

Risk of bias of included studies

Of the 61 included studies, six cohort studies and two case-control studies were rated to be of poor quality, as shown in **Table 2.**[24, 29, 40, 41, 48, 49, 55, 79] These studies mainly scored low on the *selection* domain, because of including self-selected participants, and on the *comparability* domain, because they did not adjust for confounders in statistical analysis. Twenty-one cohort studies and four case-control studies were rated to be of fair quality.[12, 19, 21, 23, 28, 30, 31, 33, 36, 37, 43, 46, 51, 53, 56, 61, 64, 65, 70, 72, 74, 78, 82, 83] Twenty-eight studies (21 cohort and seven case-control) were rated to be of good quality.[25-27, 32, 34, 35, 38, 42, 44, 45, 47, 52, 54, 59, 60, 62, 63, 66, 68, 71, 73, 76, 77, 80, 81, 84-86] As displayed in **Appendix C**, funnel plots for birth weight, birthweight <2500 grams, SGA and congenital anomalies did show asymmetry, which could be a sign of publication bias.[88]

Table 2. Risk of bias assessment using the Newcastle-Ottawa Quality Assessment Scale (NOS) of included cohort and case-control studies

	Cohort studies				
Studies	**Selection**	**Comparability**	**Outcome**	**Total score**	**Quality score**
Agmon *et al*, 2019	***		***	6	Fair
Askling *et al*, 1999	***		***	6	Fair
Bailit *et al*, 2005	***		***	6	Fair
Basso *et al*, 2001	***	*	***	7	Good
Bayraktar *et al*, 2021	***		***	6	Fair
Boissière-O'Neill *et al*, 2021	***	**	***	8	Good
Bolin *et al*, 2013	***	**	***	8	Good
Chin *et al*, 1988	***		***	6	Fair
Chin *et al*, 1989	***		**	5	Fair
Coetzee *et al*, 2011	***	**	***	8	Good
Del Mar *et al*, 2000	***		***	6	Fair
Dodds *et al*, 2006	***	**	***	8	Good
Fejzo *et al*, 2013	*		**	3	Poor
Fejzo *et al*, 2015	*		**	3	Poor
Fiaschi *et al*, 2017	***	**	***	8	Good
Getahun *et al*, 2019	***		***	6	Fair
Grooten *et al*, 2017	***	**	***	8	Good
Gu *et al*, 2021	***	**	**	7	Good
Gunay *et al*, 2020	***		***	6	Fair
Hallak *et al*, 1993	***		***	6	Fair

Table 2. *Continued*

Hastoy *et al*, 2013	***	**	***	8	Good
Hohlbein *et al*, 1961	*		**	3	Poor
Hsu *et al*, 1993	**		**	4	Poor
Kidess *et al*, 1974	***		***	6	Fair
Klebanoff *et al*, 1985	****	**	*	7	Good
Koot *et al*, 2017	***		***	6	Fair
Koudijs *et al*, 2016	***	**	***	8	Good
Kruse *et al*, 1975	*		***	4	Poor
Kuru *et al*, 2012	***		***	6	Fair
Mitsuda *et al*, 2018	**	**	***	7	Good
Mitsuda *et al*, 2019	**	**	***	7	Good
Morokuma *et al*, 2016	**	**	**	6	Fair
Muraoka *et al*, 2020	***	**	***	8	Good
Nurmi *et al*, 2020	***	**	***	8	Good
Ong *et al*, 2021	**	**	**	6	Fair
Ozay *et al*, 2019	***		***	6	Fair
Paauw *et al*, 2005	***	**	***	8	Good
Peled *et al*, 2014	***	**	***	8	Good
Rashid *et al*, 2012	***		***	6	Fair
Roseboom *et al*, 2011	***	**	***	8	Good
Salunkhe *et al*, 2021	**		***	5	Fair
Sorenson *et al*, 2000	***		***	6	Fair
Tsang *et al*, 1996	***		***	6	Fair
Ustun *et al*, 2004	**		**	4	Poor
Vandraas *et al*, 2013	***	**	***	8	Good
Vikanes *et al*, 2013	**	**	***	7	Good
Vilming *et al*, 2000	***		**	5	Fair
Wang *et al*, 2020	***	**	***	8	Good

Case-control studies

Studies	**Selection**	**Comparability**	**Exposure**	**Total score**	**Quality score**
Bashiri *et al*, 1995	**		**	4	Poor
Buyukkayaci *et al*, 2015	***		**	5	Fair
Caltekin *et al*, 2021	**		**	4	Poor
Coffey *et al*, 1957	***		**	5	Fair
Czeizel *et al*, 2003	****	**	**	8	Good

Table 2. *Continued*

Czeizel *et al*, 2005	****	**	**	8	Good
Depue *et al*, 1987	**	*	**	5	Fair
Schiff *et al*, 2004	***	*	***	7	Good
Tan *et al*, 2006	***	*	***	7	Good
Tan *et al*, 2007	****	**	***	9	Good
Vlachodimitropoulou *et al*, 2013	**	**	**	6	Fair
Zhang *et al*, 1991	****	**	**	8	Good
Zhang *et al*, 2017	***	**	**	7	Good

The NOS risk of bias assessment tool consisted of 8 items with a total maximum score of 9. A score ≥7 was considered as good quality, a score ≥5 as fair quality and a score ≤4 as poor quality.

Apgar score, asphyxia and pH at birth

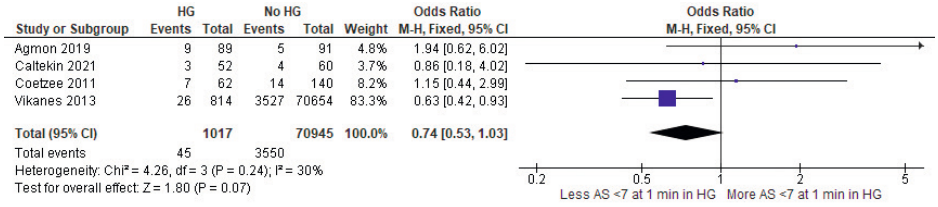
There were fifteen studies (n=3,797,409) reporting on Apgar score, see **Appendix D.**[12, 19, 29, 32, 38, 46, 53, 54, 56, 68, 71, 78, 80, 81, 83] Two studies (n=13,469) reported the mean Apgar score at 1 minute[78, 83] and three studies (n=15,707) the mean Apgar score at 5 minutes.[54, 78, 83] These five studies showed no difference between HG and control pregnancies. Studies could not be pooled in meta-analysis due to missing data.

Four studies (n=71,962) assessed Apgar scores <7 at 1 minute.[19, 29, 32, 81] Meta-analysis of these four studies showed no statistically significant association between HG and Apgar scores <7 at 1 minute (OR 0.74 95%CI:0.53-1.03, p=0.07, I^2=30%, **Figure 2.1)**. Twelve studies reported on Apgar score <7 at 5 minutes.[12, 19, 29, 32, 38, 46, 53, 56, 68, 71, 80, 81] Meta-analysis (**Figure 2.2**) showed no statistically significant association between HG-exposure and Apgar score <7 at 5 minutes (12 studies, n=3,712,472, OR 1.12, 95%CI:1.00-1.26, p=0.06, I^2=0%). One cohort study (n=180) reported on cord blood pH at birth. No differences between HG-exposed and non-exposed infants were found **(Appendix D).**[19]

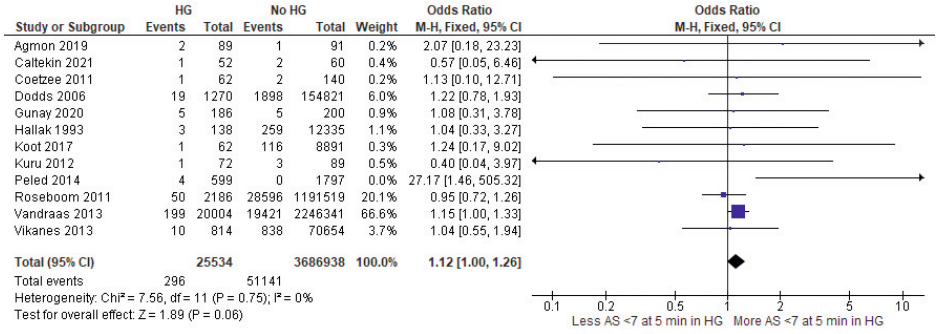
One study (n=366) reported on asphyxia which was defined as mild in case of an Apgar score between 4 and 6 and severe in case of an Apgar score less than 4.[30] Only 5 HG-exposed infants were included, and none of them were diagnosed with asphyxia.

Figure 2. Forest plots Apgar score

2.1 Forest plot Apgar score <7 at 1 minute



2.2 Forest plot Apgar score <7 at 5 minutes



Birth weight

Birth weight was reported in 25 studies (n=4,063,406), see **Appendix E**.[12, 19, 23, 24, 29-32, 37, 40, 46, 51, 53, 54, 56, 62, 65, 66, 68, 71, 72, 77, 80-82] Eight of these studies showed lower birth weight at birth in HG-exposed infants compared with non-exposed infants.[23, 29, 31, 40, 66, 68, 71, 82] Eleven studies could not be included in the meta-analysis, eight studies due to missing information on SDs [12, 23, 24, 32, 40, 80-82] and three studies because they did not report mean birth weights and SDs for the total HG group.[31, 46, 54] Meta-analysis (14 studies, n=1,211,650) showed high heterogeneity with an I^2 of %92 (**Figure 3.1).** In sensitivity analyses for study design the heterogeneity remained high (I^2>75%, **Appendix F.1)**. Sensitivity analyses for only European studies revealed lower birth weight in the HG-exposed group (mean difference -111.79 gram, 95%CI: -135.94--87.63, p<0.00001, I^2=68%, **Appendix F.2**).

Birth weight centile was reported in two cohort studies (**Appendix E)**.[64, 68] One of these studies showed lower birth weight centile in the HG-exposed group compared to the non-exposed

group (HG-exposed: 44.8±28.3 centile, HG non-exposed: 52.4±27 centile, p<0.001).[68] Meta-analysis was not possible.

Low birthweight (below 2500 or 1500 grams)

Sixteen studies (n=6,433,944) reported on HG and birth weight <2500 grams.[23, 28, 30, 32, 38, 42, 47, 52, 54, 65, 72, 73, 77, 81, 84, 85] Meta-analysis including these sixteen studies showed high heterogeneity (I%77=^{2}, **Figure 3.2**). Sensitivity analysis including only cohort studies (**Appendix F.3**) showed persistently high heterogeneity (I%89=^{2}) while the sensitivity analysis including case-control studies (**Appendix F.4**) showed birth weight <2500 grams more often in HG exposed group (four studies, n=14.524: OR 1.32, 95%CI:1.09;1.61, p=0.005, I^2=0%). Sensitivity analyses including only European studies revealed no associations **(Appendix F.5).**

The frequency of a birth weight <1500 grams was reported in two cohort studies.[23, 30] Meta-analysis of these two studies (n=489,141) showed a positive association between HG and birth weight <1500 grams (OR1.43, 95%CI:1.02-1.99, p= 0.04, I%48=^{2}, **Figure 3.3**).

SGA

SGA was reported in nineteen studies (n=1,208,478). SGA was described as birth weight <10th centile in eighteen studies and in one study as <2 SD below mean birth weight.[12, 19, 23, 27, 32, 38, 40, 42, 47, 54, 56, 61, 64, 68, 71, 80, 81, 83, 85] Meta-analysis showed high heterogeneity (nineteen studies, I%93=^{2}, **Figure 3.4).** Sensitivity analysis including cohort studies and European studies (**Appendix F.6 and F.7**) did not result in a reduction of the high heterogeneity (I^2=94 and I^2=81%).

Birth weight >4000 grams

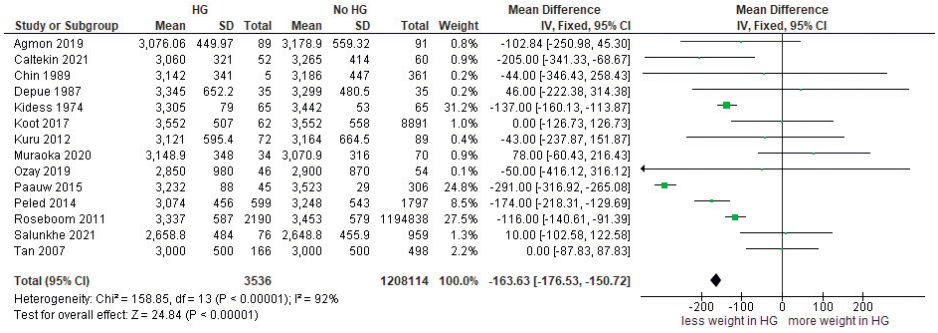
Birth weight >4000 grams was reported in two studies (n=5,503,120).[42, 73] Meta-analysis showed that HG-exposed infants were less likely to have weigh >4000 grams at birth (OR 0.74, 95%CI:0.72-0.76, p<0.00001, I%44=^{2}, **Figure 3.5**).

Large for gestational age (LGA)

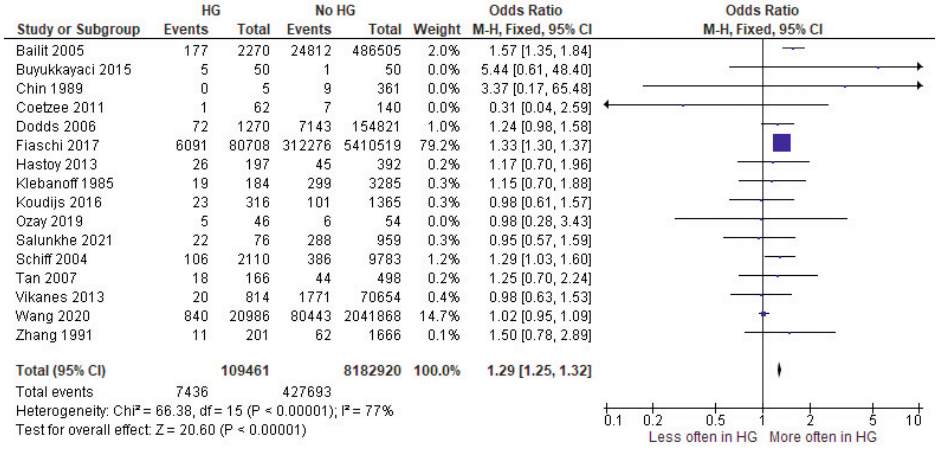
Six studies (n=8,249,645) described infants large for gestation age (LGA), which was defined as a birth weight >90th centile.[23, 42, 64, 68, 80, 83] Meta-analysis (six studies) showed high heterogeneity (I^2=100%, **Figure 3.6**). Heterogeneity remained high in sensitivity analysis that included only cohort studies (I^2=100%, **Appendix F.8**).

Figure 3. Forest plots birth weight

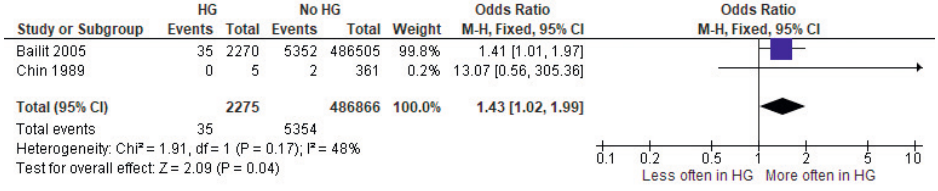
3.1 Birth weight in general



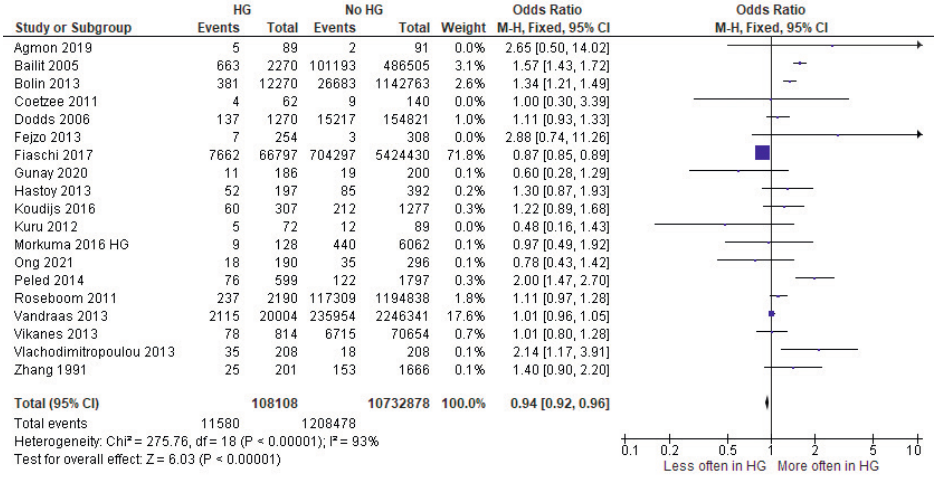
3.2 Birth weight <2500 grams



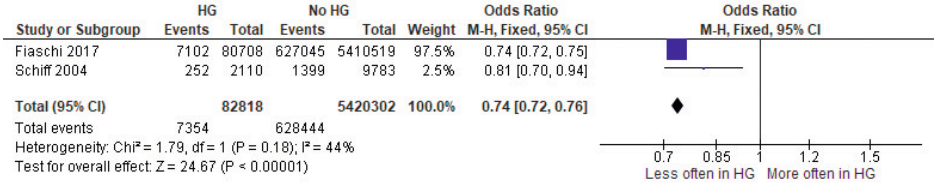
3.3 Birth weight <1500 grams



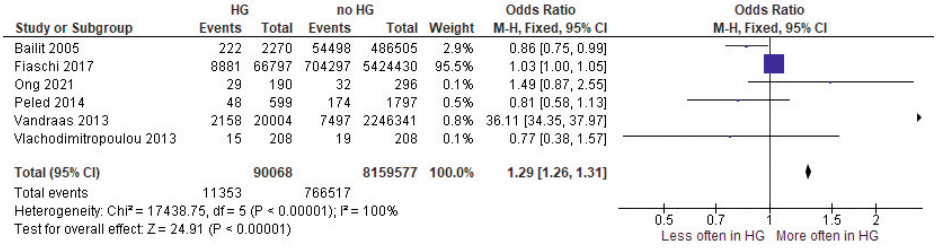
3.4 Small for gestational age



3.5 Birth weight >4000 grams



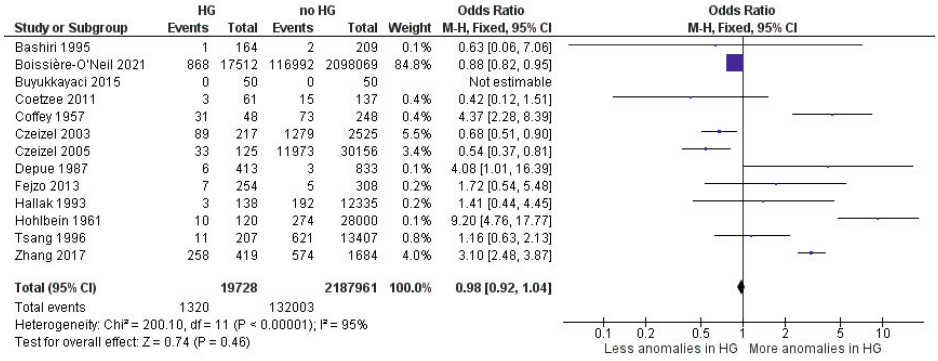
3.6 Large for gestational age



Congenital anomalies

Thirteen studies (n=2,207,689) reported on congenital anomalies (**Appendix G)**.[24, 26, 28, 32-35, 37, 40, 46, 48, 78, 86] Studies were pooled in meta-analysis which showed high heterogeneity (I%95=[2], **Figure 4)**. Sensitivity analyses for study design (cohort/case-control studies) and European origin did not result in reductions in high heterogeneity (I^2 of 90%, 95% and 97%**, Appendix H.1-H.3).**

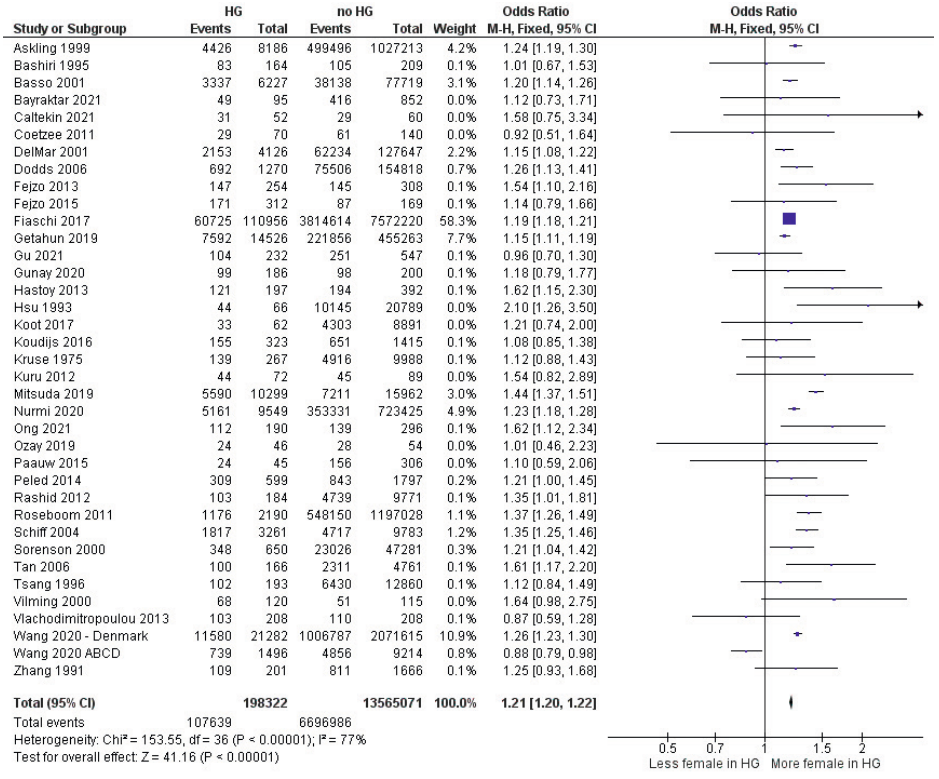
Figure 4. Forest plot Congenital anomalies



Fetal sex

Fetal sex was reported in 36 studies (n=13,565,286, **Appendix I**).[12, 21, 24, 25, 29, 32, 36, 38, 40-43, 45, 47, 49, 53-56, 59, 63-66, 68, 70, 71, 73, 74, 76, 78, 82-85, 87] Meta-analysis (**Figure 5)** had high heterogeneity (I^2=77%). Sensitivity analysis for cohort studies also had high heterogeneity (I^2=78%, **Appendix J.1**), but sensitivity analysis for case-control studies (OR 1.33, 95%CI:1.24-1.43, p<0.00001, I%39=^{2}, **Appendix J.2**) and European studies (OR 1.25, 95%CI:1.23-1.27, p<0.00001, I%30=^{2}, **Appendix J.3**) all found more female infants were born to mothers with HG.

Figure 5. Forest plot fetal sex



Gestational age at delivery and preterm birth

Twenty-five studies reported on gestational age at delivery (n=2,900,154, **Appendix K**).[12, 19, 23, 24, 29-32, 44, 46, 51, 54, 56, 60, 62, 64-66, 68, 72, 78, 80-83] Five studies reported a significantly shorter gestational age at birth in infants born to mothers with HG.[23, 29, 64, 66, 68] Only nine out of twenty-five studies were included in the meta-analysis due to missing data (no SD available) or different ways of outcome reporting in the studies. Meta-analysis, **Figure 6,** showed high heterogeneity (I%97=^{2}). Sensitivity analysis was not possible.

Preterm birth <37 weeks

Preterm birth <37 weeks was reported in 30 studies (n=11,020,932, **Appendix L**).[12, 19, 30, 32, 38, 40-43, 46, 47, 51-53, 56, 60, 64-66, 68, 71-73, 77, 78, 80, 81, 83-85] Meta-analysis, **Figure 7.1,** showed high heterogeneity (I%79=^{2}). Sensitivity analysis (**Appendix M.1**) for cohort studies had high heterogeneity (I%81=^{2}). Sensitivity analyses for case-control studies and European studies (**Appendix M.2**

and M.3) showed significantly more preterm birth in the HG-exposed group (case-control: OR 1.25, 95%CI:1.04-1.51, p=0.02, I%35=^{2}, European: OR 1.10, 95%CI:1.04-1.16, p=0.001, I%65=^{2}).

Preterm birth <34 weeks

Two studies report on preterm birth <34 weeks (n=2,882, **Appendix L**).[64, 68] When pooling these results in meta-analysis there was more preterm birth <34 weeks in the HG-exposed infants compared to non-exposed infants (OR2.81, 95%CI:1.69-4.67, p<0.0001, I%58=^{2}, **Figure 7.2**).

Preterm birth <32 weeks

Three cohort studies report on preterm birth <32 weeks (n=7,051,100, **Appendix L**).[42, 60, 80] Meta-analysis, **Figure 7.3**, showed high heterogeneity (I%96=^{2}). Sensitivity analysis was not possible.

Preterm birth <28 weeks

One study reported on preterm birth <28 weeks (n=27,042, **Appendix L**).[60] There was no difference in the frequency of preterm birth <28 weeks between the HG-exposed and non-exposed group (p=0.27).

Figure 6. Forest plot gestational age at birth

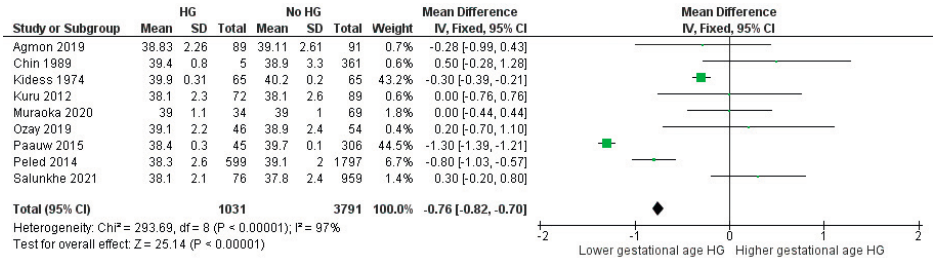
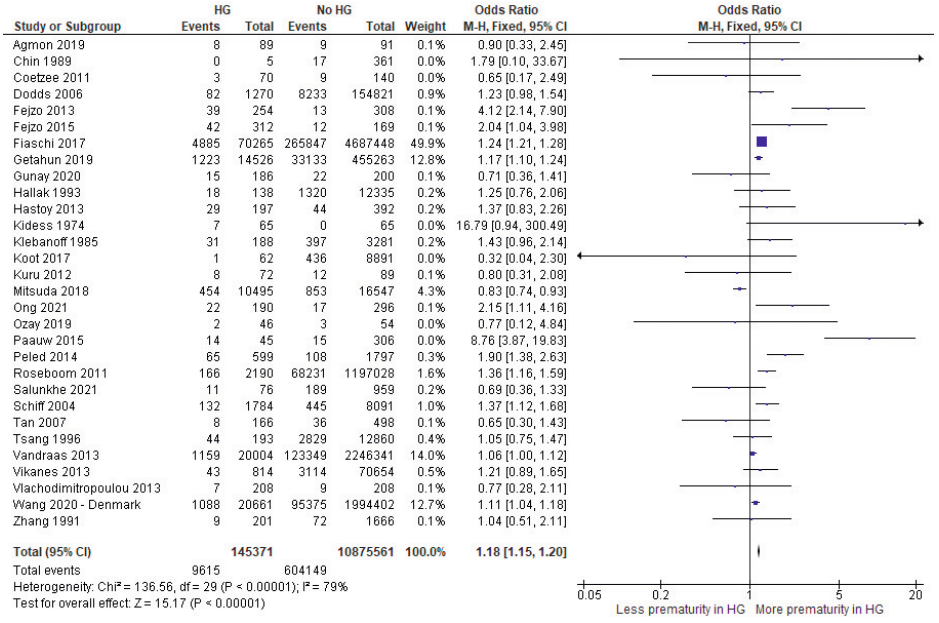
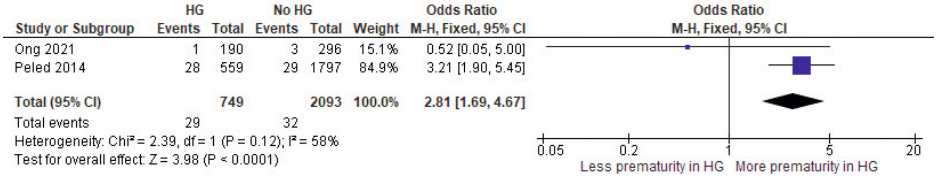


Figure 7. Forest plots preterm birth

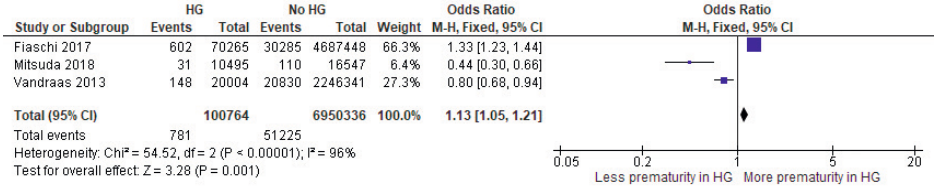
7.1 Preterm birth <37 weeks



7.2 Preterm birth <34 weeks



7.3 Preterm birth <32 weeks



Infant deaths

Fetal loss or fetal death

Eight studies (n=1,724,997) reported on fetal death, see **Appendix N**.[28, 37, 52, 54, 79] Fetal loss was defined in two studies as 'spontaneous abortion,' in the third study as 'fetal loss before 21 weeks', in the fourth study as 'miscarriage between 14 and 19 weeks or stillbirth after 20 weeks' and in the last study as 'fetal loss before 24 weeks or if the foetus weighed less than 500 grams'. The definition of fetal death was described as fetal death after 20 weeks in the first study,[23] in the second study as 'fetal death during gestation (not further described) or labour',[71] and not described in the third study.[78]When pooling these five studies in meta-analysis, we found no association between HG and the combined fetal loss or fetal death rate, see **Figure 8.1** (OR 0.97, 95%CI: 0.75-1.27, p=0.84, I^2=64%). Sensitivity analysis was not possible for European study (there was only one European study). Sensitivity analysis for cohort (**Figure 8.2** OR 1.05, 95%CI: 0.79-1.39, p=0.73, I^2=70%) and case-control studies (**Figure 8.3** OR 0.54, 95%CI: 0.24-1.25, p=0.15, I^2=17%) separately revealed no significant association between HG and fetal loss or fetal death.

Stillbirth

Nine studies reported on stillbirth (n=3,973,154, **Appendix N**).[12, 27, 30, 32, 37, 42, 54, 56, 80] The definition of stillbirth was not described in three studies, the other six studies all used a different definition, see **Appendix N**. Meta-analysis, **Figure 8.4,** showed a significantly lower chance of stillbirth in the HG exposed pregnancies (OR 0.92, 95%CI:0.85-0.99, p=0.02, I%0=^{2}).

Perinatal death

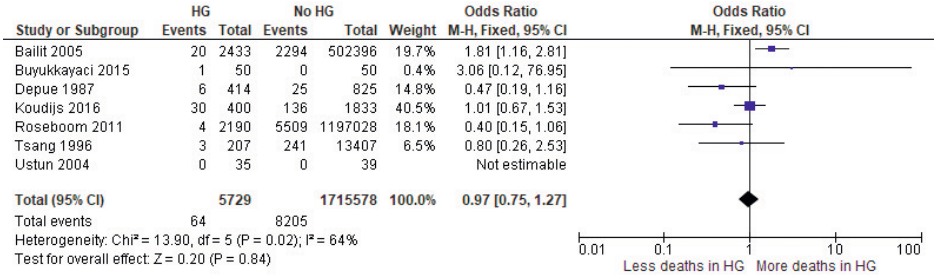
Nine studies reported on perinatal death (n=3,708,440, **Appendix N**).[19, 24, 38, 40, 47, 71, 78, 80, 81] The definition was not described in three studies and varied between the remaining six studies. Meta-analysis **(Figure 8.5)** did not show a significant association between HG and perinatal death (OR 0.97, 95%CI:0.85-1.11, p=0.64, I%30=^{2}).

Neonatal death

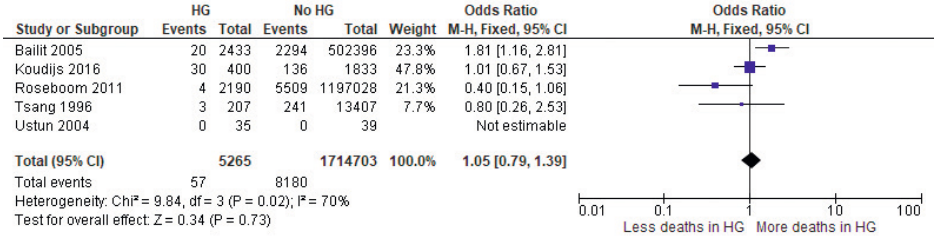
Five studies reported on neonatal death (n=3,973,154, **Appendix N**).[23, 30, 68, 71, 80] The definition of neonatal death was not described in three studies[23, 30, 68], described as death within the first week in one study,[71] and in the last study as death during the first 28 days of life.[80] Meta-analysis, **Figure 8.6,** did not show an association between HG and neonatal death (OR 1.11, 95%CI:0.90;1.35, p=0.33, I%21=^{2}).

Figure 8. Forest plots infant deaths

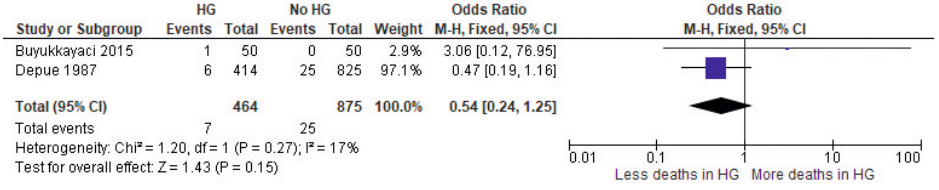
8.1 Fetal loss or fetal deaths



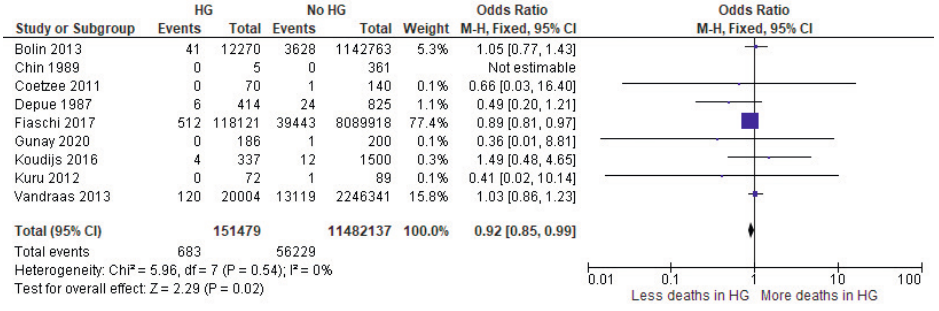
8.2 Fetal loss or fetal deaths only cohort studies



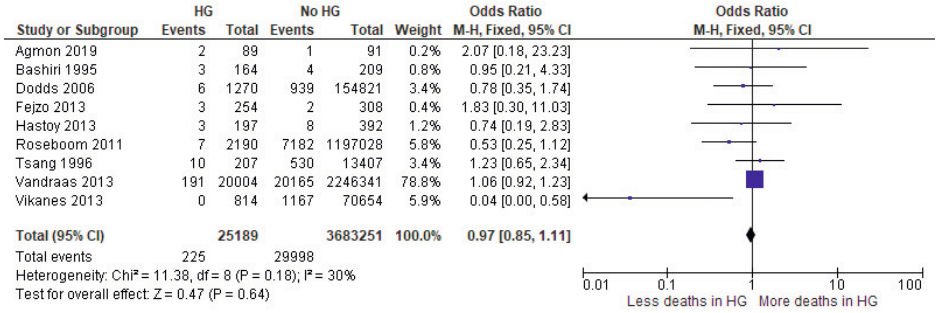
8.3 Fetal loss or fetal deaths only case-control studies



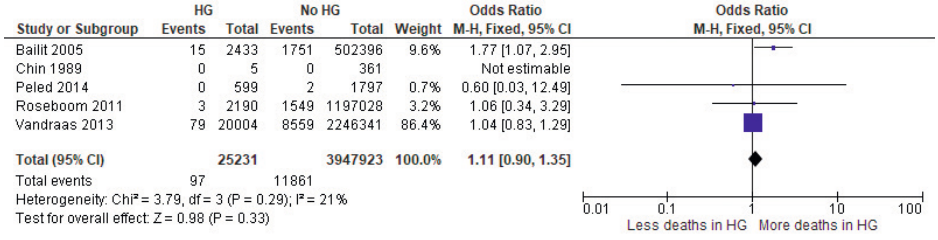
8.4 Stillbirths



8.5 Perinatal deaths



8.6 Neonatal deaths



Multiple gestation

Nine studies reported on multiple gestations (n=9,560,796, **Appendix O**).[23-25, 42, 44, 54, 59, 63, 78] Meta-analysis, **Figure 9.1,** showed high heterogeneity (I%76=^{2}). Sensitivity analysis (**Appendix P.1-P.2**) including cohort studies showed high heterogeneity (I%79=^{2}), but sensitivity analysis including only European studies showed significantly lower heterogeneity and demonstrated more multiple gestations in the HG-exposed group (OR 2.19, 95%CI:1.99-2.42, p<0.00001, I%29=^{2}).

Neonatal complications

One study (n=2,396) reported on four different neonatal complications: respiratory distress syndrome; necrotizing enterocolitis; jaundice requiring phototherapy and hypoglycaemia (**Appendix Q**).[68] Only respiratory distress syndrome was significantly less common in the HG-exposed group (2.7% versus 1.2%, p=0.01).

Two studies reported on resuscitation (n=4,289,344, **Appendix Q**).[32, 42] Meta-analysis (**Figure 9.2**) showed that resuscitation was more often necessary in the HG-exposed infants (OR 1.07, 95%CI:1.05-1.10, p<0.00001, I%0=^{2}).

Neonatal intensive care unit (NICU) admission and length of stay in hospital

Results of studies reporting on NICU admission and length of stay in hospital are shown in **Appendix R.**[32, 42, 46, 64, 66, 68, 71] Meta-analysis including 7 studies (n=6,509,702) showed that there were more NICU admissions in HG-exposed infants compared to non-exposed infants (OR 1.20 95%CI:1.14-1.26, p<0.00001, I^2=75%), see **Figure 9.3.**

Only one study(n=351) reported on the length of stay in hospital and showed a longer mean stay of 2.9 days in the HG-exposed infants compared to a mean stay of 1.8 days in the non-exposed infants.[66]

Newborn measurements

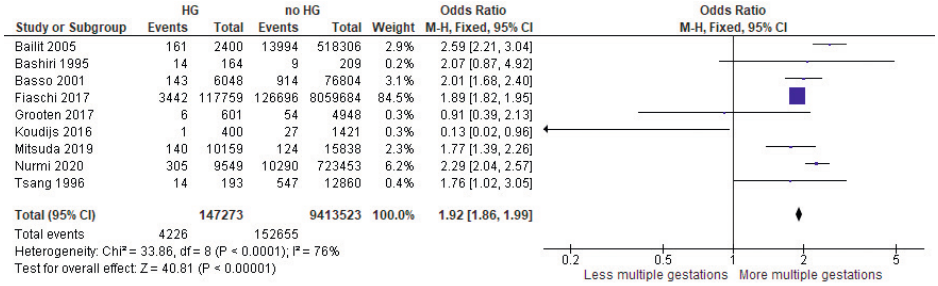
Two studies reported on newborn measurements (n=328, **Appendix S**).[32, 51] Both report the crown-heel length and the head circumference in centimetres. Only the head circumference in neonates born at term was smaller in the HG exposed infants in one of the studies (HG-exposed: 34.8 cm IQR 30.2-38, non-exposed 35.5cm IQR 31-39, p=0.02).[32] Meta-analysis was not possible due to different ways of outcome reporting. One of the studies (n=198) also reported the frequency of the head circumference being respectively small or large for gestational age and found no differences between the HG-exposed and non-exposed group.[32] One study (n=130) reported on the biparietal diameter, but did not state if there was a significant difference between HG-exposed (9.3±0.1 cm) and non-exposed newborns (9.2±0.1 cm).[51]

Placental abruption

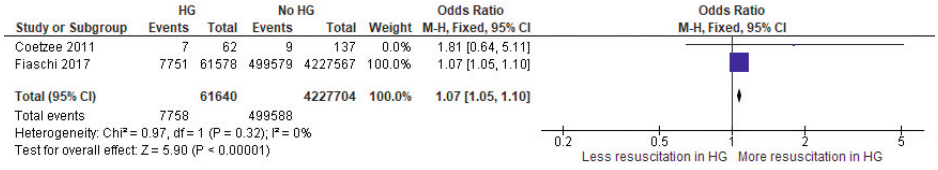
Six studies reported on placental abruption (n=9,370,211, **Appendix T**).[12, 19, 27, 30, 42, 68] Meta-analysis, **Figure 9.4,** showed a higher risk of placental abruption in the HG-exposed group (OR 1.15, 95%CI:1.05-1.25, p=0.002, I%14=[2]).

Figure 9. Forest plots multiple gestations, resuscitation, NICU admission and placental abruption

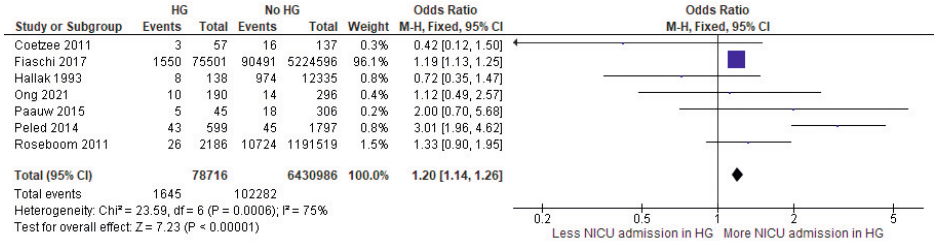
9.1 Forest plot multiple gestation



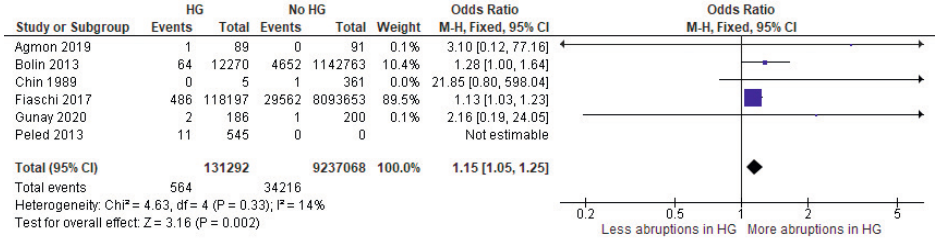
9.2 Forest plot neonatal resuscitation



9.3 Forest plot NICU admission



9.4 Forest plot placental abruption



Subgroup analysis of studies performed after 2001

There were 28 studies conducted after 2001 (six case-control and 22 cohort studies) and 33 for 2001 (seven case-control and 26 cohort studies). [12, 19, 29, 32, 39-41, 44, 45, 47, 54, 56, 58-65, 70-72, 76, 77, 83, 86, 87] Subgroup analysis could be performed for seventeen outcomes (**Table 3.1 and Appendix U**).

Subgroup analysis revealed a lower birth weight (8 studies, n=1,199,384, mean difference -99.71 gram, 95%CI: -121.82--77.59, p<0.00001, I^2=64%, **Appendix U.4**), more SGA infants (11 studies, n=1,207,784, OR1.13, 95%CI:1.02-1.26, p=0.02, I%31=^{2}, **Appendix U.8)** and more female infants in the HG-exposed group (18 studies, n=1,980,302, OR1.31, 95%CI:1.28-1.35, p<0.00001, I%61=^{2}, **Appendix U.14).**

Subgroup analysis showed high heterogeneity in case of congenital anomalies (I%80=^{2}, **Appendix U.11**), multiple gestations (I%81=^{2}, **Appendix U.24**), preterm birth <37 weeks (I%77=^{2}, **Appendix U.30**) and preterm birth <32 weeks (I%96=^{2}, **Appendix U.32**).

Subgroup analysis revealed no differences in Apgar score <7 at 1 minute, Apgar score <7 at 5 minutes, birth weight <2500 grams, LGA infants, gestational age at delivery, fetal loss of fetal death, stillbirth, perinatal death, NICU admission and placental abruption (**Appendix U**).

Subgroup analysis including only studies performed before 2001 for the same outcomes, revealed high heterogeneity in nine outcomes, no effect in two outcomes and was not possible to perform in two outcomes (**Table 3.1**). Four outcomes showed significant correlations with more Apgar score <7 at 5 minutes, more multiple gestations, more placental abruptions and less stillbirths in HG exposed infants.

Subgroup analysis women hospitalized for HG compared with non-exposed

To assess the possibility of HG severity affecting perinatal outcomes, we performed a subgroup analysis only including hospitalized women with HG and compared this subgroup with women without HG (the non-exposed group). There were 32 studies, 26 cohort and six case-control studies, that only included women hospitalized for HG or which had a subgroup of women hospitalized for HG.[12, 21, 23-27, 30-32, 34, 36, 38, 40-43, 46, 47, 49, 51, 56, 62, 64, 66-68, 70, 73, 74, 76, 77, 81-83] Subgroup analysis could be performed for sixteen outcomes **(Table 3.2, Appendix V**).

In offspring born to women hospitalized for HG there was less often an Apgar score <7 at 1 minute (two studies, n=71,670, OR 0.67, 95%CI:0.47-0.97, p=0.03, I%23=^{2}, **Appendix V.1**), birth

weight <2500 grams was more common (nine studies, n=6,221,275, OR 1.33, 95%CI:1.30-1.37, p<0.00001, I%13=^{2}, **Appendix V.4**), there were more female infants (23 studies, n=9,658,736, OR 1.20, 95%CI:1.18-1.21, p<0.00001, I%53=^{2}, **Appendix V.8**), there was less often stillbirth (six studies, n=9,364,195, OR 0.90, 95%CI:0.83-0.98, p=0.01, I%0=^{2}, **Appendix V.10**), less often perinatal death (five studies, n=229,083, OR 0.44, 95%CI:0.26-0.77, p=0.004, I%55=^{2}, **Appendix V.11**), more often neonatal death (three studies, n=507,591, OR 1.69, 95%CI:1.02-2.80, p=0.04, I%0=^{2}, **Appendix V.12**), more often placental abruption (five studies, n=9,368,180, OR1.14, 95%CI:1.05-1.25, p=0.002, I%30=^{2}, **Appendix V.15**) and more often preterm birth <37 weeks in the HG-exposed group (twenty studies, 5,485,642, OR 1.24 95%CI:1.21-1.27, p<0.00001, I^2=71%, **Appendix V.16**)

Subgroup analysis for hospitalization revealed no effect in case of Apgar score <7 at 5 minutes (six studies, n=242,975, OR 1.21, 95%CI:0.88-1.66, p=0.25, I%0=^{2}, **Appendix V.2**) and LGA (5 studies, n=5,983,177, OR 1.02, 95%CI:1.00-1.04, p=0.06, I%64=^{2}, **Appendix V.6**).

Subgroup analysis for hospitalization showed high heterogeneity in case of birth weight (I%95=^{2}, **Appendix V.3**), SGA(I%95=^{2}, **Appendix V.5**), congenital anomalies(I%81=^{2}, **Appendix V.7**), gestational age at delivery(I%98=^{2}, **Appendix V.9**), multiple gestations(I%80=^{2}, **Appendix V.13**) and NICU admission(I%83=^{2}, **Appendix V.14**).

Subgroup analysis mild versus severe HG

Six studies, five cohort studies and one case-control study, made a distinction between mild and severe HG.[19, 31, 46, 47, 54, 73] All studies used different definitions, see **Table 4**. Subgroup analysis could be performed for birth weight, birth weight <2500 grams, SGA, fetal sex, gestational age at delivery and preterm birth <37 weeks (**Appendix W**).

When comparing mild HG with no HG, heterogeneity remained high in case of fetal sex, none of the other outcomes showed an effect (**Table 3.3**).

When comparing severe HG with no HG, birth weight was significantly lower (mean difference -128.43 gram, 95%CI:-206.29- -0.81, p=0.05, I%0=^{2}) and significantly more female infants (OR 1.44, 95%CI:1.29-1.60, p<0.00001, I%0=^{2}) were born in the severe HG group (**Table 3.4**). Birth weight <2500 grams, SGA, gestational age at delivery and preterm birth <37 weeks did not show significant differences.

When comparing severe HG with mild HG there was a significantly lower birth weight in the severe HG group (mean difference -104.29 gram, 95%CI:-207.78- -0.81, p=0.001, I%67=^{2}), all other outcomes did not show an effect (**Table 3.5**).

For eleven outcomes subgroup analysis was not possible due to the fact that only one study made a distinction between mild and severe HG. One study (n=89) showed no significant difference when comparing mild with severe HG on Apgar score at 1 minute (mild n=4/46, severe n=5/43, p=0.417, **Appendix D**), Apgar score at 5 minutes (mild n=0/46, severe n=2/43, p=0.198, **Appendix D**), pH at birth (mild: 7.30±0.95, severe: 7.29±0.92, p=0.23, **Appendix D**) perinatal deaths (mild n=0/46, severe n=2/43, p=0.192, **Appendix N)** and placental abruption (mild n=0/46, severe n=1/43, p=0.202, **Appendix T**).[19]

Another study (n=2,233) made a distinction between mild and severe HG but only compared this to the non HG-exposed group. They found similar Apgar scores at 5 minutes between groups (all three groups 9±0, **Appendix D**), no difference in fetal loss in the severe or mild HG group compared to the non-exposed group (non-exposed:7.5%, mild:7.9% p=0.75, severe:4.3% p=0.44, **Appendix N**), no difference in stillbirth in the severe or mild HG group compare to the non-exposed group (non-exposed:1.2%, mild:1% p=0.74, severe:2.7% p=0.26, **Appendix N**).[54]

A third study (n=12,473) made a distinction between mild and severe HG and found no significant relations on congenital anomalies (non-exposed:1.6%, mild:2.5%, severe:2% p>0.05, **Appendix G**) and NICU admission(non-exposed:7.9%, mild:10%, severe:4.1% p>0.05, **Appendix Q**).[46]

Table 3. Overview of results meta-analysis compared with subgroup analysis

3.1 Subgroup analysis studies performed after 2001

Outcome	Meta-analysis	Subgroup analysis after 2001	Subgroup analysis before 2001
Apgar <7 at 1 minute	No effect: OR 0.74 95%CI:0.53-1.03, p=0.07, I^2=30%	No effect: OR 1.31, 95%CI:0.68-2.51, p=0.42, I^2=0%	Not possible, only one study available
Apgar <7 at 5 minutes	No effect: OR 1.12, 95%CI:1.00-1.26, p=0.06, I^2=0%	No effect: OR 0.95, 95%CI:0.73-1.24, p=0.70, I^2=0%	More apgar scores <7 at 5 minutes: OR 1.17, 95%CI:1.03-1.33, p=0.02, I^2=0%
Birth weight	High heterogeneity: I^2=92%	Lower birth weight: mean difference -99.71 gram, 95%CI:-121.82--77.59, p<0.00001, I^2=64%	High heterogeneity: I^2=94%
Birth weight <2500 grams	High heterogeneity: I^2=77%	No effect: OR 0.95, 95%CI:0.73-1.24, p=0.70, I^2=0%	High heterogeneity: I^2=85%
Small for gestational age	High heterogeneity: I^2=93%	More SGA: OR 1.13, 95%CI:1.02-1.26, p=0.02, I^2=31%	High heterogeneity: I^2=97%
Large for gestational age	High heterogeneity: I^2=100%	No effect: OR 1.16, 95%CI:0.76-1.78, p=0.49, I^2=52%	High heterogeneity: I^2=100%
Congenital anomalies	High heterogeneity: I^2=95%	High heterogeneity: I^2=80%	High heterogeneity: I^2=88%
Fetal sex- female	High heterogeneity: I^2=78%	More often: OR 1.31, 95%CI:1.28-1.35, p<0.00001, I^2=61%	High heterogeneity: I^2=76%
Gestational age at delivery	High heterogeneity: I^2=97%	No effect: mean difference 0.06, 95%CI:-0.20-0.33, p=0.63, I^2=0%	High heterogeneity: I^2=99%
Fetal loss or fetal death	No effect: OR 0.97, 95%CI:0.75-1.27, p=0.84, I^2=64%	No effect: OR 0.82, 95%CI:0.57-1.19, p=0.30, I^2=64%	No effect: OR 1.19, 95%CI:0.81-1.73, p=0.38, I^2=64%
Stillbirth	Less often: OR 0.92, 95%CI:0.85-0.99, p=0.02, I^2=0%	No effect: OR 1.01, 95%CI:0.38-2.66, p=0.99, I^2=0%	Less often: OR 0.91, 95%CI:0.85-0.99, p=0.02, I^2=35%
Perinatal death	No effect: OR 0.97, 95%CI:0.85-1.11, p=0.64, I^2=30%	No effect: OR 0.68, 95%CI:0.38-1.19, p=0.17, I^2=0%	No effect: OR 0.99, 95%CI:0.87-1.14, p=0.92, I^2=44%
Multiple gestations	High heterogeneity: I^2=76%	High heterogeneity: I^2=81%	More multiple gestations: OR 1.91, 95%CI:1.85-1.98, p<0.00001, I^2=73%
NICU admission	More often: OR 1.20 95%CI:1.14-1.26, p<0.00001, I^2=75%	No effect: OR 1.15, 95%CI:0.82-1.62, p=0.42, I^2=32%	High heterogeneity: I^2=86%
Placental abruption	More often: OR 1.15, 95%CI:1.05-1.25, p=0.002, I^2=14%	No effect: OR 2.48, 95%CI:0.36-16.93, p=0.35, I^2=0%	More often: OR 1.14 95%CI:1.05-1.24, p=0.002, I^2=50%
Prematurity <37 weeks	High heterogeneity: I^2=79%	High heterogeneity: I^2=77%	High heterogeneity: I^2=78%
Prematurity <32 weeks	High heterogeneity: I^2=96%	High heterogeneity: I^2=96%	Not possible, only one study available

3.2 Subgroup analysis hospitalization compared with non-exposed

Outcome	Meta-analysis	Subgroup analysis
Apgar <7 at 1 minute	No effect: OR 0.74 95%CI:0.53-1.03, p=0.07, I^2=30%	Less often: OR 0.67, 95%CI:0.47-0.97, p=0.03, I^2=23%
Apgar <7 at 5 minutes	No effect: OR 1.12, 95%CI:1.00-1.26, p=0.06, I^2=0%	No effect: OR 1.21, 95%CI:0.88-1.66, p=0.25, I^2=0%
Birth weight	High heterogeneity: I^2=92%	High heterogeneity: I^2=95%
Birth weight <2500 grams	High heterogeneity: I^2=77%	More often: OR 1.33, 95%CI:1.30-1.37, p<0.00001, I^2=13%
Small for gestational age	High heterogeneity: I^2=93%	High heterogeneity: I^2=95%
Large for gestational age	High heterogeneity: I^2=100%	No effect: OR 1.02, 95%CI:1.00-1.04, p=0.06, I^2=64%
Congenital anomalies	High heterogeneity: I^2=95%	High heterogeneity: I^2=81%
Fetal sex- female	High heterogeneity: I^2=78%	More often: OR 1.20, 95%CI:1.18-1.21, p<0.00001, I^2=53%
Gestational age at delivery	High heterogeneity: I^2=97%	High heterogeneity: I^2=98%
Stillbirth	Less often: OR 0.92, 95%CI:0.85-0.99, p=0.02, I^2=0%	Less often: OR 0.90, 95%CI:0.83-0.98, p=0.01, I^2=0%
Perinatal death	No effect: OR 0.97, 95%CI:0.85-1.11, p=0.64, I^2=30%	Less often: OR 0.44, 95%CI:0.26-0.77, p=0.004, I^2=55%
Neonatal death	No effect: OR 1.11, 95%CI:0.90-1.35, p=0.33, I^2=21%	More often: OR 1.69, 95%CI:1.02-2.80, p=0.04, I^2=0%
Multiple gestations	High heterogeneity: I^2=76%	High heterogeneity: I^2=80%
NICU admission	More often: OR 1.20 95%CI:1.14-1.26, p<0.00001, I^2=75%	High heterogeneity: I^2=83%
Placental abruption	More often: OR 1.15, 95%CI:1.05-1.25, p=0.002, I^2=14%	More often: OR 1.14, 95%CI:1.05-1.25, p=0.002, I^2=30%
Prematurity <37 weeks	High heterogeneity: I^2=79%	More often: OR 1.24 95%CI:1.21-1.27, p<0.00001, I^2=71%

3.3 Subgroup analysis severity HG: mild HG compared with no HG

Outcome	Meta-analysis	Subgroup analysis
Birth weight	High heterogeneity: I^2=92%	No effect: mean difference 0.42, 95%CI:-46.80-47.64, p=0.99, I^2=0%
Birth weight <2500 grams	High heterogeneity: I^2=77%	No effect: OR 0.96, 95%CI:0.65-1.41, p=0.84, I^2=0%
Small for gestational age	High heterogeneity: I^2=93%	No effect: OR 1.19, 95%CI:0.91-1.56, p=0.21, I^2=0%
Fetal sex- female	High heterogeneity: I^2=78%	High heterogeneity: I^2=85%
Gestational age at delivery	High heterogeneity: I^2=97%	No effect: mean difference 0.13, 95%CI:-0.07-0.33, p=0.20, I^2=0%
Prematurity <37 weeks	High heterogeneity: I^2=79%	No effect: OR 1.37, 95%CI:0.69-2.74, p=0.37, I^2=0%

3.4 Subgroup analysis severity HG: severe HG compared with no HG

Outcome	Meta-analysis	Subgroup analysis
Birth weight	High heterogeneity: I^2=92%	Lower birth weight: mean difference -128.43, 95%CI:-206.29- -0.81, p=0.05, I^2=0%
Birth weight <2500 grams	High heterogeneity: I^2=77%	No effect: OR 1.61, 95%CI:0.80-3.25, p=0.18, I^2=38%
Small for gestational age	High heterogeneity: I^2=93%	No effect: OR 1.53, 95%CI:0.94-2.49, p=0.08, I^2=0%
Fetal sex- female	High heterogeneity: I^2=78%	More often: OR 1.44, 95%CI:1.29-1.60, p<0.00001, I^2=0%
Gestational age at delivery	High heterogeneity: I^2=97%	No effect: mean difference -0.07, 95%CI:-0.40-0.26, p=0.68, I^2=16%
Prematurity <37 weeks	High heterogeneity: I^2=79%	No effect: OR 1.03, 95%CI:0.59-1.81, p=0.92, I^2=0%

3.5 Subgroup analysis severity HG: mild HG compared with severe HG

Outcome	Subgroup analysis
Birth weight	Lower birth weight in severe HG: mean difference -104.29, 95%CI:-207.78- -0.81, p=0.001, I^2=67%
Birth weight <2500 grams	No effect: OR 1.61, 95%CI:0.80-3.25, p=0.18, I^2=29%
Small for gestational age	No effect: OR 1.31, 95%CI:0.77-2.23, p=0.32, I^2=16%
Fetal sex- female	No effect: OR 0.98, 95%CI:0.82-1.17, p=0.84, I^2=60%
Gestational age at delivery	No effect: mean difference -0.16, 95%CI:-0.62-0.31, p=0.51, I^2=0%
Prematurity <37 weeks	No effect: OR 0.73, 95%CI:0.32-1.70, p=0.47, I^2=0%

Rows were gray-colored in case MA and subgroup analysis showed similar effect

Table 4. Definition severe HG

Study	Definition severe HG
Agmon	≥2 of the following: (1) ≥3 hospitalizations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria
Chin 1988	Heavy ketonuria (>3+), increase in urea and creatinine concentrations, serum electrolyte disturbance and/or increase in hematocrit (> 0.43)
Hallak	Ketonuria, increased BUN and hematocrit, and/or abnormal electrolytes.
Hastoy	Weight gain less than 7kg
Koudijs	>5% weight loss compared to pre pregnancy weight
Schiff	Hospitalized ≥3 times or admitted with a metabolic disturbance or hospitalized for ≥3 days

DISCUSSION

Main findings

Our systematic review demonstrated that HG is associated with several adverse perinatal outcomes including low birth weight and preterm birth. We found some evidence that suggests severe HG has a larger impact on adverse outcomes than milder HG. Possible improvements in care in the last two decades (after 2001) did not translate in any appreciable improvements in perinatal outcomes among women with HG in studies conducted in the more recent time frame, compared to older studies.

Strengths and limitations

This systematic review has several strengths. The research protocol was prospectively published online. We were able to include a large number of studies on multiple perinatal outcomes. More than 80% of included studies were of fair to good quality. We were able to conduct 23 meta-analyses, in case of high heterogeneity we performed sensitivity analyses and we explored potential effects on outcomes of mild and severe HG. Considering these strengths we feel our systematic review gives a complete overview on perinatal outcomes of infants born to mothers with hyperemesis gravidarum.

There were several limitations of this review. Eleven out of 23 meta-analyses revealed high heterogeneity, which limited our ability to provide aggregate point estimates. This heterogeneity may have been caused by variation in HG definitions (i.e. based on ICD codes, medical records and self-reports) or in reporting of perinatal outcomes, a problem that has been previously identified in HG research.[89] Finally, we were not able to assess the role of maternal weight gain during pregnancy, or of treatment modalities, nor were we able to investigate the role of protracted HG on outcomes, each of which would could have produced relevant input for antenatal management of women with HG.

Comparison with existing literature

Low gestational weight gain and maternal undernutrition can lead to lower birth weight and preterm birth.[7] Meta-analysis showed that HG was associated with an increased risk of birth weight <1500 gram and a lower risk of birth weight >4000 gram, although these effects were modest in size (respectively OR 1.43 and OR 0.74). The fact that HG can lead to low pregnancy weight gain and maternal undernutrition can be a major contributor to the effects on birth weight and preterm birth.[3, 4] Altered vascular development, diminished angiogenic growth factor expression, and reduced placental glucose, amino acid, and lipid transport are

all associated with maternal undernutrition and lower birth weight and preterm birth[90-96]. Unfortunately, the included studies did not report enough data on maternal dietary factors, or maternal weight or weight change in pregnancy and we were therefore unable to investigate whether maternal undernutrition could be the causative mechanism underlying HG's effects on perinatal outcomes.

Our meta-analysis confirmed that HG is associated with a reduced risk of stillbirth, which may seem at odds with increased low birth weight. Younger maternal age [97, 98], increased monitoring and induction of labour could each contribute to the relatively lower risk of still birth in pregnancies complicated by HG.[99] Others have suggested that nausea and vomiting symptoms might be the direct result of sound placental function, and could underlie HG's protective effects for stillbirth.[100-102] Due to limited data we were unable to investigate these theories in this systematic review.

We aimed to investigate the effect of HG severity on perinatal outcomes. Some studies based severity on maternal weight difference, others on the number of hospitalizations or on biochemical determinants. We found significantly lower birth weight after severe HG in comparison to mild HG (mean difference -104.29, 95%CI:-207.78- -0.81, p=0.001, I%67=^{2}), for five other outcomes (birth weight <2500 grams, SGA, gestational age at delivery and preterm birth <37 weeks) subgroup analysis showed no alterations in outcomes after severe HG. Previous work of ours suggested that HG's detrimental effects on perinatal outcomes were largely explained by maternal characteristics.[103] Other explanations for the lack of increase in adverse perinatal outcomes after severe HG can be due to the impossibility of identifying women with a poor prognosis. We defined severe HG as hospitalisation for HG, but many other factors than actual disease severity could have resulted in hospitalisation: social and financial determinants can affect a decision to admit. Therefore, a more appropriate definition of severe HG, for example based on factors indicating a poor prognosis, would be helpful to resolve whether there is a 'dose-effect' relation between HG and perinatal outcomes. Unfortunately, such factors are yet unknown and therefore not included in the recently published international definition of hyperemesis gravidarum.[104]

Conclusions and implications

This systematic review and meta-analysis showed that HG is associated with a range of adverse perinatal outcomes, and that there may be larger effects among those more severely affected by HG. Future studies should investigate the suspected mediating role of maternal undernutrition in adverse outcomes among mothers with HG, and also assess which role

prompt treatment with anti-emetics may play in improving adverse perinatal outcomes. Our findings could be an argument for clinicians to offer increased surveillance to women affected with HG.

Acknowledgements

None

Funding

This systematic review did not receive any funding.

Contribution to authorship

KN, LAWJ, RCP. and TJR. conceived and designed the study. JL performed the electronic search. KN and LAWJ screened titles and abstracts for eligibility and performed data extraction. LAWJ performed all statistical analysis and drafted the manuscript. All authors contributed in interpreting the results and revising the manuscript and approved the final draft of this manuscript.

REFERENCES

1. Einarson TR, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: a meta analysis. J Popul Ther Clin Pharmacol. 2013;20(2):e171-83.
2. Jansen LAW, Koot MH, van't Hooft J, Dean CR, Bossuyt PMM, Ganzevoort W, et al. The Windsor Definition for Hyperemesis Gravidarum: a multistakeholder International Consensus Definition. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2021.
3. Fairweather DVI. Nause and vomiting in pregnancy. American Journal of Obstetrics and Gynecology. 1968;102(1):135-75.
4. Niebyl JR. Nausea and Vomiting in Pregnancy. N Engl J Med. 2010;363(16):1544-50.
5. Boelig RC, Barton SJ, Saccone G, Kelly AJ, Edwards SJ, Berghella V. Interventions for treating hyperemesis gravidarum. Cochrane Database Syst Rev. 2016(5):CD010607.
6. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. N Engl J Med. 2008;359(1):61-73.
7. Han Z, Lutsiv O, Mulla S, Rosen A, Beyene J, McDonald SD, et al. Low gestational weight gain and the risk of preterm birth and low birthweight: a systematic review and meta-analyses. Acta Obstet Gynecol Scand. 2011;90(9):935-54.
8. Harding JE. The nutritional basis of the fetal origins of adult disease. Int J Epidemiol. 2001;30(1):15-23.
9. McDonald SD, Han Z, Mulla S, Beyene J, Knowledge Synthesis G. Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analyses. Bmj. 2010;341:c3428.
10. Nijland MJ, Ford SP, Nathanielsz PW. Prenatal origins of adult disease. Curr Opin Obstet Gynecol. 2008;20(2):132-8.
11. Veenendaal MV, van Abeelen AF, Painter RC, van der Post JA, Roseboom TJ. Consequences of hyperemesis gravidarum for offspring: a systematic review and meta-analysis. BJOG. 2011;118(11):1302-13.
12. Gunay T, Turgut A, Ayaz Bilir R, Hocaoglu M, Demircivi Bor E. Comparative Analysis of Maternal and Fetal Outcomes of Pregnancies Complicated and Not Complicated with Hyperemesis Gravidarum Necessitating Hospitalization. Medeniyet med. 2020;35(1):8-14.
13. Koudijs HM, Savitri AI, Browne JL, Amelia D, Baharuddin M, Grobbee DE, et al. Hyperemesis gravidarum and placental dysfunction disorders. BMC Pregnancy and Childbirth. 2016;16(1):374.
14. Özay AC, Özay ÖE. The Effect of Hyperemesis Gravidarum on Gestational Diabetes and Pregnancy Outcomes. Zeynep Kamil Med J. 2019;50(1):50-3.
15. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient–clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ Open. 2021;11(1):e041254.
16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
17. Wells G SB, O'Connell D, Peterson J, Welch V, Losos M, et al. . The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. 2014.
18. Higgins JPT TJ CJ, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). 2020. Available from: www.training.cochrane.org/handbook.
19. Agmon N, Sade S, Pariente G, Rotem R, Weintraub AY. Hyperemesis gravidarum and adverse pregnancy outcomes. Arch Gynecol Obstet. 2019;300(2):347-53.

20. Almond D, Edlund L, Joffe M, Palme M. An adaptive significance of morning sickness? Trivers-Willard and Hyperemesis Gravidarum. Econ Hum Biol. 2016;21:167-71.

21. Askling J, Erlandsson G, Kaijser M, Akre O, Ekbom A. Sickness in pregnancy and sex of child. Lancet. 1999;354(9195):2053.

22. Axt-Fliedner R. Hyperemesis gravidarum: Sex ratio in pregnancies. [German]. Gynakologe. 2004;37(8):763-4.

23. Bailit JL. Hyperemesis gravidarium: Epidemiologic findings from a large cohort. Am J Obstet Gynecol. 2005;193(3 Pt 1):811-4.

24. Bashiri A, Neumann L, Maymon E, Katz M. Hyperemesis gravidarum: epidemiologic features, complications and outcome. Eur J Obstet Gynecol Reprod Biol. 1995;63(2):135-8.

25. Basso O, Olsen J. Sex ratio and twinning in women with hyperemesis or pre-eclampsia. Epidemiology. 2001;12(6):747-9.

26. Boissiere-O'Neill T, Schnitzer ME, Lewin A, Bilodeau-Bertrand M, Ayoub A, Auger N. Original article: is the protective association between hyperemesis gravidarum and birth defects biased by pregnancy termination? Ann Epidemiol. 2021;59:10-5.

27. Bolin M, Akerud H, Cnattingius S, Stephansson O, Wikstrom AK. Hyperemesis gravidarum and risks of placental dysfunction disorders: a population-based cohort study. Bjog. 2013;120(5):541-7.

28. Buyukkayaci Duman N, Ozcan O, Bostanci MO. Hyperemesis gravidarum affects maternal sanity, thyroid hormones and fetal health: a prospective case control study. Arch Gynecol Obstet. 2015;292(2):307-12.

29. Caltekin MD, Caltekin I, Onat T, Kirmizi DA, Baser E, Yalcin SE, et al. Can Subclinical Inflammatory Markers Predict Birth Time and Birth Weight in Hyperemesis Gravidarum?: A Comparative Study and Comprehensive Current Literature Review. Med Bull Haseki. 2021;59:139-44.

30. Chin RK. Antenatal complications and perinatal outcome in patients with nausea and vomiting-complicated pregnancy. Eur J Obstet Gynecol Reprod Biol. 1989;33(3):215-9.

31. Chin RK, Lao TT. Low birth weight and hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 1988;28(3):179-83.

32. Coetzee RL, Cormack B, Sadler L, Bloomfield FH. Pregnancy and neonatal outcomes following hyperemesis gravidarum. J Dev Orig Health Dis. 2011;2(2):81-8.

33. Coffey VP, Jessop WJE. A study of 137 cases of anencephaly. British journal of preventive and social medicine. 1957;11(4):174-80.

34. Czeizel AE, Puho E, Acs N, Banhidy F. Inverse association between severe nausea and vomiting in pregnancy and some congenital abnormalities. Am J Med Genet A. 2006;140(5):453-62.

35. Czeizel AE, Sarkozi A, Wyszynski DF. Protective effect of hyperemesis gravidarum for nonsyndromic oral clefts. Obstet Gynecol. 2003;101(4):737-44.

36. del Mar Melero-Montes M, Jick H. Hyperemesis gravidarum and the sex of the offspring. Epidemiology. 2001;12(1):123-4.

37. Depue RH, Bernstein L, Ross RK, Judd HL, Henderson BE. Hyperemesis gravidarum in relation to estradiol levels, pregnancy outcome, and other maternal factors: a seroepidemiologic study. Am J Obstet Gynecol. 1987;156(5):1137-41.

38. Dodds L, Fell DB, Joseph KS, Allen VM, Butler B. Outcomes of pregnancies complicated by hyperemesis gravidarum. Obstet Gynecol. 2006;107(2 Pt 1):285-92.

39. Fejzo M, Kam A, Laguna A, MacGibbon K, Mullin P. Analysis of neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum reveals increased reporting of autism spectrum disorder. Reproductive Toxicology. 2019;84:59-64.

40. Fejzo MS, Magtira A, Schoenberg FP, MacGibbon K, Mullin P, Romero R, et al. Antihistamines and other prognostic factors for adverse outcome in hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2013;170(1):71-6.

41. Fejzo MS, Magtira A, Schoenberg FP, Macgibbon K, Mullin PM. Neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2015;189:79-84.

42. Fiaschi L, Nelson-Piercy C, Gibson J, Szatkowski L, Tata LJ. Adverse Maternal and Birth Outcomes in Women Admitted to Hospital for Hyperemesis Gravidarum: a Population-Based Cohort Study. Paediatric and Perinatal Epidemiology. 2018;32(1):40-51.

43. Getahun D, Fassett MJ, Jacobsen SJ, Xiang AH, Takhar HS, Wing DA, et al. Autism Spectrum Disorders in Children Exposed in Utero to Hyperemesis Gravidarum. American Journal of Perinatology. 2019;03:03.

44. Grooten IJ, Den Hollander WJ, Roseboom TJ, Kuipers EJ, Jaddoe VW, Gaillard R, et al. Helicobacter pylori infection: a predictor of vomiting severity in pregnancy and adverse birth outcome. Am J Obstet Gynecol. 2017;216(5):512.e1-.e9.

45. Gu L, Mo M, Si S, Luo W, Shao B, Xin X, et al. Association of nausea and vomiting of pregnancy with infant growth in the first 24 months of life. Arch Gynecol Obstet. 2021;304(2):429-38.

46. Hallak M, Tsalamandris K, Dombrowski MP, Isada NB, Pryde PG, Evans MI. Hyperemesis gravidarum. Effects on fetal outcome. Journal of Reproductive Medicine. 1996;41(11):871-4.

47. Hastoy A, Lien Tran P, Lakestani O, Barau G, Gerardin P, Boukerrou M. [Hyperemesis gravidarum and pregnancy outcomes]. J Gynecol Obstet Biol Reprod (Paris). 2015;44(2):154-63.

48. Hohlbein R. [Hyperemesis gravidarum as the cause of infantile malformations]. Med Klin. 1961;56:93-5.

49. Hsu CD, Witter FR. Fetal sex and severe hyperemesis gravidarum. Int J Gynaecol Obstet. 1993;40(1):63-4.

50. Kallen B. Hyperemesis during pregnancy and delivery outcome: a registry study. Eur J Obstet Gynecol Reprod Biol. 1987;26(4):291-302.

51. Kidess E, Klein M. [The outcome of pregnancies complicated by hyperemesis gravidarum (author's transl)]. Geburtshilfe und Frauenheilkunde. 1974;34(3):181-5.

52. Klebanoff MA, Koslowe PA, Kaslow R, Rhoads GG. Epidemiology of vomiting in early pregnancy. Obstet Gynecol. 1985;66(5):612-6.

53. Koot MH, Grooten IJ, Sebert S, Koiranen M, Jarvelin MR, Kajantie E, et al. Hyperemesis gravidarum and cardiometabolic risk factors in adolescents: a follow-up of the Northern Finland Birth Cohort 1986. Bjog. 2017;124(7):1107-14.

54. Koudijs HM, Savitri AI, Browne JL, Amelia D, Baharuddin M, Grobbee DE, et al. Hyperemesis gravidarum and placental dysfunction disorders. BMC Pregnancy Childbirth. 2016;16(1):374.

55. Kruse HJ, Adomssent S, Herre HD. Body measurements and morphologic criteria of maturity in newborn infants of mothers with hyperemesis gravidarum. [German]. Zentralblatt fur Gynakologie. 1975;97(18):1105-9.

56. Kuru O, Sen S, Akbayir O, Goksedef BP, Ozsurmeli M, Attar E, et al. Outcomes of pregnancies complicated by hyperemesis gravidarum. Arch Gynecol Obstet. 2012;285(6):1517-21.

57. Lu QB, Wang ZP, Gao LJ, Gong R, Sun XH, Wang M, et al. Nausea and vomiting in early pregnancy and the risk of neural tube defects: a case-control study. Sci. 2015;5:7674.

58. McCarthy FP, Khashan AS, North RA, Moss-Morris R, Baker PN, Dekker G, et al. A prospective cohort study investigating associations between hyperemesis gravidarum and cognitive, behavioural and emotional well-being in pregnancy. PLoS ONE. 2011;6(11):e27678.

59. Mitsuda N, Eitoku M, Maeda N, Fujieda M, Suganuma N. Severity of Nausea and Vomiting in Singleton and Twin Pregnancies in Relation to Fetal Sex: The Japan Environment and Children's Study (JECS). J Epidemiol. 2019;29(9):340-6.

60. Mitsuda N, Eitoku M, Yamasaki K, Sakaguchi M, Yasumitsu-Lovell K, Maeda N, et al. Nausea and vomiting during pregnancy associated with lower incidence of preterm births: the Japan Environment and Children's Study (JECS). BMC Pregnancy Childbirth. 2018;18(1):268.

61. Morokuma S, Shimokawa M, Kato K, Sanefuji M, Shibata E, Tsuji M, et al. Relationship between hyperemesis gravidarum and small-for-gestational-age in the Japanese population: the Japan Environment and Children's Study (JECS). BMC Pregnancy Childbirth. 2016;16:247.

62. Muraoka M, Takagi K, Ueno M, Morita Y, Nagano H. Fetal Head Growth during Early to Mid-Gestation Associated with Weight Gain in Mothers with Hyperemesis Gravidarum: A Retrospective Cohort Study. Nutrients. 2020;12(6):03.

63. Nurmi M, Rautava P, Gissler M, Vahlberg T, Polo-Kantola P. Incidence and risk factors of hyperemesis gravidarum: A national register-based study in Finland, 2005-2017. Acta Obstetricia et Gynecologica Scandinavica. 2020;99(8):1003-13.

64. Ong J, Sadananthan SA, Soh SE, Ng S, Yuan WL, Aris IM, et al. Increasing nausea and vomiting of pregnancy is associated with sex-dependent differences in early childhood growth: the GUSTO mother-offspring cohort study. BMC Pregnancy Childbirth. 2021;21(1):578.

65. Ozay AC, Ozay OE. The effect of hyperemesis gravidarum on gestational diabetes and pregnancy outcomes. [Turkish]. Zeynep Kamil Tip Bulteni. 2019;50(1):50-3.

66. Paauw JD, Bierling S, Cook CR, Davis AT. Hyperemesis gravidarum and fetal outcome. JPEN J Parenter Enteral Nutr. 2005;29(2):93-6.

67. Peled Y, Melamed N, Hiersch L, Hadar E, Wiznitzer A, Yogev Y. Pregnancy outcome in hyperemesis gravidarum--the role of fetal gender. J Matern Fetal Neonatal Med. 2013;26(17):1753-7.

68. Peled Y, Melamed N, Hiersch L, Pardo J, Wiznitzer A, Yogev Y. The impact of total parenteral nutrition support on pregnancy outcome in women with hyperemesis gravidarum. J Matern Fetal Neonatal Med. 2014;27(11):1146-50.

69. Poeran-Bahadoer S, Jaddoe VWV, Gishti O, Grooten IJ, Franco OH, Hofman A, et al. Maternal vomiting during early pregnancy and cardiovascular risk factors at school age: the Generation R Study. J Dev Orig Health Dis. 2020;11(2):118-26.

70. Rashid M, Rashid MH, Malik F, Herath RP. Hyperemesis gravidarum and fetal gender: a retrospective study. J Obstet Gynaecol. 2012;32(5):475-8.

71. Roseboom TJ, Ravelli AC, van der Post JA, Painter RC. Maternal characteristics largely explain poor pregnancy outcome after hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2011;156(1):56-9.

72. Salunkhe AH, Pratinidhi AK, Salunkhe JA, Kakade SV, Mohite VR, Patange RP. Antenatal Risk Scoring Scale for Predication of Low Birth Weight and Its Validity. Indian J Community Med. 2019;44(2):97-101.

73. Schiff MA, Reed SD, Daling JR. The sex ratio of pregnancies complicated by hospitalisation for hyperemesis gravidarum. Bjog. 2004;111(1):27-30.

74. Sorensen HT, Thulstrup AM, Mortensen JT, Larsen H, Pedersen L. Hyperemesis gravidarum and sex of child. Lancet. 2000;355(9201):407.

75. Syn NL, Chan SY, Chia EWY, Ong WX, Phua D, Cai S, et al. Severity of nausea and vomiting in pregnancy and early childhood neurobehavioural outcomes: The Growing Up in Singapore Towards Healthy Outcomes study. Paediatric and Perinatal Epidemiology. 2020;23:23.

76. Tan PC, Jacob R, Quek KF, Omar SZ. The fetal sex ratio and metabolic, biochemical, haematological and clinical indicators of severity of hyperemesis gravidarum. Bjog. 2006;113(6):733-7.

77. Tan PC, Jacob R, Quek KF, Omar SZ. Pregnancy outcome in hyperemesis gravidarum and the effect of laboratory clinical indicators of hyperemesis severity. J Obstet Gynaecol Res. 2007;33(4):457-64.

78. Tsang IS, Katz VL, Wells SD. Maternal and fetal outcomes in hyperemesis gravidarum. Int J Gynaecol Obstet. 1996;55(3):231-5.

79. Ustun Y, Engin-Ustun Y, Dokmeci F, Soylemez F. Serum concentrations of lipids and apolipoproteins in normal and hyperemetic pregnancies. J Matern Fetal Neonatal Med. 2004;15(5):287-90.

80. Vandraas KF, Vikanes AV, Vangen S, Magnus P, Stoer NC, Grjibovski AM. Hyperemesis gravidarum and birth outcomes-a population-based cohort study of 2.2 million births in the Norwegian Birth Registry. Bjog. 2013;120(13):1654-60.

81. Vikanes AV, Stoer NC, Magnus P, Grjibovski AM. Hyperemesis gravidarum and pregnancy outcomes in the Norwegian Mother and Child Cohort - a cohort study. BMC Pregnancy Childbirth. 2013;13:169.

82. Vilming B, Nesheim BI. Hyperemesis gravidarum in a contemporary population in Oslo. Acta Obstetricia et Gynecologica Scandinavica. 2000;79(8):640-3.

83. Vlachodimitropoulou Koumoutsea E, Gosh S, Manmatharajah B, Ray A, Igwe-Omoke N, Yoong W. Pregnancy outcomes in severe hyperemesis gravidarum in a multi-ethnic population. J Obstet Gynaecol. 2013;33(5):455-8.

84. Wang H, Rolls ET, Du X, Du J, Yang D, Li J, et al. Severe nausea and vomiting in pregnancy: psychiatric and cognitive problems and brain structure in children. BMC Med. 2020;18(1):228.

85. Zhang J, Cai WW. Severe vomiting during pregnancy: antenatal correlates and fetal outcomes. Epidemiology. 1991;2(6):454-7.

86. Zhang Y, Li Z, Zhang L, Liu J, Jin L, Ren A. Association between severe nausea and vomiting in early pregnancy and the risk of neural tube defects in Northern China. Birth Defects Res Part A Clin Mol Teratol. 2018;110(5):406-12.

87. Bayraktar B, Balikoglu M, Bayraktar MG, Kanmaz AG. The Effects of Hyperemesis Gravidarum on the Oral Glucose Tolerance Test Values and Gestational Diabetes. Prague Med Rep. 2021;122(4):285-93.

88. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Bmj. 2011;343:d4002.

89. Koot MH, Boelig RC, Van't Hooft J, Limpens J, Roseboom TJ, Painter RC, et al. Variation in hyperemesis gravidarum definition and outcome reporting in randomised clinical trials: a systematic review. Bjog. 2018;125(12):1514-21.

90. Belkacemi L, Nelson DM, Desai M, Ross MG. Maternal undernutrition influences placental-fetal development. Biol Reprod. 2010;83(3):325-31.

91. Che L, Yang Z, Xu M, Xu S, Che L, Lin Y, et al. Maternal nutrition modulates fetal development by inducing placental efficiency changes in gilts. BMC Genomics. 2017;18(1):213.

92. Heasman L, Clarke L, Firth K, Stephenson T, Symonds ME. Influence of restricted maternal nutrition in early to mid gestation on placental and fetal development at term in sheep. Pediatr Res. 1998;44(4):546-51.

93. Munro HN. Placental factors conditioning fetal nutrition and development. Am J Clin Nutr. 1981;34(Suppl 4):756-9.

94. Rasby RJ, Wettemann RP, Geisert RD, Rice LE, Wallace CR. Nutrition, body condition and reproduction in beef cows: fetal and placental development, and estrogens and progesterone in plasma. J Anim Sci. 1990;68(12):4267-76.

95. Rosso P. Placental growth, development, and function in relation to maternal nutrition. Fed Proc. 1980;39(2):250-4.

96. Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal nutrition and fetal development. J Nutr. 2004;134(9):2169-72.

97. Flenady V, Koopmans L, Middleton P, Froen JF, Smith GC, Gibbons K, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. Lancet. 2011;377(9774):1331-40.

98. Huang L, Sauve R, Birkett N, Fergusson D, van Walraven C. Maternal age and risk of stillbirth: a systematic review. CMAJ. 2008;178(2):165-72.

99. Po G, Oliver EA, Reddy UM, Silver RM, Berghella V. The impact of induction of labor at 39 weeks in low-risk women on the incidence of stillbirth. Am J Obstet Gynecol. 2020;222(1):88-90.

100. Petitti DB. Nausea and pregnancy outcome. Birth. 1986;13(4):223-6.

101. Weigel RM, Weigel MM. Nausea and vomiting of early pregnancy and pregnancy outcome. A meta-analytical review. Br J Obstet Gynaecol. 1989;96(11):1312-8.

102. Boneva RS, Moore CA, Botto L, Wong LY, Erickson JD. Nausea during pregnancy and congenital heart defects: a population-based case-control study. Am J Epidemiol. 1999;149(8):717-25.

103. Roseboom TJ, Ravelli AC, van der Post JA, Painter RC. Maternal characteristics largely explain poor pregnancy outcome after hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2011;156(1):56-9.

104. Jansen LAW, Koot MH, Van't Hooft J, Dean CR, Bossuyt PMM, Ganzevoort W, et al. The windsor definition for hyperemesis gravidarum: A multistakeholder international consensus definition. Eur J Obstet Gynecol Reprod Biol. 2021;266:15-22.

Appendix A. Search strategy

Ovid MEDLINE <1946 to February 08, 2022>
Search 2022-02-09

#	Searches	Results
1	hyperemesis gravidarum/	1707
2	((*morning sickness/ and (*vomiting/ or *nausea/)) or (*vomiting/ and *nausea/)) and pregnancy/	477
3	(hypereme* adj15 (pregnanc* or pregnant or gestat* or gravidi* or gravidar* or trimester* or maternal or prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or in-utero)).tw,kf.	1833
4	((pernicious* or serious* or sever* or excessiv*) adj2 (vomiting or nause*) adj9 (gravidar* or gravidit* or gestat* or pregnanc* or pregnant or trimester*)).tw,kf.	318
5	(nausea adj2 vomit* adj3 pregnan*).tw,kf.	890
6	**or/1-5 [hyperemesis gravidarum]**	**3105**
7	exp pregnancy outcome/ [incl stillbirth, live birth, spontaneous abortion]	80661
8	prenatal exposure delayed effects/ or maternal exposure/	39460
9	fetus/ or exp fetal heart/ or exp child/ or exp infant/ or puberty/ or schools/ or pediatrics/ [child /fetus]	2766192
10	child mortality/ or fetal mortality/ or exp infant mortality/	33310
11	embryo loss/ or fetal diseases/ or fetal macrosomia/ or fetal growth retardation/ or fetal hypoxia/ or fetal nutrition disorders/ or exp fetal death/ or exp fetal membranes, premature rupture/ or exp obstetric labor, premature/ or oligohydramnios/ or hydrops fetalis/ or perinatal death/ or placenta diseases/ or abruptio placentae/ or placental insufficiency/ or infant, newborn, diseases/ or asphyxia neonatorum/ or exp infant, premature, diseases/ or neonatal sepsis/ or jaundice neonatal/ or exp infant death/ [pregnancy complications/infant death]	214719
12	exp birth weight/ or fetal weight/ or cephalometry/ or crown-rump length/ or fetal distress/ or apgar score/	81925
13	child nutritional physiological phenomena/ or infant nutritional physiological phenomena/ or prenatal nutritional physiological phenomena/ or "growth and development"/ or exp human development/ or "embryonic and fetal development"/ or embryonic development/ or fetal development/ or fetal movement/ or fetal organ maturity/ or fetal viability/ or sex determination processes/ or sex differentiation/ or sexual development/ or language development/ or psychology, developmental/ or psychology, educational/ or exp education, special/ or exp child behavior/ or behavioral symptoms/ or neurobehavioral manifestations/	205292
14	adolescent health/ or child health/ or infant health/	6451
15	neurodevelopmental disorders/ or exp "attention deficit and disruptive behavior disorders"/ or exp autism spectrum disorder/ or obsessive-compulsive disorder/ or exp tic disorders/ or exp psychomotor performance/ or motor skills disorders/ or child behavior disorders/	229050

Continued

#	Searches	Results
16	exp aptitude tests/ or behavior rating scale/ or neuropsychological tests/ or language tests/ or exp "memory and learning tests"/ or stroop test/ or trail making test/	127023
17	sex factors/ or sex ratio/ or sex determination analysis/	290472
18	exp testicular diseases/	40079
19	exp musculoskeletal system/ab or exp heart/ab or exp nervous system/ab or genitalia/ab or abdominal wall/ab or urinary tract/ab or kidney/ab or urinary bladder/ab or kidney diseases/cn	88450
20	exp *congenital abnormalities/ or exp congenital abnormalities/et, ep	527110
21	abnormalities, severe teratoid/ or exp cardiovascular abnormalities/ or exp nervous system malformations/ or hydrocephalus/ or exp musculoskeletal abnormalities/ or exp bone diseases, developmental/ or cleft lip/ or exp digestive system abnormalities/ or exp respiratory system abnormalities/ or exp urogenital abnormalities/ or exp hydronephrosis/	522798
22	((pregnancy or gestat*) adj outcom*).tw,kf.	29042
23	((perinat* or peri-nat* or birth*1 or childbirth* or deliver* or labo?r* or obstetric*) adj3 outcome*).tw,kf.	36309
24	((perinat* or peri-nat*) adj3 (complicat* or health or morbidit* or cancer* or malignan* or neoplas*)).tw,kf.	10213
25	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj18 expos*).tw,kf.	30672
26	(maternal adj2 expos*).tw,kf.	8230
27	((prenat* or pre-nat* or antenat* or ante-nat* or in-utero or intra-uterine or intrauterine) adj3 (factor* or variabl* or parameter* or circumstanc* or condition*)).tw,kf.	5965
28	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj life).tw,kf.	1964
29	(DOHAD* or FOAD* or (early adj3 origin*)).tw,kf. or (development* adj3 origin* adj4 (health* or diseas* or adult)).tw,kf,jw.	6048
30	(offspring* or progeny or (born adj2 mother*)).tw,kf.	118792
31	(f?etal or f?etus* or neonat* or neo-nat* or new*born* or new-born* or child or child*1 or children* or schoolchild* or childhood or infant* or infanc* or toddler* or prekindergarten* or kindergarten* or preschool* or school-age* or schoolage* or high-school* or highschool* or elementary school* or graders or puber* or teens or teenager* or youth or juvenil* or adolescence or adulthood or young adult* or adult life or older age* or "early life" or later-life or "later in life").tw,kf. [child filter]	2742979
32	(((perinat* or peri-nat* or intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (mortalit* or death* or demise)) or stillbirth* or stillborn* or asphyx* or miscarriag* or IUFD or (spontan* adj3 abort*) or ((embry* or pregnancy) adj2 loss*) or liveborn* or (live adj3 (birth* or born*))).tw,kf.	111411

Continued

#	Searches	Results
33	(((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (growth* or develop* or brain or movement*)) or ((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj12 growth adj2 (restrict* or retard*)) or FGR* or IUGR* or SFGR* or SIUGR*).tw,kf.	31206
34	(placent* adj3 (insufficien* or d*sfunct* or inflammat* or abruptio*)).tw,kf.	8677
35	((PROM and ruptur* and (membran* or amnio*)) or PPROM* or EPPROM*1 or ((prematur* or pre-matur* or i?matur* or preterm* or pre-term* or pre-labo?r or prelabo?r) adj6 ruptur* adj4 (amnio* or membran*)) or chorioamn* or amnionit* or intraamnio* or funisit*).tw,kf. or (((ruptur* adj2 (amnio* or membran*)) or ROM).tw,kf. and (pregnan* or gestat* or gravidit* or trimester* or intrauterine* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*).mp.)	14626
36	(prematurity or ((preterm* or pre-term* or prematur* or pre-matur*) adj3 (labo?r or deliver* or birth* or childbirth*)) or PTB or PTBs or TPTB* or SPTB* or VPTB* or EPTB* or PTL or TPTL* or PTD*).tw,kf.	87937
37	(((small* or large* or deliver* or labo?r* or birth or childbirth) adj4 gestat* adj2 (age or ages)) or (gestat* adj ("at birth" or "at deliver*")) or birth age* or SGA or LGA).tw,kf.	35171
38	(((birth or births or childbirth* or born or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or peri-nat* or intrauterine or intra-uterine or in-utero) adj2 (underweight* or weight* or overweight* or siz* or length*)) or birthweight* or LBW* or VLBW* or ELBW*).tw,kf.	90273
39	((((head or cephal* or body or arm or arms or leg*) adj4 (circumfer* or measur* or siz* or small* or larg*)) or cephalometr* or anthropometr* or body mass* or BMI) adj9 (birth or births or childbirth* or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or per-nat* or intrauterine or intra-uterine or in-utero)).tw,kf.	10988
40	(neurodevelop* or ((brain or neurol* or neural or sex* or grow*) adj2 develop*) or "ages and stages" or DDST or (developmental adj4 (outcome* or test* or quotient* or index or indices or scor* or scale*)) or ((behav* or neurocognit* or cognit* or neurobehav* or neuropsychomot* or psychomotor* or psycho-motor* or neuromotor* or neuro-motor* or sensor?motor* or sensory motor* or visuomotor or visual motor or neuro-sensory or neurosensory) adj3 (abilit* or outcom* or problem* or develop*)) or ((executive or motor) adj3 (function* or d#sfunc* or deficit* or problem*)) or interference control or psychointellect* or intellect* or intelligen* or IQ or DQ or psycholinguist* or linguist*).tw,kf.	503983
41	(((language or learning or speech or reading or memory) adj3 (skill* or test* or scale* or scor* or task* or cognitiv*)) or verbal skill* or wording or naming or (numeric* adj3 memory) or 5-digit or digit-span or letter-digit).tw,kf.	118582

Continued

#	Searches	Results
42	(Binet* or Wechsler* or WAIS or WASI* or WIAT* or WDD or WDR or WRR or WISC or (WISC* not Wisconsin*) or WPPSI* or WRIT or CANTAB or complex figure or RCF or RCOF or RCFOSS or GIT-2 or FSIQ or VIQ or PIQ or (assess* adj2 batter*) or ((mobile or assesment*) adj3 ABC) or m-ABC or mABC or kABC or CELF or (mental development adj3 (index or indic* or scor*)) or MDI or WJ-R or Ba?ley* or BSID* or NEPSY or Beery or Basic Concept Scale or BBCS* or CMS or continuous performance* or Serial Addition or PDI or RBMT or Stroop or CNT or PPVT* or Everyday Attention or TOMI or MPC or CNT or Trail Making or Brunet or LMT or LMTs or TMT or TMTs or TMTa or Achievement Test or WJ-SAT or WJ IV or WRAT* or sensory profil* or ITSP or SSP or SPNL or CBCL).tw,kf.	117713
43	(((female or male or women or men or males or ratio* or distribut* or proportion* or factor*) adj3 (sex or gender*)) or ((male or males) adj1 female*) or girls or boys or ((deliver* or parturit* or birth or born*) adj2 (girl* or boy* or femal* or male or males))).tw,kf.	327258
44	((testi* adj3 (cancer* or neoplas* or malignan* or tumo?r* or undescen* or descen*)) or cryptorch*).tw,kf.	35934
45	(congenit* or anomal* or malformat* or deformit* or d#smorph* or aplas* or d#splas* or hypoplas* or atres* or agenes*).tw,kf.	665896
46	((prenat* or pre-nat* or antenat* or ante-nat* or perinat or peri-nat* or birth or anatomic* or morphological* or isolated or chromosom* or nonchromosom* or cardiac or noncardiac or extracard* or cardio* or heart or outflow tract or OFT or conotrunc* or cono-trunc* or septal or septum or endocard* cushion* or atrioventric* or atrio-ventr* or AV or musc*skelet* or skelet* or bone or bones or osseous or spine or spinal or limb or limbs or extremit* or foot or feet or hand or hands or cranio* or orofacial or facial* or palat*2 or mouth or lip or lips or (digest* adj2 (system or tract*)) or GI or intestin* or duoden* or esophag* or oesophag* or trach*esophag* or abdominal or respirator* or pulmonar* or lung or diaphragm* or hemidiaphragm* or sex or sexual or genit* or urogenit* or kidney* or uret* or renal or bladder or neural-tube* or nervous system or CNS or brain) adj3 (abnormalit* or defect*1)).tw,kf.	204999

Continued

#	Searches	Results
47	(Down* syndrome or CHD or Fallot* or Ebstein* or coarct* or (aort* adj1 arch*) or double outlet* or DORV or HLHS or HLV or HRHS or univentricular or uni-ventricular or single ventricle* or ((common arterial or arterios*) adj2 (trunk or truncus)) or VSD or (common adj3 (septum or septal) adj3 canal*) or AVSD or CAVC or scimitar* or TAPVC* or PAPVC* or encephaloc?el* or cephaloc?el* or meningoencephaloc?el* or notoencephaloc?ele or craniac?el* or ((cereb* or mening*) adj2 hernia*) or anencephal* or acrani* or aprosencephal* or ((spin* or cranium or crania) adj2 (bifid* or open)) or d#sraphi* or rachischis* or crani*-schis* or crani*schis* or myeloc?ele* or mening*myeloc?el* or hydrocephal* or hydro-cephal* or ventricul*-megal* or ventricul*megal* or holoprosencephal* or holo-prosencephal* or arhinencephal* or achondroplasi* or thanatophor* or osteochondrod#splas* or osteod#splast* or osteod#d#stroph* or chondrod#splas* or chondrod#stroph* or ((limb or limbs) adj2 reduct*) or talipes or clubfoot or club-foot or cleft or clefts or gastro?chis* or gastro-schis* or (umbilic* adj2 hernia*) or omphaloc?el* or exomphal* or ((cystic or polycyst* or multicyst*) adj2 (kidney* or (renal adj2 (diseas* or disorder*)))) or PKD or MCKD or megacystis or hydronephro* or Smith-Lemli-Opitz or micromelia or ectromeli* or (hydrops adj3 f?etalis) or hypospad* or hip dislocat*).tw,kf.	255421
48	**or/7-47 [perinatal and long term offspring outcomes]**	**5838396**
49	**6 and 48 [HG + perinatal and long term offspring outcome]**	**1214**
50	(editorial or "systematic review").pt. or (editorial or reply or (case-report not case-report-survey) or two-cases or three-cases or four-cases or five-cases or 2-cases or 3-cases or 4-cases or 5-cases).ti. or cochrane.jw. or ((review.pt. or case reports/ or case report*.jw. or (review or overview).ti. or (search* adj15 (literatur* or ((electronic* or medical or biomedical) adj3 database*) or medline or pubmed or embase or psyc?info or exhaustiv* or systematic*)).tw,kf,kw.) not (cohort studies/ or longitudinal studies/ or prospective studies/ or retrospective studies/ or cross-sectional studies/ or case-control studies/ or (case-control* or cohort* or retrospectiv* or prospectiv* or crosssection* or cross-section* or population-based or ((chart* or record* or retrospectiv*) adj3 review*)).tw,kf,kw.)) or (exp animals/ not exp humans/) or animal.jw. or (rodent* or rabbit* or mice or mouse or murine or rat or rats or (animal* adj3 (experiment* or model))).ti. [filter for original human studies]	10904951
51	**49 not 50 [HG + perinatal and long term offspring outcome - original human studies]**	**711**
52	**remove duplicates from 51 [HG + perinatal and long term offspring outcome -original human studies - duplicates removed]**	**709**

Embase Classic+Embase <1947 to 2022 February 08>
Search 2022-02-09

#	Searches	Results
1	hyperemesis gravidarum/	3315
2	((*"nausea and vomiting"/ or hyperemesis.dj. or (*vomiting/ and *nausea/)) and (pregnancy/ or pregnancy complication/ or prenatal period/ or prenatal exposure/)) or ((*vomiting/ or *nausea/) and (prenatal exposure/ or prenatal period/))	835
3	(hypereme* adj15 (pregnanc* or pregnant or gestat* or gravidi* or gravidar* or trimester* or maternal or prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or in-utero)).tw,kw.	2513
4	((pernicious* or serious* or sever* or excessiv*) adj2 (vomiting or nause*) adj9 (gravidar* or gravidit* or gestat* or pregnanc* or pregnant or trimester*)).tw,kw.	432
5	(nausea adj2 vomit* adj3 pregnan*).tw,kw.	1185
6	**or/1-5 [HG]**	**5046**
7	apgar score/ or exp birth weight/ or crown rump length/ or cephalometry/ or exp fetus maturity/ or fetus outcome/ or fetus weight/ or live birth/ or exp perinatal morbidity/ or placenta weight/ or pregnancy outcome/ ["parameters concerning the fetus, newborn and pregnancy"]	281613
8	maternal exposure/ or prenatal exposure/	29755
9	fetus/ or fetus brain/ or fetus heart/ or child/ or juvenile/ or exp infant/ or preschool child/ or school child/ or toddler/ or adolescence/ or adulthood/ or exp childhood/ or high school/ or kindergarten/ or middle school/ or primary school/ or pediatrics/ or progeny/ [fetus/child]	3614327
10	childhood mortality/ or embryo mortality/ or fetus mortality/ or infant mortality/ or exp perinatal mortality/ or prenatal mortality/ [child/fetus mortality]	71048
11	exp "immature and premature labor"/ or fetus disease/ or prenatal disorder/ or chorioamnionitis/ or dysmaturity/ or fetal malnutrition/ or fetus distress/ or fetus hypoxia/ or fetus malformation/ or exp hydramnios/ or exp intrauterine growth retardation/ or macrosomia/ or exp oligohydramnios/ or premature fetus membrane rupture/ or fetus wastage/ or spontaneous abortion/ or exp child death/ or embryo death/ or exp fetus death/ or perinatal death/ or placenta disorder/ or placenta insufficiency/ or solutio placentae/ or infant disease/ or newborn disease/ or dysmaturity/ or immaturity/ or large for gestational age/ or neonatal respiratory distress syndrome/ or neonatal stress/ or newborn apnea/ or newborn hypoxia/ or newborn infection/ or newborn sepsis/ or newborn vomiting/ or perinatal asphyxia/ or perinatal stress/ or prematurity/ or retrolental fibroplasia/ or newborn assessment/ or newborn intensive care/ or lung dysplasia/ or encephalomalacia/	462167
12	"growth, development and aging"/ or nerve cell differentiation/ or human development/ or adolescent development/ or language development/ or psychosocial development/ or speech development/ or exp postnatal development/ or prenatal development/ or embryo development/ or fetus development/ or fetal well being/ or fetus lung maturation/ or exp postnatal development/ or special education/ or exp sexual development/ or child behavior/ or neurobehavioral manifestations/ or neurodevelopment* outcome*.dq.	716008

Continued

#	Searches	Results
13	adolescent health/ or child health/	40253
14	behavior disorder/ or attention deficit disorder/ or congenital behavior disorder/ or exp disruptive behavior/ or oppositional defiant disorder/ or exp autism/ or exp tic/ or obsessive compulsive disorder/ or exp learning disorder/ or exp "disorders of higher cerebral function"/ or attention disturbance/ or developmental coordination disorder/ or exp intellectual impairment/ or language disability/ or exp developmental language disorder/ or developmental disorder/ or developmental delay/	1144031
15	behavior assessment/ or aptitude test/ or learning test/ or exp neuropsychological test/ or exp cognition assessment/ or developmental screening/	151368
16	"gender and sex"/ or gender/ or sex/ or sex difference/ or sex factor/ or sex ratio/ or exp sex determination/	854640
17	exp testis disease/	80421
18	exp *congenital disorder/ or exp congenital disorder/et, ep	987614
19	congenital malformation/ or exp "head and neck malformation"/ or exp limb malformation/ or exp cardiovascular malformation/ or severe teratoid abnormality/ or exp nervous system malformation/ or exp digestive system malformation/ or exp male genital tract malformation/ or exp musculoskeletal system malformation/ or hydronephrosis/	811231
20	((pregnancy or gestat*) adj outcom*).tw,kw.	41832
21	((perinat* or peri-nat* or birth*1 or childbirth* or deliver* or labo?r* or obstetric*) adj3 outcome*).tw,kw.	54107
22	((perinat* or peri-nat*) adj3 (complicat* or health or morbidit* or cancer* or malignan* or neoplas*)).tw,kw.	14553
23	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj18 expos*).tw,kw.	40025
24	(maternal adj2 expos*).tw,kw.	10597
25	((prenat* or pre-nat* or antenat* or ante-nat* or in-utero or intra-uterine or intrauterine) adj3 (factor* or variabl* or parameter* or circumstanc* or condition*)).tw,kw.	8328
26	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj life).tw,kw.	3495
27	(DOHAD* or FOAD* or (early adj3 origin*)).tw,kw. or (development* adj3 origin* adj4 (health* or diseas* or adult)).tw,kw,jw.	6669
28	(offspring* or progeny or (born adj2 mother*)).tw,kw.	144372
29	(f?etal or f?etus* or neonat* or neo-nat* or new*born* or new-born* or child or child*1 or children* or schoolchild* or childhood or infant* or infanc* or toddler* or prekindergarten* or kindergarten* or preschool* or school-age* or schoolage* or high-school* or highschool* or elementary school* or graders or puber* or teens or teenager* or youth or juvenil* or adolescence or adulthood or young adult* or adult life or older age* or "early life" or later-life or "later in life").tw,kw. [child filter]	3642221

Continued

#	Searches	Results
30	(((perinat* or peri-nat* or intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (mortalit* or death* or demise)) or stillbirth* or stillborn* or asphyx* or miscarriag* or IUFD or (spontan* adj3 abort*) or ((embry* or pregnancy) adj2 loss*) or liveborn* or (live adj3 (birth* or born*))).tw,kw.	165191
31	(((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (growth* or develop* or brain or movement*)) or ((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj12 growth adj2 (restrict* or retard*)) or FGR* or IUGR* or SFGR* or SIUGR*).tw,kw.	45969
32	(placent* adj3 (insufficien* or d*sfunct* or inflammat* or abruptio*)).tw,kw.	14134
33	((PROM and ruptur* and (membran* or amnio*)) or PPROM* or EPPROM*1 or ((prematur* or pre-matur* or i?matur* or preterm* or pre-term* or pre-labo?r or prelabo?r) adj6 ruptur* adj4 (amnio* or membran*)) or chorioamn* or amnionit* or intraamnio* or funisit*).tw,kw. or (((ruptur* adj2 (amnio* or membran*)) or ROM).tw,kw. and (pregnan* or gestat* or gravidit* or trimester* or intrauterine* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*).mp.)	21956
34	(prematurity or ((preterm* or pre-term* or prematur* or pre-matur*) adj3 (labo?r or deliver* or birth* or childbirth*)) or PTB or PTBs or TPTB* or SPTB* or VPTB* or EPTB* or PTL or TPTL* or PTD*).tw,kw.	128119
35	(((small* or large* or deliver* or labo?r* or birth or childbirth) adj4 gestat* adj2 (age or ages)) or (gestat* adj ("at birth" or "at deliver*")) or birth age* or SGA or LGA).tw,kw.	52693
36	(((birth or births or childbirth* or born or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or peri-nat* or intrauterine or intra-uterine or in-utero) adj2 (underweight* or weight* or overweight* or siz* or length*)) or birthweight* or LBW* or VLBW* or ELBW*).tw,kw.	124766
37	((((head or cephal* or body or arm or arms or leg*) adj4 (circumfer* or measur* or siz* or small* or larg*)) or cephalometr* or anthropometr* or body mass* or BMI) adj9 (birth or births or childbirth* or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or per-nat* or intrauterine or intra-uterine or in-utero)).tw,kw.	16774
38	(neurodevelop* or ((brain or neurol* or neural or sex* or grow*) adj2 develop*) or "ages and stages" or DDST or (developmental adj4 (outcome* or test* or quotient* or index or indices or scor* or scale*)) or ((behav* or neurocognit* or cognit* or neurobehav* or neuropsychomot* or psychomotor* or psycho-motor* or neuromotor* or neuro-motor* or sensor?motor* or sensory motor* or visuomotor or visual motor or neuro-sensory or neurosensory) adj3 (abilit* or outcom* or problem* or develop*)) or ((executive or motor) adj3 (function* or d#sfunc* or deficit* or problem*)) or interference control or psychointellect* or intellect* or intelligen* or IQ or DQ or psycholinguist* or linguist*).tw,kw.	660413
39	(((language or learning or speech or reading or memory) adj3 (skill* or test* or scale* or scor* or task* or cognitiv*)) or verbal skill* or wording or naming or (numeric* adj3 memory) or 5-digit or digit-span or letter-digit).tw,kw.	155524

Continued

#	Searches	Results
40	(Binet* or Wechsler* or WAIS or WASI* or WIAT* or WDD or WDR or WRR or WISC or (WISC* not Wisconsin*) or WPPSI* or WRIT or CANTAB or complex figure or RCF or RCOF or RCFOSS or GIT-2 or FSIQ or VIQ or PIQ or (assess* adj2 batter*) or ((mobile or assesment*) adj3 ABC) or m-ABC or mABC or kABC or CELF or (mental development adj3 (index or indic* or scor*)) or MDI or WJ-R or Ba?ley* or BSID* or NEPSY or Beery or Basic Concept Scale or BBCS* or CMS or continuous performance* or Serial Addition or PDI or RBMT or Stroop or CNT or PPVT* or Everyday Attention or TOMI or MPC or CNT or Trail Making or Brunet or LMT or LMTs or TMT or TMTs or TMTa or Achievement Test or WJ-SAT or WJ IV or WRAT* or sensory profil* or ITSP or SSP or SPNL or CBCL).tw,kw.	158466
41	(((female or male or women or men or males or ratio* or distribut* or proportion* or factor*) adj3 (sex or gender*)) or ((male or males) adj1 female*) or girls or boys or ((deliver* or parturit* or birth or born*) adj2 (girl* or boy* or femal* or male or males))).tw,kw.	491736
42	((testi* adj3 (cancer* or neoplas* or malignan* or tumo?r* or undescen* or descen*)) or cryptorch*).tw,kw.	50388
43	(congenit* or anomal* or malformat* or deformit* or d#smorph* or aplas* or d#splas* or hypoplas* or atres* or agenes*).tw,kw.	923861
44	((prenat* or pre-nat* or antenat* or ante-nat* or perinat or peri-nat* or birth or anatomic* or morphological* or isolated or chromosom* or nonchromosom* or cardiac or noncardiac or extracard* or cardio* or heart or outflow tract or OFT or conotrunc* or cono-trunc* or septal or septum or endocard* cushion* or atrioventric* or atrio-ventr* or AV or musc*skelet* or skelet* or bone or bones or osseous or spine or spinal or limb or limbs or extremit* or foot or feet or hand or hands or cranio* or orofacial or facial* or palat*2 or mouth or lip or lips or (digest* adj2 (system or tract*)) or GI or intestin* or duoden* or esophag* or oesophag* or trach*esophag* or abdominal or respirator* or pulmonar* or lung or diaphragm* or hemidiaphragm* or sex or sexual or genit* or urogenit* or kidney* or uret* or renal or bladder or neural-tube* or nervous system or CNS or brain) adj3 (abnormalit* or defect*1)).tw,kw.	269192

Continued

#	Searches	Results
45	(Down* syndrome or CHD or Fallot* or Ebstein* or coarct* or (aort* adj1 arch*) or double outlet* or DORV or HLHS or HLV or HRHS or univentricular or uni-ventricular or single ventricle* or ((common arterial or arterios*) adj2 (trunk or truncus)) or VSD or (common adj3 (septum or septal) adj3 canal*) or AVSD or CAVC or scimitar* or TAPVC* or PAPVC* or encephaloc?el* or cephaloc?el* or meningoencephaloc?el* or notoencephaloc?ele or craniac?el* or ((cereb* or mening*) adj2 hernia*) or anencephal* or acrani* or aprosencephal* or ((spin* or cranium or crania) adj2 (bifid* or open)) or d#sraphi* or rachischis* or crani*-schis* or crani*schis* or myeloc?ele* or mening*myeloc?el* or hydrocephal* or hydro-cephal* or ventricul*-megal* or ventricul*megal* or holoprosencephal* or holo-prosencephal* or arhinencephal* or achondroplasi* or thanatophor* or osteochondrod#splas* or osteod#splast* or osteod#d#stroph* or chondrod#splas* or chondrod#stroph* or ((limb or limbs) adj2 reduct*) or talipes or clubfoot or club-foot or cleft or clefts or gastro?chis* or gastro-schis* or (umbilic* adj2 hernia*) or omphaloc?el* or exomphal* or ((cystic or polycyst* or multicyst*) adj2 (kidney* or (renal adj2 (diseas* or disorder*)))) or PKD or MCKD or megacystis or hydronephro* or Smith-Lemli-Opitz or micromelia or ectromeli* or (hydrops adj3 f?etalis) or hypospad* or hip dislocat*).tw,kw.	354380
46	**or/7-45 [perinatal and long term offspring outcomes]**	**8694421**
47	**6 and 46 [HG + perinatal and long term offspring outcome]**	**2644**
48	editorial/ or "systematic review"/ or (editorial or conference abstract or conference review).pt. or (editorial or reply or (case-report not case-report-survey) or two-cases or three-cases or four-cases or five-cases or 2-cases or 3-cases or 4-cases or 5-cases).ti. or cochrane.jw. or ((review.pt. or review/ or case report/ or case report*.jw. or (review or overview).ti. or (search* adj15 (literatur* or ((electronic* or medical or biomedical) adj3 database*) or medline or pubmed or embase or psyc?info or exhaustiv* or systematic*)).tw,kw.) not (cohort analysis/ or longitudinal study/ or prospective study/ or retrospective study/ or exp case control study/ or cross-sectional study/ or (case-control* or cohort* or retrospectiv* or prospectiv* or crosssection* or cross-section* or population-based or ((chart* or record* or retrospectiv*) adj3 review*)).tw,kw.)) or ((exp animal/ or animal experiment/ or exp animal model/ or nonhuman/ or exp female animal/) not human/) or exp veterinary medicine/ or animal*.jw. or (rodent* or rabbit* or mice or mouse or murine or rat or rats or (animal* adj3 (experiment* or model))).ti. [filter for original human studies]	18100190
49	**47 not 48 [HG + perinatal and long term offspring outcome - original human studies]**	**1270**
50	**Remove duplicates from 49 [HG + perinatal and long term offspring outcome -original human studies - duplicates removed]**	**1252**
51	**50 not medline.cr. [HG + perinatal and long term offspring outcome - original human studies - duplicates removed - embase records only]**	**1066**

Appendix B. Study characteristics of excluded studies

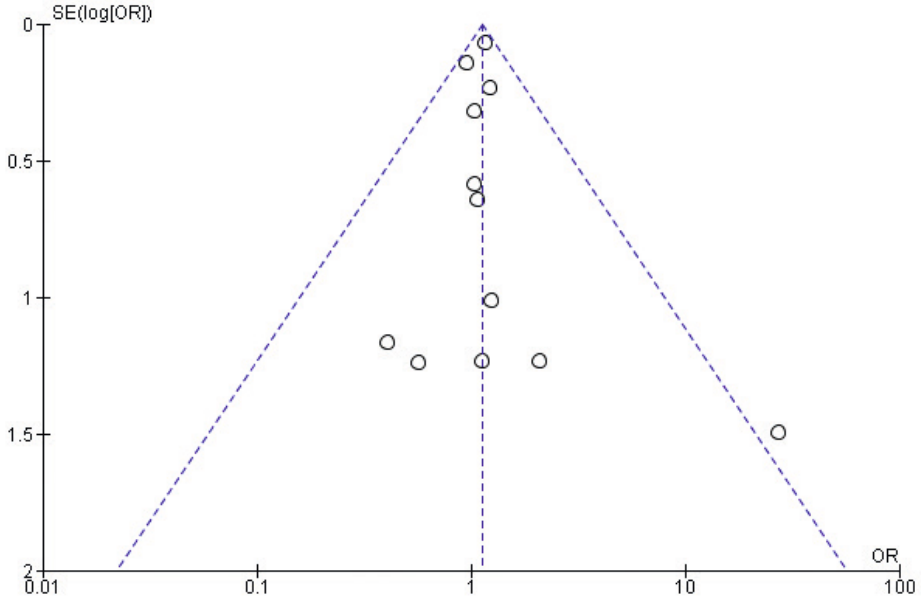
	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Reason of exclusion
1	Almond *et al.* (2016)	Sweden	Retrospective cohort	1987-2005	ICD-9 and ICD-10	17.840/ 1.652.699	Fetal gender Pregnancy loss	No additional data received, so not included in tables and meta-analysis
2	Axt-Fliedner *et al.* (2004)	USA	Retrospective case-control	1987-1996	*From medical records:* Admitted with HG in the first trimester	2.110/11.893	Fetal gender	No additional data received, so not included in tables and meta-analysis
3	Fejzo *et al.* (2019)	USA	Self-selected online survey cohort	2007-2017	*Self-reported:* HG diagnosis with treatment with IV fluids or TPN/ nasogastric feeding tube	267/360	Prematurity	Not included in tables and meta-analysis, because over overlap in data with Fejzo 2015.
4	Kallen *et al.* (1987)	Sweden	Retrospective cohort	1973-1981	ICD-8	3.068/ unknown	Birth weight Congenital anomalies Fetal gender Gestational age at delivery Infant deaths Multiple pregnancy Prematurity	No additional data received, so not included in tables and meta-analysis.
5	Lu *et al.* (2015)	China	Retrospective case-control	2006-2008	*Self-reported:* Questionnaires, severe NVP defined as continuous symptoms of nausea and vomiting with the need for bedrest or hospitalization	81/672	Congenital anomalies	Not included in tables and meta-analysis, because over overlap in data with Zhang 2017.

Continued

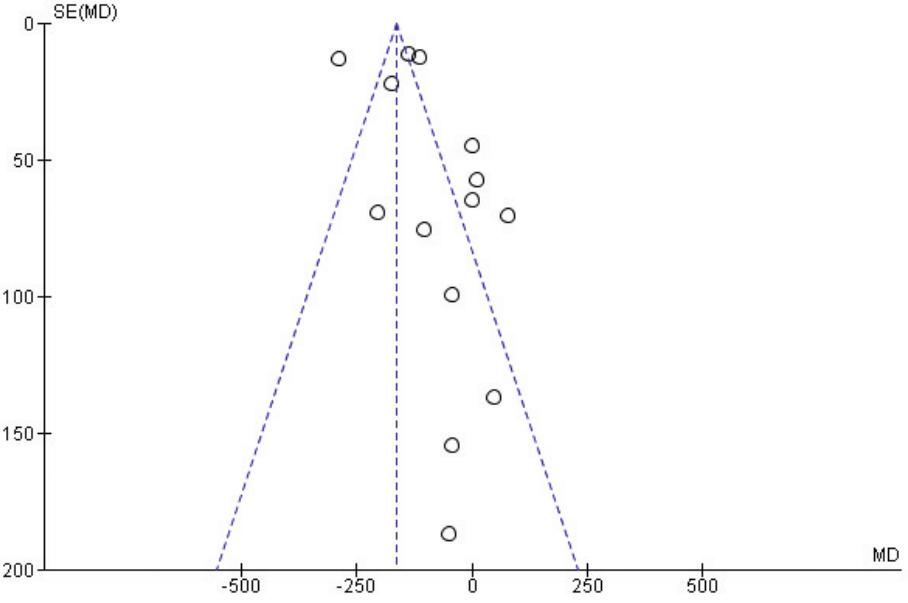
	Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Perinatal outcomes	Reason of exclusion
6	McCarthy *et al.* (2011)	New Zealand, Australia, Ireland, United Kingdom	Prospective cohort	2004-2008	*Self-reported:* Interview, HG diagnosis with treatment with IV fluids or TPN/ nasogastric feeding tube or loss of >5% of prepregnancy weight	164/3.423	Birth weight Fetal gender Prematurity	Only odds ratios provided. No additional data received, so not included in tables and meta-analysis.
7	Peled *et al.* (2013)	Israel	Retrospective cohort	1994-2008	*From medical records:* Hospitalization for HG	545/2.180	Apgar Birth weight Fetal gender Gestational age at delivery Infant deaths NICU admission Placental abruption Postpartum neonatal complications	Not included in tables and meta-analysis, because over overlap in data with Peled 2014.
8	Poeran *et al.* (2019)	The Netherlands	Prospective cohort	2001-2005	*Self-reported:* Daily vomiting	463/4.769	Birth weight Fetal gender Gestational age at delivery	Not included in tables and meta-analysis, because over overlap in data with Roseboom 2011 and Grooten 2017
9	Syn *et al.* (2020)	Malaysia	population-based prospective longitudinal cohort	2009-2010	*Self-reported and medical records:* structured interview-administered questionnaire combined with medical records. Regular vomiting with inability to retain meals	190/486	Birth weight Fetal gender Gestational age at delivery	Not included in tables and meta-analysis, because over overlap in data with Fejzo 2015.

Appendix C. Funnel plots

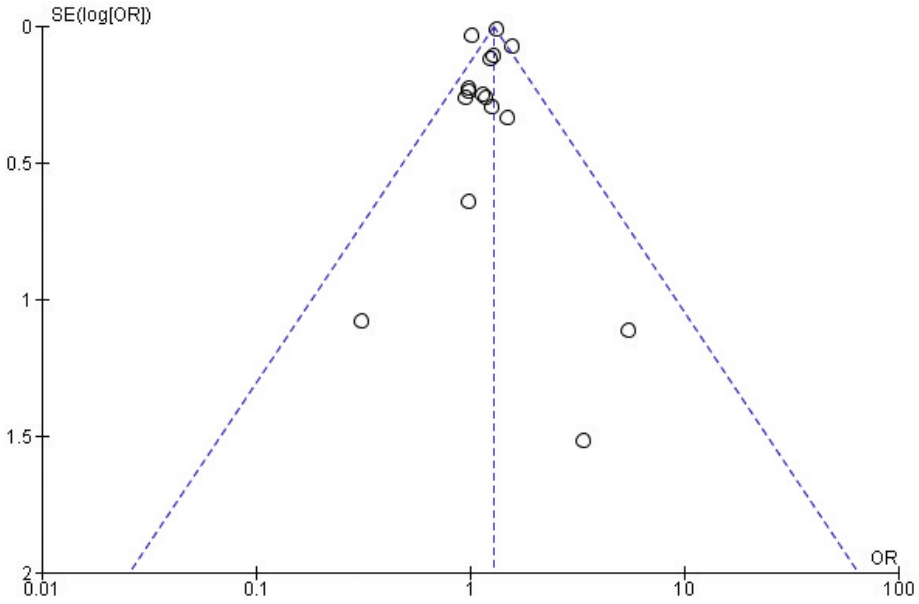
C.1 Funnel plot Apgar score <7 at 5 minutes



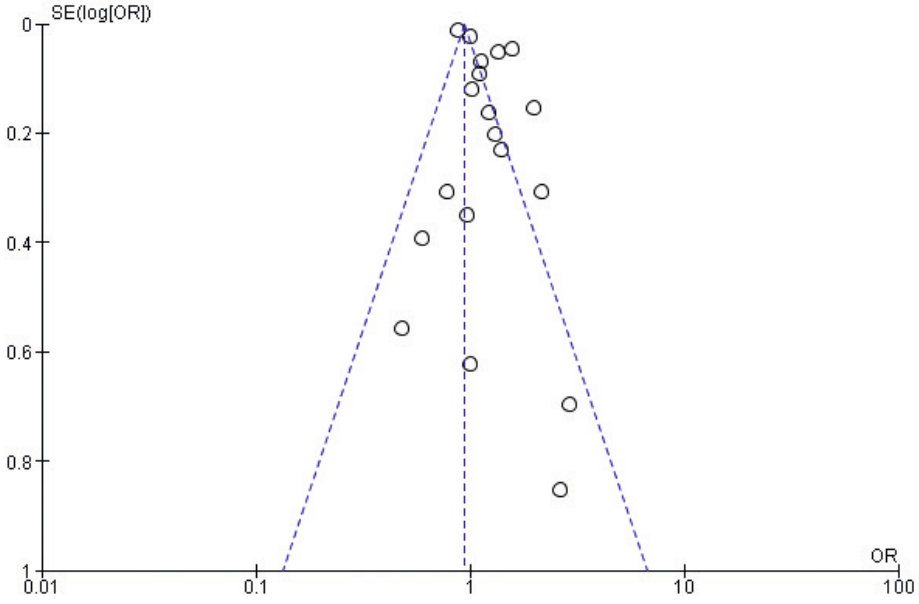
C.2 Funnel plot birth weight in grams



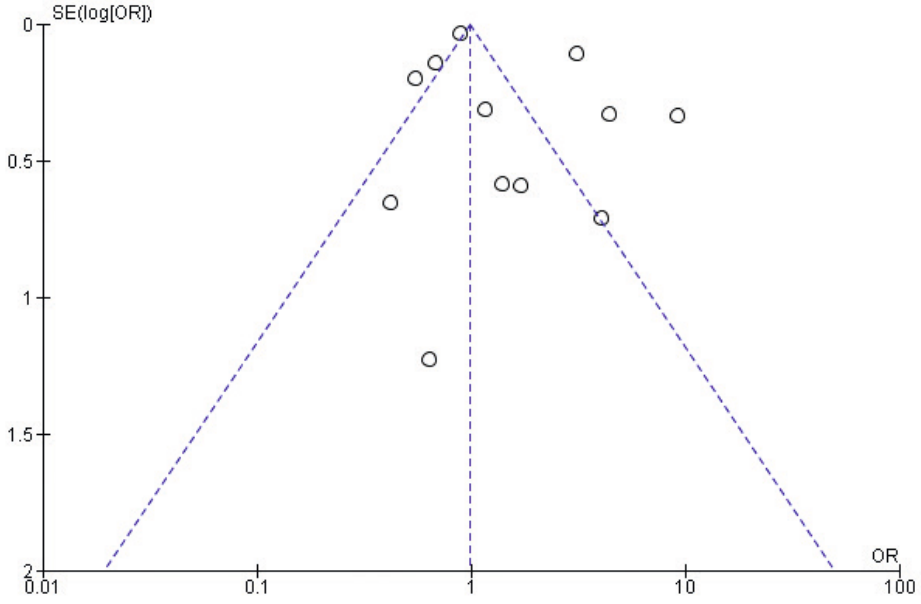
C.3 Funnel plot birth weight <2500 grams



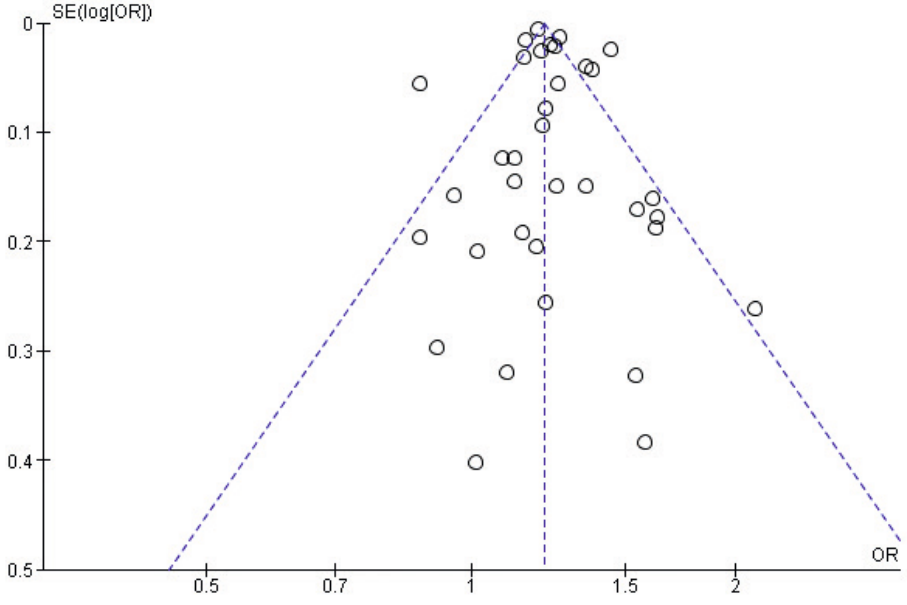
C.4 Funnel plot small for gestational age



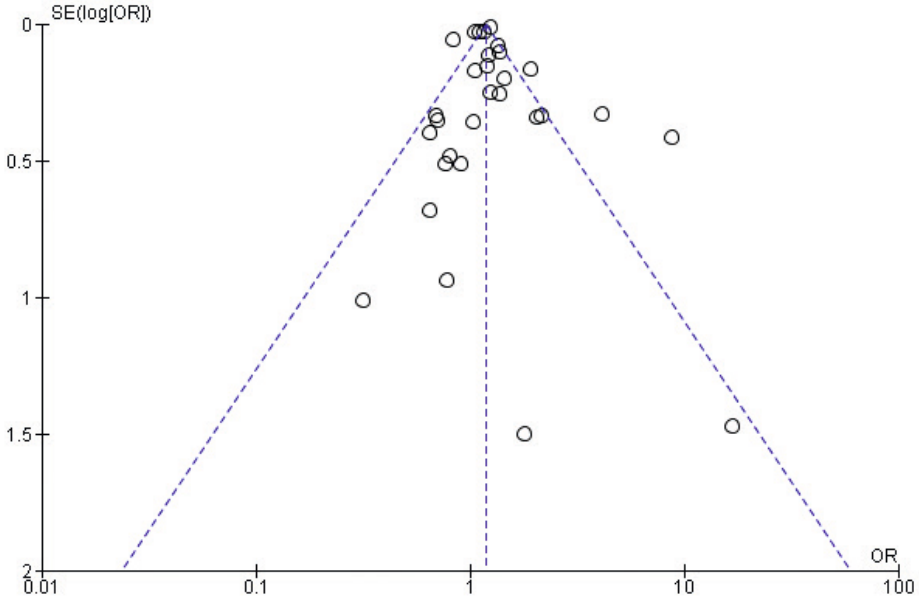
C.5 Funnel plot congenital anomalies



C.6 Funnel plot fetal sex



C.7 Funnel plot prematurity <37 weeks



Appendix D. Overview of results of included studies reporting on Apgar score, asphyxia and pH

Study	HG-exposed/ total sample size	HG-exposed Mean ± SD; Frequencies (%)	Non-exposed Mean ± SD; Frequencies (%)	P
	Apgar <7 at 1 minute			
Agmon*	Total:89/180	9 (10.1%)		0.256
	Mild HG: 46/137	4 (8.7%)	5 (5.6%)	not shown
	Severe HG: 43/ 134	5 (11.6%)		not shown
Caltekin	52/112	3 (5.8%)	4 (6.7%)	0.845
Coetzee	62/202	7 (11.3%)	14 (10%)	not significant
Vikanes	814/71,468	26 (3.2%)	3,527 (5.0%)	significant
	Apgar <7 at 5 minutes			
Agmon*	Total:89/180	2 (2.2%)	1 (1.1%)	0.5
	Mild HG: 46/137	0 (0%)		not shown
	Severe HG: 43/ 134	2 (4.7%)		not shown
Caltekin	52/112	1 (1.9%)	2 (3.3%)	0.641
Coetzee	62/202	1 (1.6%)	2 (1.4%)	not significant
Dodds	Total: 1,270/156,091	19 (1.5%)	1,898 (1.2%)	not significant
	1-2 admissions for HG: 1,182/155,903	18 (1.5%)		not significant
	3+ admissions for HG: 88/154,909	1 (1.2%)		not significant
	Weight gain ≥ 7kg: 885/128,720	5 (0.6%)	1,471 (1.2%)	not significant
	Weight gain < 7kg: 144/127,979	8 (5.6%)		significant
Gunay	186/386	5 (2.7%)	5 (2.5%)	not significant
Hallak**	Total: 138/12,473	3 (2.2%)	259 (2.1%)	not shown
	Mild HG: 40/12,375	2 (5%)		not significant
	Severe HG: 98/12,433	1 (1.0%)		not significant
Koot	62/8,953	1 (1.6%)	116 (1.3%)	0.86
Kuru	72/161	1 (1.4%)	3 (3.4%)	not significant
Peled 2014	Total: 599/2,396	4 (0.7%)	0 (0%)	<0.001
	No TPN: 422/2,219	3 (0.7%)		not shown
	TPN: 122/1,919	1 (0.6%)		not shown
Roseboom	2,186/1,193,705	50 (2.3%)	28,596 (2.4%)	0.77
Vandraas	20,004/2,266,345	199 (1.0%)	19,421 (0.9%)	0.715
Vikanes	814/71,468	10 (1.2%)	838 (1.2%)	not significant

Continued

Study	HG-exposed/ total sample size	HG-exposed Mean ± SD; Frequencies (%)	Non-exposed Mean ± SD; Frequencies (%)	P
Apgar score at 1 minute				
Tsang	193/13,053	7.5	7.5	not significant
Vlachodimitropoulou	208/416	9	9	not significant
Apgar score at 5 minutes				
Koudijs***	Total: 400/2,238	9 ±0	9 ±0	Not shown
	Mild HG: 354/2,192	9 ±0		0.24
	Severe HG: 46/1,884	9 ±0		0.67
Tsang	193/13,053	8.5	8.5	not significant
Vlachodimitropoulou	208/416	9	9	not significant
pH at birth				
Agmon*	Total: 89/180	7.3 ±0.95	7.29 ±0.92	0.23
	Mild HG: 46/137	7.31 ±0.10		not shown
	Severe HG: 43/137	7.30 ±0.89		not shown
Asphyxia				
Chin 1989****	Mild asphyxia: 5/366	0 (0%)	9 (2.5%)	not shown
	Severe asphyxia: 5/366	0 (0%)	1 (0.3%)	not shown

* Severe HG: ≥2 of the following: (1) ≥3 hospitalizations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria
** Severe HG: presence of ≥1 of the following: ketonuria, increased BUN and hematocrit or abnormal electrolytes
*** Severe HG: >5% weight loss compared to pre pregnancy weight
**** Mild asphyxia: Apgar 4-6, severe asphyxia: Apgar <4

Appendix E. Overview of results of included studies reporting on birth weight

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
		Birth weight in grams			
Agmon	89/180	3,076.06 ± 449.97	3,178.90 ± 559.32	0.176	
Bailit	2,270/488,775	3,255	3,38	**<0.0001**	No SD provided
Bashiri	164/373	3,122	3,145	not significant	No SD provided
Caltekin	52/112	3,060 ± 321	3,265 ± 414	**0.004**	
Chin 1988*	Severe HG: 46/8,848	2,950 ± 475	3,221 ± 452	**<0.001**	
	Mild HG: 26/8,828	3,159 ± 314		**<0.05**	
Chin 1989	5/366	3,142 ± 341	3,186 ± 447	not significant	
Coetzee	All: 62/202	3,440 (1,980-4,815)	3,540 (1,010-5,250)	not significant	No SD provided
	Born at term: 59/189	3,465 (1,980-4,815)	3,580 (2,040-5,250)	not significant	
Depue study 2	35/70	3,345 ± 652.2	3,299 ± 480.5	0.79	
Fejzo 2013	254/562	3,236.69	3,446.01	**<0.0001**	No SD provided
Gunay	186/386	3,250 (2,850-3,610)	3,275 (2,915-3,640)	0.698	No SD provided
Hallak**	Mild HG: 40/12,375	3,110 ± 641	3,160 ± 652	not significant	
	Severe HG: 98/ 12,433	3,093 ± 636		not significant	
Kidess	65/130	3,305 ± 79	3,442 ± 53	not shown	
Koot	62/8,953	3,552 ± 507	3,552 ± 558	0.99	
Koudijs****	Mild HG: 354/2,187	3,116 ± 464.4	3,100 ± 476,6	0.6	
	Severe HG: 46/1,879	3,046 ± 485.7		0.5	
Kuru	72/161	3,121 ± 595.4	3,164 ± 664.5	0.67	
Muraoka	34/104	3,148.9 ± 348	3,070.9 ± 316	0.98	
Ozay	46/100	2,850 ± 980	2,900 ± 870	0.32	
Paauw	45/351	3,232 ± 88	3,523 ± 29	**<0.05**	
Peled 2014	Total: 599/2,396	3,074 ± 456	3,248 ± 543	**<0.001**	
	No TPN: 422/2,219	3,056 ± 443		not shown	
	TPN: 122/1,919	3,145 ± 487		not shown	

Continued

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
Roseboom	Total: 2,190/ 1,197,028	3,337 ± 587	3,453 ± 579	**<0.0001**	
	Live born only: 2,186/ 1,191,519	3,337 ± 587	3,459 ± 569	**<0.0001**	
Salunkhe	76/1,035	2,658.8 ± 484	2,648.8 ± 455.9	not shown	
Tan 2007	166/664	3,000 ± 500	3,000 ± 500	0.3	
Vandraas	20,004/2,266,345	3,501.1	3,525.9	not shown	No SD provided
Vikanes	Total: 814/71,468	3,567	3,602	not significant	No SD provided
	HG 1st trimester: 484/71,138	3,591		not significant	
	HG 2nd trimester: 173/70,827	3,506		not significant	
	HG 1st and 2nd trimester: 114/ 70,768	3,563		not significant	
Vilming	120/235	3,440	3,578	**0.04**	No SD provided
		Birth weight centile			
Ong	Severe vomiting with hospitalization: 67/363	51.3 (22.2-80.7)	53.5 (27.9-75.5)	not significant	No SD available
	Severe vomiting: 123/ 419	54.2 (31.1-78.5)		not significant	
Peled 2014	Total: 599/2,396	44.8 ± 28.3	52.4 ±27	**<0.001**	
	No TPN: 422/2,219	43.6 ± 28.6		not shown	
	TPN: 122/1,919	49.6 ± 27.1		not shown	
		Birth weight <2500 grams			
Bailit	2,270/488,775	177 (7.8%)	24,812 (5.1%)	**<0.0001**	
Buyukkayaci	50/100	5 (10%)	1 (2%)	>0.05	
Chin 1989	5/366	0 (0%)	9 (2.5%)	not significant	

Continued

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
Coetzee	All: 62/202	1 (1.6%)	7 (5%)	not significant	
	Born at term: 59/189	1 (1.7%)	2 (1.5%)	not significant	
Dodds	All: 1,270/ 156,091	72 (5.7%)	7,143 (2.6%)	not significant	
	1-2 admissions: 1,182/ 155,903	64 (5.4%)		not significant	
	3+ admissions: 88/154,909	8 (9.1%)		**significant**	
	Weight gain ≥ 7kg: 885/ 128,720	30 (3.4%)	5,326 (4.2)	not significant	
	Weight gain < 7kg: 144/ 127,979	18 (12.5%)		**significant**	
Fiaschi	total: 80,708/ 5,491,227	6,091 (7.5%)	312,276 (5.7%)	not shown	
	1 admission: 57,030/ 5,467,549	4,087 (7.2%)		**significant**	
	readmission: 23,678/ 5,434,197	2,004 (8.5%)		**significant**	
Hastoy***	Total:197/589	26 (13.2%)	45 (11.5%)	0.52	
	Mild HG: 137/529	14 (10%)		not significant	
	Severe HG: 60/452	12 (20%)		not significant	
Klebanoff	184/3,469	19 (10.1%)	299 (9.1%)	not significant	
Koudijs****	Total: 316/1,681	23 (7.3%)	101 (7.4%)	not shown	
	Mild HG: 280/1,645	21 (7.5%)		0.92	
	Severe HG:36/1,401	2 (5.6%)		0.68	
Ozay	46/100	5 (10.9%)	6 (11.1%)	0.4	
Salunkhe	76/1,035	22 (28.9%)	288 (30.0%)	not shown	
Schiff	2,110/11,893	106 (5%)	386 (4%)	not shown	
Tan 2007	166/664	18 (10.8%)	44 (8.8%)	0.44	

Continued

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
Vikanes	814/71,468	20 (2.5%)	1.771 (2.5%)	not significant	
Wang	20,986/ 2,062,854	840 (4.0%)	80,443 (3.9%)	not shown	
Zhang	201/1,867	11 (5.5%)	62 (3.7%)	not significant	
		Birth weight <1500 grams			
Bailit	2,270/488,775	35 (1.54%)	5,352 (1.1%)	**0.04**	
Chin 1989	5/366	0 (0%)	2 (0.6)	not significant	
		Small for gestational age			
Agmon	89/180	5 (5.6%)	2 (2.2%)	0.213	birth weight < 10th centile
Bailit	2,270/488,775	663 (29.21%)	101,193 (20.8%)	**<0.0001**	birth weight < 10th centile
Bolin	12,270/ 1,155,033	381 (3.1%)	26,683 (2.4%)	**significant**	>2 SD below mean birth weight
Coetzee	All: 62/202	4 (6.5%)	9 (6.4%)	not significant	birth weight < 10th centile
	Born at term: 59/189	4 (6.8%)	7 (5.4%)	not significant	birth weight < 10th centile
Dodds	All: 1,270/156,091	137 (10.8%)	15,217 (10%)	not significant	birth weight < 10th centile
	1-2 admissions: 1,182/155,903	126 (10.7%)		not significant	
	3+ admissions: 88/ 154,909	8 (9.1%)		not significant	
	Weight gain ≥ 7kg: 885/128,720	81 (9.2%)	12,541 (10%)	not significant	
	Weight gain < 7kg: 144/ 127,979	21 (14.6%)		not significant	
Fejzo 2013	254/562	7 (2.8%)	3 (1%)	0.1976	birth weight < 10th centile
Fiaschi	Total: 66,797/ 5,491,227	7,662 (11.5%)	704,297 (16.1%)	not shown	birth weight < 10th centile
	1 admission: 47,037/ 4,435,152	5,202 (11%)		significant	
	readmission: 19,760/ 4,407,875	2,460 (12.5%)		significant	
Gunay	186/386	11 (5.9%)	19 (9.5%)	0.253	birth weight < 10th centile

Continued

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
Hastoy***	Total:197/589	52 (26.4%)	85 (21.7%)	0.2	birth weight < 10th centile
	Mild HG: 137/529	32 (23.4%)		not significant	
	Severe HG: 60/452	20 (33.3%)		**significant**	
Koudijs****	Total: 307/1,584	60 (19.5%)	212 (16.6%)	not shown	birth weight < 10th centile
	Mild HG: 273/1,550	54 (19.8%)		0.2	
	Severe HG:34/1,311	6 (17.6%)		0.87	
Kuru	72/161	5 (6.9%)	12 (13.5%)	0.16	birth weight < 10th centile
Morokuma*****	HG: 128/6,190	9 (7.0%)	440 (7.3%)	not significant	birth weight < 10th centile
	Severe NVP: 832/7,975	59 (7.1%)	518 (7.3%)	not significant	
Ong	Total: 190/486	18 (9.5%)	35 (11.8%)	not shown	birth weight < 10th centile
	Severe vomiting with hospitalization: 67/363	8 (11.9%)		not significant	
	Severe vomiting: 123/419	10 (8.1%)		not significant	
Peled 2014	Total: 599/2,396	76 (12.7%)	122 (6.8%)	**<0.001**	birth weight < 10th centile
	No TPN: 422/2,219	62 (14.7%)		not shown	
	TPN: 122/1,919	40 (9.5%)		not shown	
Roseboom	Total: 2,190/ 1,197,028	237 (10.8%)	117,309 (9.8%)	0.11	birth weight < 10th centile
	Live born only: 2,186/ 1,191,519	236 (10.8%)	115,577 (9.7%)	0.09	
Vandraas	20,004/ 2,266,345	2,115 (10.6%)	235,954 (10.5%)	0.757	birth weight < 10th centile
Vikanes	814/71,468	78 (9.6%)	6,715 (9.5%)	not significant	birth weight < 10th centile
Vlachodimitropoulou	208/416	35 (16.8%)	18 (8.7%)	**<0.05**	birth weight < 10th centile
Zhang	201/1,867	25 (12.4%)	153 (9.2%)	not significant	birth weight < 10th centile

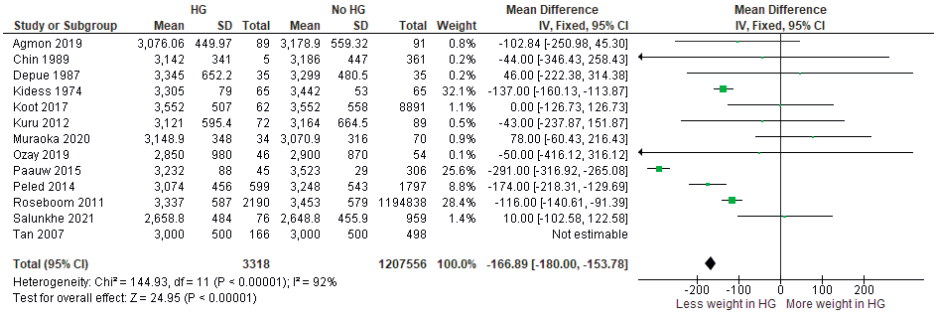
Continued

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comments
		Birth weight >4000 grams			
Fiaschi	Total: 80,708/ 5,491,227	7,102 (8.8%)	627,045 (11.6%)	not shown	
	1 admission: 57,030/ 5,467,549	5,147 (9%)		**significant**	
	Readmission: 23,678/ 5,434,197	1,955 (8.3%)		**significant**	
Schiff	2,110/11,893	252 (11.9%)	1,399 (14.3%)	0.002	
		Large for gestational age			
Bailit	2,270/488,775	222 (9.8%)	54,498 (11.2%)	**0.04**	birth weight > 90th centile
Fiaschi	Total: 66,797/ 5,491,227	8,881 (13.3%)	704,297 (16.1%)	not shown	birth weight > 90th centile
	1 admission: 47,037/ 4,435,152	6,408 (13.6%)		not significant	
	Readmission: 19,760/ 4,407,875	2,473 (12.5%)		**significant**	
Ong	Total: 190/486	29 (15.3%)	32 (10.8%)	not shown	birth weight > 90th centile
	Severe vomiting with hospitalization: 67/363	12 (17.9%)		not significant	
	Severe vomiting: 123/419	17 (13.8%)		not significant	
Peled 2014	Total: 599/2,396	48 (8.2%)	174 (9.7%)	0.3	birth weight > 90th centile
	No TPN: 422/2,219	40 (9.5%)		not shown	
	TPN: 122/1,919	9 (5.1%)		not shown	
Vandraas	20,004/2,266,345	2,158 (10.8%)	7,497 (0,3%)	0.36	birth weight > 90th centile
Vlachodimitropoulou	208/416	15 (7.2%)	19 (9.1%)	not significant	birth weight > 90th centile

* Severe HG: heavy ketonuria (>3+), increase in urea and creatinine concentrations, serum electrolyte disturbance or/and increase in hematocrit (> 0.43)
**Severe HG: heavy ketonuria (>3+), abnormal serum BUN, sodium or potassium or increase in hematocrit (> 0.43)
*** Severe HG: weight gain less than 7kg
***** Severe NVP: vomiting and unable to eat. HG: severe NVP and >5% weight loss compared to pre pregnancy weight

Appendix F. Sensitivity analysis birth weight

F.1 Sensitivity analysis birth weight only cohort studies

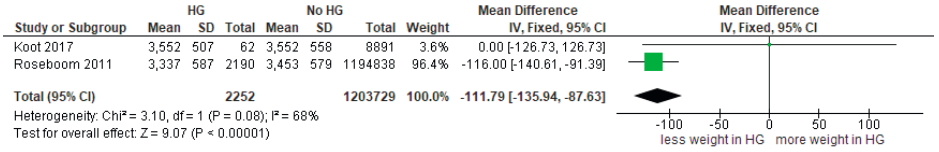


Study or Subgroup	HG Mean	HG SD	HG Total	No HG Mean	No HG SD	No HG Total	Weight	Mean Difference IV, Fixed, 95% CI
Agmon 2019	3,076.06	449.97	89	3,178.9	559.32	91	0.8%	-102.84 [-250.98, 45.30]
Chin 1989	3,142	341	5	3,186	447	361	0.2%	-44.00 [-346.43, 258.43]
Depue 1987	3,345	652.2	35	3,299	480.5	35	0.2%	46.00 [-222.38, 314.38]
Kidess 1974	3,305	79	65	3,442	53	65	32.1%	-137.00 [-160.13, -113.87]
Koot 2017	3,552	507	62	3,552	558	8891	1.1%	0.00 [-126.73, 126.73]
Kuru 2012	3,121	595.4	72	3,164	664.5	89	0.5%	-43.00 [-237.87, 151.87]
Muraoka 2020	3,148.9	348	34	3,070.9	316	70	0.9%	78.00 [-60.43, 216.43]
Ozay 2019	2,850	980	46	2,900	870	54	0.1%	-50.00 [-416.12, 316.12]
Paauw 2015	3,232	88	45	3,523	29	306	25.6%	-291.00 [-316.92, -265.08]
Peled 2014	3,074	456	599	3,248	543	1797	8.8%	-174.00 [-218.31, -129.69]
Roseboom 2011	3,337	587	2190	3,453	579	1194838	28.4%	-116.00 [-140.61, -91.39]
Salunkhe 2021	2,658.8	484	76	2,648.8	455.9	959	1.4%	10.00 [-102.58, 122.58]
Tan 2007	3,000	500	166	3,000	500	498		Not estimable
Total (95% CI)			**3318**			**1207556**	**100.0%**	**-166.89 [-180.00, -153.78]**

Heterogeneity: Chi² = 144.93, df = 11 (P < 0.00001); I² = 92%
Test for overall effect: Z = 24.95 (P < 0.00001)

Mean Difference IV, Fixed, 95% CI
-200 -100 0 100 200
Less weight in HG More weight in HG

F.2 Sensitivity analysis birth weight only European studies

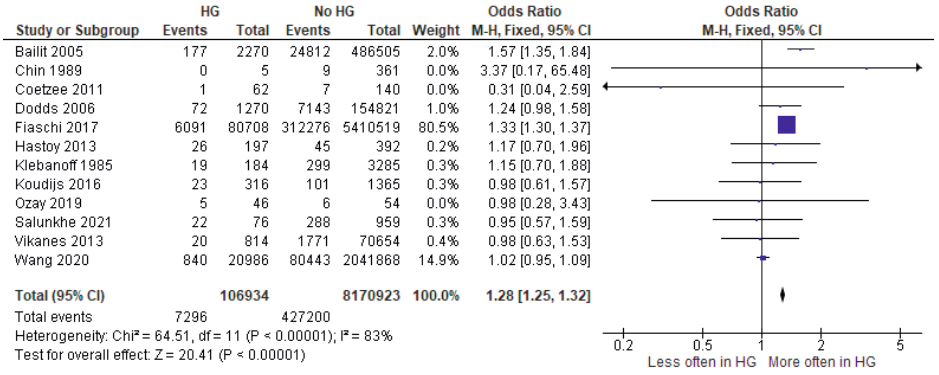


Study or Subgroup	HG Mean	HG SD	HG Total	No HG Mean	No HG SD	No HG Total	Weight	Mean Difference IV, Fixed, 95% CI
Koot 2017	3,552	507	62	3,552	558	8891	3.6%	0.00 [-126.73, 126.73]
Roseboom 2011	3,337	587	2190	3,453	579	1194838	96.4%	-116.00 [-140.61, -91.39]
Total (95% CI)			**2252**			**1203729**	**100.0%**	**-111.79 [-135.94, -87.63]**

Heterogeneity: Chi² = 3.10, df = 1 (P = 0.08); I² = 68%
Test for overall effect: Z = 9.07 (P < 0.00001)

Mean Difference IV, Fixed, 95% CI
-100 -50 0 50 100
less weight in HG more weight in HG

F.3 Sensitivity analysis birth weight <2500 grams only cohort studies

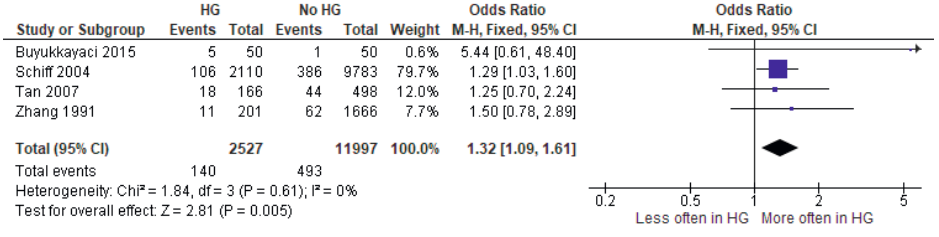


Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bailit 2005	177	2270	24812	486505	2.0%	1.57 [1.35, 1.84]
Chin 1989	0	5	9	361	0.0%	3.37 [0.17, 65.48]
Coetzee 2011	1	62	7	140	0.0%	0.31 [0.04, 2.59]
Dodds 2006	72	1270	7143	154821	1.0%	1.24 [0.98, 1.58]
Fiaschi 2017	6091	80708	312276	5410519	80.5%	1.33 [1.30, 1.37]
Hastoy 2013	26	197	45	392	0.2%	1.17 [0.70, 1.96]
Klebanoff 1985	19	184	299	3285	0.3%	1.15 [0.70, 1.88]
Koudijs 2016	23	316	101	1365	0.3%	0.98 [0.61, 1.57]
Ozay 2019	5	46	6	54	0.0%	0.98 [0.28, 3.43]
Salunkhe 2021	22	76	288	959	0.3%	0.95 [0.57, 1.59]
Vikanes 2013	20	814	1771	70654	0.4%	0.98 [0.63, 1.53]
Wang 2020	840	20986	80443	2041868	14.9%	1.02 [0.95, 1.09]
Total (95% CI)		**106934**		**8170923**	**100.0%**	**1.28 [1.25, 1.32]**
Total events	7296		427200			

Heterogeneity: Chi² = 64.51, df = 11 (P < 0.00001); I² = 83%
Test for overall effect: Z = 20.41 (P < 0.00001)

Odds Ratio M-H, Fixed, 95% CI
0.2 0.5 1 2 5
Less often in HG More often in HG

F.4 Sensitivity analysis birth weight <2500 grams only case-control studies

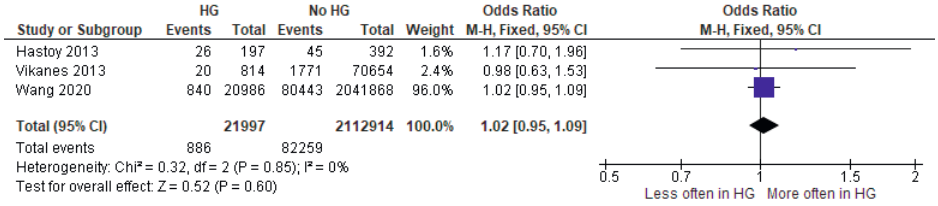


Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Buyukkayaci 2015	5	50	1	50	0.6%	5.44 [0.61, 48.40]
Schiff 2004	106	2110	386	9783	79.7%	1.29 [1.03, 1.60]
Tan 2007	18	166	44	498	12.0%	1.25 [0.70, 2.24]
Zhang 1991	11	201	62	1666	7.7%	1.50 [0.78, 2.89]
Total (95% CI)		**2527**		**11997**	**100.0%**	**1.32 [1.09, 1.61]**
Total events	140		493			

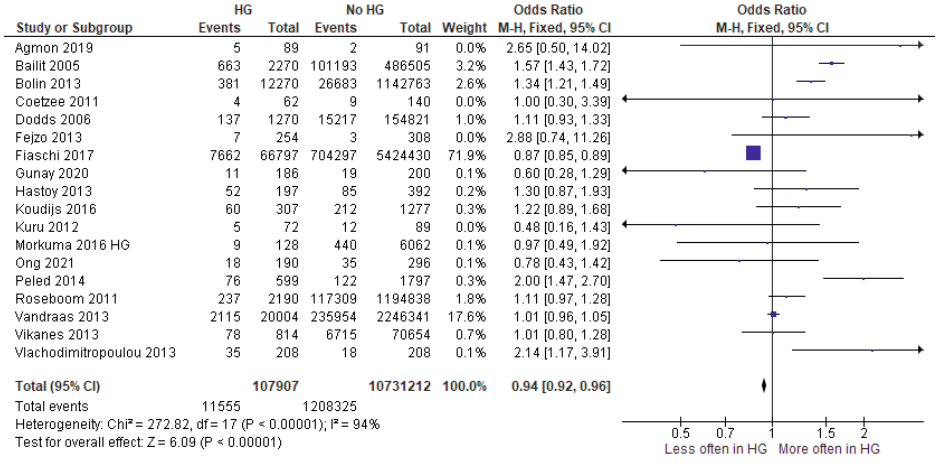
Heterogeneity: Chi² = 1.84, df = 3 (P = 0.61); I² = 0%
Test for overall effect: Z = 2.81 (P = 0.005)

Odds Ratio M-H, Fixed, 95% CI
0.2 0.5 1 2 5
Less often in HG More often in HG

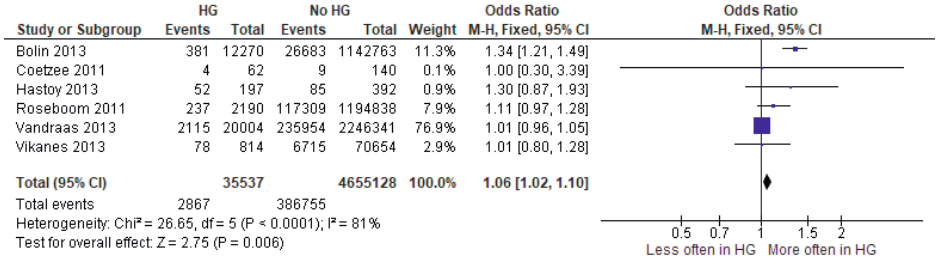
F.5 Sensitivity analysis birth weight <2500 grams only European studies



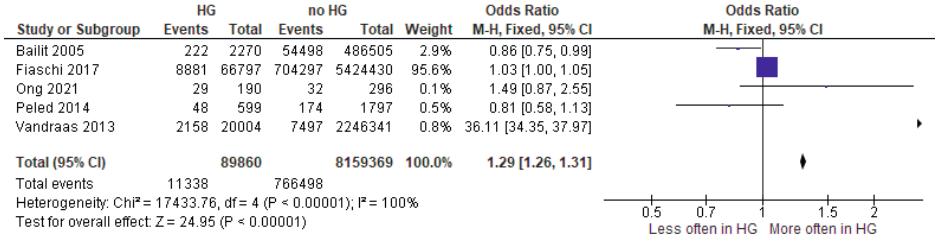
F.6 Sensitivity analysis small for gestational age only cohort studies



F.7 Sensitivity analysis small for gestational age only European studies



F.8 Sensitivity analysis large for gestational age only cohort studies



Appendix G. Overview of results of included studies reporting on congenital anomalies

Study	Definition	HG exposed/total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
Bashiri	All congenital malformations	164/373	1 (0.6%)	2 (1%)	not significant
Boissière-O'Neil	Birth defects documented at delivery	Total: 17,512/2,115,581	868 (5.0%)	116,992 (5.6%)	significant
		HG with metabolic disturbances: 2,366/ 2,100,435	120 (5.1%)		not significant
		HG without hospitalization: 1,561/2,099,630	71 (4.5%)		significant
		HG with 1 hospitalization: 12,780/ 2,110,849	633 (5.0%)		significant
		HG with 2 hospitalization: 2,127/ 2,098,069	106 (5.0%)		not significant
		HG with ≥3 hospitalizations: 1,044/ 2,099,113	58 (5.6%)		not significant
Buyukkayaci	Not described	50/100	0 (0%)	0 (0%)	not significant
Coetzee	Not described	61/198	3 (4.9%)	15 (10.9%)	not significant
Coffey	Anencephaly	48/296	31 (64.6%)	73 (29.4%)	not shown
Czeizel 2003	Cleft lip with or without cleft palate	217/2,742	89 (41%)	1,279 (50.7%)	not shown
Czeizel 2005	All congenital abnormalities reported in the first 3 months after birth	125/30,281	33 (26.4%)	11,973 (39.7%)	not shown
Depue	Central nervous system malformations	413/1,246	6 (1.5%)	3 (0.4%)	**0.03**
Fejzo 2013	Birth defects	254/562	7 (2.8%)	5 (1.6%)	not significant
Hallak *	Not described	Total: 138/12,473	3 (2.2%)	192 (1.6%)	not shown
		Mild HG: 40/12,375	1 (2.5%)		not significant
		Severe HG: 98/12,433	2 (2.0%)		not significant
Hohlbein	Not described	120/28,120	10 (8.3%)	274 (1.0%)	not shown
Tsang	Not described	207/13,614	11 (5.3%)	621 (4.6%)	not significant
Zhang 2017	Anencephaly, spina bifida and encephalocele	Total: 419/2,103	258 (61.6%)	574 (34.1%)	not shown
		Anencephaly: 419/2,103	116 (27.7%)	221 (13.1%)	not shown
		Spina bifida: 419/2,103	110 (26.3%)	278 (16.5%)	not shown
		Encephalocele: 419/ 2,103	32 (7.6%)	75 (4.5%)	not shown

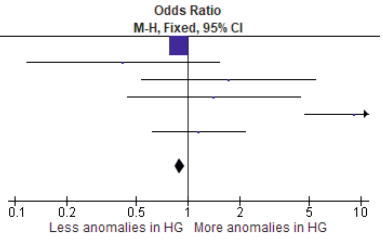
*Severe HG: if one or more of the following conditions was present: heavy ketonuria (>3+), abnormal serum BUN, sodium or potassium or increase in hematocrit (> 0.43)

Appendix H. Sensitivity analysis congenital anomalies

H.1 Sensitivity analysis congenital anomalies only cohort studies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Boissière-O'Neil 2021	868	17512	116992	2098069	98.0%	0.88 [0.82, 0.95]
Coetzee 2011	3	61	15	137	0.5%	0.42 [0.12, 1.51]
Fejzo 2013	7	254	5	308	0.2%	1.72 [0.54, 5.48]
Hallak 1993	3	138	192	12335	0.2%	1.41 [0.44, 4.45]
Hohlbein 1961	10	120	274	28000	0.1%	9.20 [4.76, 17.77]
Tsang 1996	11	207	621	13407	1.0%	1.16 [0.63, 2.13]
Total (95% CI)		**18292**		**2152256**	**100.0%**	**0.90 [0.84, 0.96]**
Total events	902		118099			

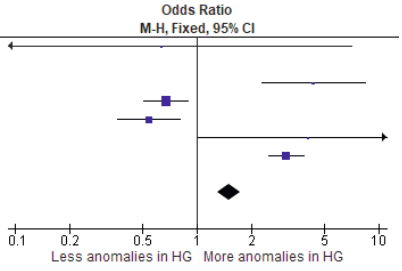
Heterogeneity: Chi² = 52.06, df = 5 (P < 0.00001); I² = 90%
Test for overall effect: Z = 3.19 (P = 0.001)



H.2 Sensitivity analysis congenital anomalies only case-control studies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bashiri 1995	1	164	2	209	0.6%	0.63 [0.06, 7.06]
Buyukkayaci 2015	0	50	0	50		Not estimable
Coffey 1957	31	48	73	248	2.9%	4.37 [2.28, 8.39]
Czeizel 2003	89	217	1279	2525	40.9%	0.68 [0.51, 0.90]
Czeizel 2005	33	125	11973	30156	24.9%	0.54 [0.37, 0.81]
Depue 1987	6	413	3	833	0.7%	4.08 [1.01, 16.39]
Zhang 2017	258	419	574	1684	30.1%	3.10 [2.48, 3.87]
Total (95% CI)		**1436**		**35705**	**100.0%**	**1.50 [1.30, 1.74]**
Total events	418		13904			

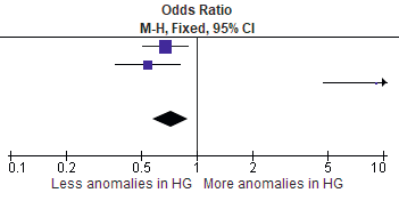
Heterogeneity: Chi² = 109.64, df = 5 (P < 0.00001); I² = 95%
Test for overall effect: Z = 5.49 (P < 0.00001)



H.3 Sensitivity analysis congenital anomalies only European studies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Czeizel 2003	89	217	1279	2525	61.5%	0.68 [0.51, 0.90]
Czeizel 2005	33	125	11973	30156	37.4%	0.54 [0.37, 0.81]
Hohlbein 1961	10	120	274	28000	1.1%	9.20 [4.76, 17.77]
Total (95% CI)		**462**		**60681**	**100.0%**	**0.72 [0.58, 0.90]**
Total events	132		13526			

Heterogeneity: Chi² = 59.55, df = 2 (P < 0.00001); I² = 97%
Test for overall effect: Z = 2.91 (P = 0.004)



Appendix I. Overview of results of included studies reporting on fetal sex

Study	HG exposed/total sample size	HG-exposed Female	Non-exposed Female	P
		Frequencies (%)	Frequencies (%)	
Askling	Total: 8,186/ 1,035,399	4,426 (54%)	499,496 (48.6%)	not shown
	1st trimester: 5,926/1,033,139	3,299 (55.7%)		**<0.000001**
	2nd and 3rd trimester: 2,260/1,029473	1,127 (49.9%)		>0.21
Bashiri	164/373	83 (50.6%)	105 (50.2%)	not significant
Basso	Total:6,227/83,946	3,337 (53.6%)	38,138 (49.1%)	not shown
	Only singleton: 5,941/81,831	3,179 (53.5%)	37,216 (49.0%)	not shown
	Only twins: 286/2,114	158 (55.2%)	921 (50.4%)	not shown
Bayraktar	95/947	49 (51.6%)	416 (48.8%)	Not significant
Caltekin	52/112	31 (59.6%)	29 (48.3%)	0.315
Coetzee	70/210	29 (41.4%)	61 (43.6%)	not significant
DelMar	Total 4,126/131,773	2,153 (52.2%)	62,234 (48.8%)	not shown
	1st trimester: 3,249	1,739 (53.5%)		not shown
	2nd and 3rd trimester: 877/ 128,524	414 (47.2%)		not shown
	Not hospitalized 1st trimester: 2,642/130,289	1,413 (53.5%)		not shown
	Hospitalized 1st trimester: 607/128,254	326 (53.7%)		not shown
Dodds	1,270/156,088	692 (54.5%)	75,506 (48.8%)	not shown
Fejzo 2013	254/562	147(57.9%)	145 (47.2%)	0.0138
Fejzo 2015	312/481	171 (54.8%)	87 (51.5%)	0.71
Fiaschi	Total: 110,956/7,683,176	60,725 (54.7%)	3.814.614 (49.6%)	not shown
	1 admission: 78,601/7,761,777	42,613 (54.2%)		not shown
	Readmission: 32,355/7,715,531	18,112 (56.0%)		not shown
Getahun	14,526/469,789	7,592 (52.3%)	221,856 (48.7%)	**<0.001133**
Gu	232/779	104 (44.8%)	251 (45.9%)	not shown
Gunay	186/386	99 (53.2%)	97 (48.5%)	0.361
Hastoy	197/589	121 (61.4%)	194 (49.5%)	**0.01**
Hsu	66/20,864	44 (66.7%)	10,145 (48.8%)	**<0.01**
Koot	62/8,953	33(53.2%)	4,303 (48.4%)	0.44
Koudijs *	Total:323/1,738	155 (48.0%)	651 (46%)	not shown
	Mild HG: 286/1,701	133 (46.5%)		not shown
	Severe HG: 37/1,452	22 (59.5%)		not shown
Kruse	267/10,255	139 (52.1%)	4,916 (49.2%)	not shown
Kuru	72/161	44 (61.1%)	45 (50.6%)	0.16
Mitsuda 2019	Total: 10,299/26,261	5,590 (54.3%)	7,211 (45.2%)	**<0.01**
	Only singleton: 10,019/25,733	5,422 (54.1%)	7,104 (45.2%)	not shown
	Only twins:280/528	168 (60%)	107 (43.1%)	not shown

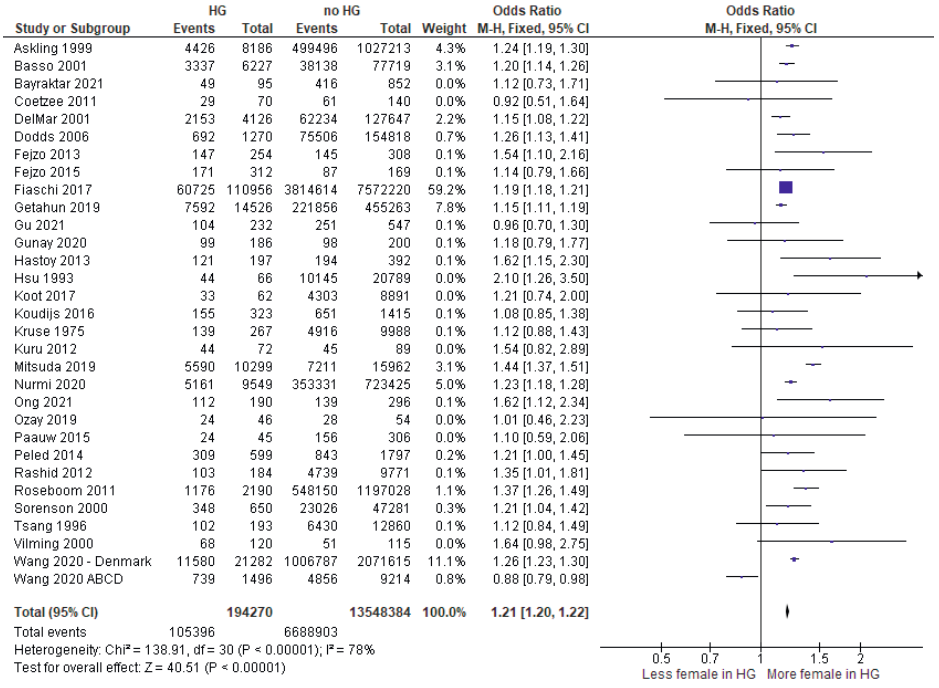
Continued

Study	HG exposed/total sample size	HG-exposed Female Frequencies (%)	Non-exposed Female Frequencies (%)	P
Nurmi	Total: 9,549/732,974	5,161 (54%)	353,331 (48.8%)	**significant**
	Only singleton: 9,244/722,381	4,978 (53.9%)	348,249 (48.8%)	**significant**
Ong	Total: 190/486	112 (58.9%)	139 (58.2%)	not shown
	Severe vomiting with hospitalization: 67/363	39 (58.2%)		not shown
	Severe vomiting: 123/419	73 (59.3%)		not shown
Ozay	46/100	24 (52.2%)	28 (51.9%)	0.18
Paauw	45/351	24 (53.3%)	156 (51%)	not significant
Peled 2014	Total: 599/2,396	309 (51.6%)	843 (46.9%)	**0.05**
	TPN: 122/1,919	100 (56.5%)		not shown
R68ashid	184/9,955	103 (56%)	4,739 (48.5%)	**<0.01**
Roseboom	2,190/ 1,199,218	1,176 (53.7%)	548,150 (48.8%)	**<0.0001**
Schiff **	Total: 3,261/13,044	1,817 (55.7%)	4,717 (48.2%)	not shown
	1st trimester: 2,110/11,893	1,214 (57.5%)		**significant**
	2nd trimester: 1,053/10,836	554 (52.6%)		**significant**
	3rd trimester: 98/9,881	49 (50%)		not significant
	Mild HG: 684/10,467	400 (58.5%)		**significant**
	Severe HG: 1,426/11,209	814 (57.1%)		**significant**
	1-2 admissions: 1,906/11,689	1,099 (57.7%)		**significant**
	3+ admissions: 204/9,987	115 (56.4%)		not significant
	1-2 days hospitalized: 1,271/11,054	693 (54.5%)		**significant**
	3+ days hospitalized: 839/10,622	521 (62.1%)		**significant**
Sorenson	650/47,931	348 (53.5%)	23,026 (48.7%)	**0.018**
Tan 2006	166/4,927	100 (60.2%)	2,311 (48.5%)	**0.004**
Tsang	193/13,053	102 (52.8%)	6,430 (50%)	not significant
Vlachodimitropoulou	208/416	103 (49.5%)	110 (52.9%)	not significant
Vilming	120/235	68 (56.6%)	51 (44.3%)	not shown
Wang	USA: 1,496/10,710	739 (49.4%)	4,856 (52.7%)	**0.018**
	Denmark: 21,282/2,092,897	11,580 (54.4%)	1,006,787 (48.6%)	not shown
Zhang 1991	201/1,867	109 (54.2%)	811 (48.7%)	not significant

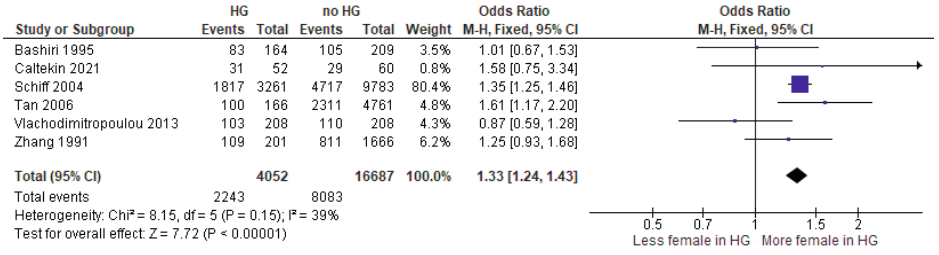
* Severe HG:>5% weight loss compared to pre pregnancy weight
** Severe HG: hospitalized ≥3 times or admitted with a metabolic disturbance or hospitalized for ≥3 days

Appendix J. Sensitivity analysis fetal sex

J.1 Sensitivity analysis fetal sex only cohort studies



J.2 Sensitivity analysis fetal sex only case-control studies

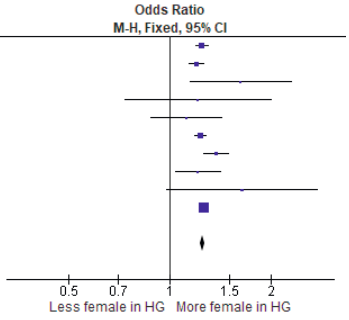


J.3 Sensitivity analysis fetal sex only European studies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Askling 1999	4426	8186	499496	1027213	17.1%	1.24 [1.19, 1.30]
Basso 2001	3337	6227	38138	77719	12.4%	1.20 [1.14, 1.26]
Hastoy 2013	121	197	194	392	0.2%	1.62 [1.15, 2.30]
Koot 2017	33	62	4303	8891	0.1%	1.21 [0.74, 2.00]
Kruse 1975	139	267	4916	9988	0.6%	1.12 [0.88, 1.43]
Nurmi 2020	5161	9549	353331	723425	19.9%	1.23 [1.18, 1.28]
Roseboom 2011	1176	2190	548150	1197028	4.4%	1.37 [1.26, 1.49]
Sorenson 2000	348	650	23026	47281	1.4%	1.21 [1.04, 1.42]
Vilming 2000	68	120	51	115	0.1%	1.64 [0.98, 2.75]
Wang 2020 - Denmark	11580	21282	1006787	2071615	43.9%	1.26 [1.23, 1.30]
Total (95% CI)		**48730**		**5163667**	**100.0%**	**1.25 [1.23, 1.27]**
Total events	26389		2478392			

Heterogeneity: $Chi^2 = 12.56$, df = 9 (P = 0.18); $I^2 = 28\%$

Test for overall effect: Z = 24.27 (P < 0.00001)



Appendix K. Overview of results of included studies reporting on gestational age at delivery

Study	HG exposed/total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P	Comment
Agmon*	Total: 89/180	38.83 ±2.26	39.11 ±2.61	0.44	
	Mild HG: 46/137	39.06 ±1.55		not shown	
	Severe HG: 43/134	38.59 ±2.82		not shown	
Bailit	Singleton: 2,270/488,505	39	39.2	**<0.0001**	No SD available
	Multiple: 163/13,891	35.6	35.8	0.32	
Bashiri	164/373	38.8	39.1	not significant	No SD available
Caltekin	52/112	38 (34-42)	39 (35-42)	**<0.001**	No SD available
Chin 1988**	Mild HG: 26/8,828	39.2 ±1.5	39.4 ±1.9	not shown	Total group not available
	Severe HG: 46/8,848	38.9 ±2.4		not shown	
Chin 1989	5/366	39.4 ±0.8	38.9 ±3.3	not shown	
Coetzee	All: 70/210	40.1 (35.7-41.9)	39.7 (31.6-43.1)	0.72	No SD available
	Spontaneous birth: 50/129	40.2 (35.7-41.6)	39.9 (31.6-41.7)	0.56	
	Born at term: 67/197	40.1 (37.1-41.9)	39.9 (37.3-42.4)	0.84	
Grooten	601/5,549	40 (39-40.9)	40.1 (39.1-41)	not shown	No SD available
Gunay	186/386	38.6 (37.6-40)	39 (37.5-40.05)	0.927	No SD available
Hallak***	Mild HG: 40/12,375	39.0 ±3.4	38.9 ±2.5	not shown	
	Severe HG: 98/12,433	39.0 ±2.4		not shown	
Kidess	65/130	39.9 ±0.31	40.2 ±0.2	not shown	
Koudijs****	Mild HG: 354/2,192	39.3 ±1.9	39.1 ±2.4	0.24	
	Severe HG: 46/1,884	39.4 ±2.7		0.67	
Kuru	72/161	38.1 ±2.3	38.1 ±2.6	0.91	
Mitsuda 2018	10,518/27,114	39.4 (37.0-41.3)	39.4 (36.9-41.3)	not shown	No SD available
Muraoka	34/103	39.0 ±1.1	39.0 ±1.0	0.13	
Ong	Severe vomiting with hospitalization: 67/363	38 (37-39)	39 (38-39)	**significant**	Total group not available
	Severe vomiting: 123/419	38 (37-39)		not significant	No SD available
Ozay	46/100	39.1 ±2.2	38.9 ±2.4	0.79	
Paauw	45/351	38.4 ±0.3	39.7 ±0.1	**<0.05**	

Continued

Study	**HG exposed/total sample size**	**HG-exposed** Mean ± SD; Median (IQR); Frequencies (%)	**Non-exposed** Mean ± SD; Median (IQR); Frequencies (%)	**P**	**Comment**
Peled 2014	Total: 599/2,396	38.3 ±2.6	39.1 ±2.0	**<0.001**	
	No TPN: 422/2,219	38.2 ±2.8		not shown	
	TPN: 122/1,919	38.7 ±2.1		not shown	
Salunkhe	76/1,035	38.1 ±2.1	37.8 ±2.4	not shown	
Tsang	193/13,053	38	38	not significant	No SD available
Vandraas	20,004/2,266,345	39.8	40	not significant	No SD available
Vikanes	Total: 814/71,468	39.7	40	significant	No SD available
	HG 1st trimester: 484/71,138	39.7		significant	
	HG 2nd trimester: 173/70,827	39.7		significant	
	HG 1st and 2nd trimester: 114/70,768	39.6		significant	
Vlachodimitropoulou	208/416	39 (38-40)	39 (38-40)	not significant	No SD available
Vilming	120/235	39 (20-42)	39.5 (28-43)	>0.05	No SD available

* Severe HG: ≥2 of the following: (1) ≥3 hospitalizations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria
** Severe HG: ≥1 of the following: heavy ketonuria (>3+). increase in urea and creatinine concentrations. serum electrolyte disturbance. increase in hematocrit (> 0.43)
*** Severe HG: ketonuria, increased BUN and hematocrit, and/or abnormal electrolytes
**** Severe HG:>5% weight loss compared to pre pregnancy weight

Appendix L. Overview of results of included studies reporting on preterm birth

Study	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
	<37 weeks			
Agmon *	Total: 89/180	8 (9%)	9 (9.9%)	0.836
	Mild HG: 46/137	4 (8.7%)		not shown
	Severe HG: 43/134	4 (9.3%		not shown
Chin 1989	5/366	0 (0%)	17 (4.8%)	not shown
Coetzee	All: 70/210	3 (4.3%	9 (6.4%)	not significant
	Spontaneous births: 50/129	2 (4%)	2 (2.5%)	not significant
Dodds	All: 1,270/156,091	82 (6.5%)	8,233 (5.4%)	not shown
	1-2 admissions: 1,182/155,903	75 (6.4%)		not shown
	3+ admissions: 88/154,909	7 (8%)		not shown
	Weight gain ≥ 7kg: 885/128,720	36 (4.1%)	6,134 (4.9%)	not significant
	Weight gain < 7kg: 144/127,979	20 (13.9%)		**significant**
Fejzo 2013	254/562	39 (15.4%)	13(4.2%)	**<0.0001**
Fejzo 2015	312/481	42 (13.5%)	12 (7.1%)	0.05019
Fiaschi	Total: 70,265/ 4,757,713	4,885 (6.9%)	265,847 (5.7%)	not shown
	1 admission: 49,580/4,737,028	3,357 (6.8%)		not shown
	Readmission: 20,685/4,708,133	1,528 (7.4%)		not shown
Getahun	14,526/469,789	1,223 (8.4%)	33,133 (7.3%)	**<0.001**
Gunay	186/386	15 (8.1%)	22 (11%)	0.388
Hallak **	Total: 138/12,473	18 (13.0%)	1,320 (10.7%)	not shown
	Mild HG: 40/12,375	7 (17.5%)		not significant
	Severe HG: 98/12,433	11 (11.2%)		not significant
Hastoy	Total:197/589	29 (14.7%)	44 (11.2%)	0.22
	Weight gain>7kg: 137/529	18 (13.1%)		not significant
	Weight gain <7kg: 60/452	11 (18.3%)		not significant
Kidess	65/130	7 (10.8%)	0 (0%)	not shown
Klebanoff	188/3,469	31 (16.5%)	397 (12.1%)	not significant
Koot	62/8,953	1 (1.6%)	436 (4.9%)	0.23
Kuru	72/161	8 (11.1%)	12 (13.5%)	0.65
Mitsuda 2018	10,495/27,042	454 (4.3%)	853 (5.2%)	not shown

Continued

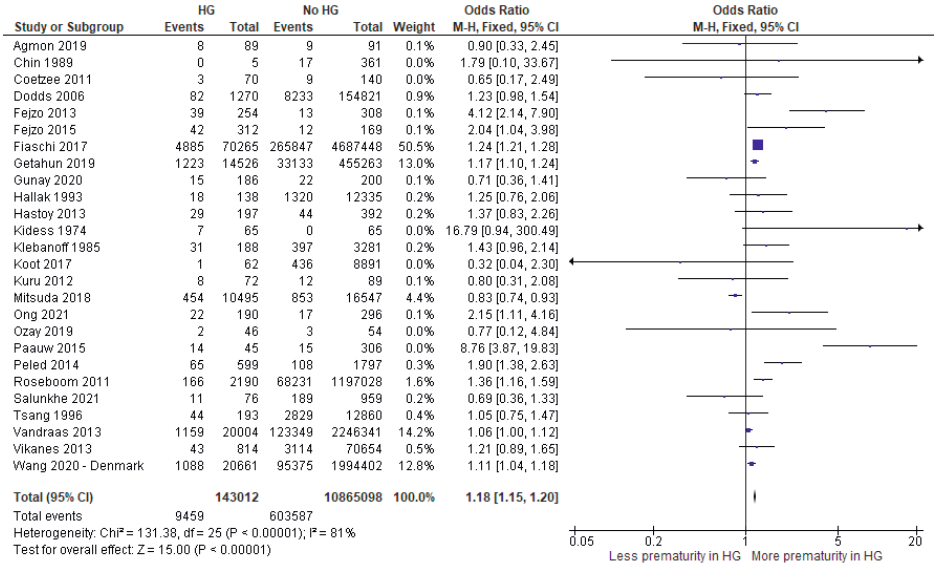
Study	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
Ong	Total: 190/486	22 (11.6%)	17 (5.7%)	not shown
	Severe vomiting with hospitalization: 67/363	7 (10.5%)		not significant
	Severe vomiting: 123/419	15 (12.2%)		<0.05
Ozay	46/100	2 (4.3%)	3 (5.6%)	0.33
Paauw	45/351	14 (31.1%)	15 (4.9%)	<0.05
Peled 2014	Total: 599/2,396	65 (10.9%)	108 (6.0%)	**<0.001**
	No TPN: 422/2,219	54 (12.8%)		not shown
	TPN: 122/1,919	11 (6.2%)		not shown
Roseboom	2,190/1,199,218	166 (7.6%)	68,231 (5.7%)	**<0.0001**
Salunkhe	76/1,035	11 (14.5%)	189 (19.7%)	not shown
Schiff	1,784/9,875	132 (7.4%)	445 (5.5%)	**0.002**
Tan 2007	166/664	8 (4.8%)	36 (7.2%)	0.37
Tsang	193/13,407	44 (23%)	2,829 (22%)	not significant
Vandraas	20,004/2,266,345	1,159 (5.8%)	123,349 (5.5%)	0.07
Vikanes	814/71,468	43 (5.3%)	3,114 (4.4%)	not significant
Vlachodimitropoulou	208/416	7 (3.4%)	9 (4.3%)	not significant
Wang	Denmark: 20,661/2,015,063	1,088 (5.3%)	95,375 (4.8%)	not shown
Zhang 1991	201/1,867	9 (4.5%)	72 (4.3%)	not significant
	<34 weeks			
Ong	Total: 190/486	1 (0.5%)	3 (1.0%)	not shown
	Severe vomiting with hospitalization: 67/363	1 (1.5%)		not shown
	Severe vomiting: 123/419	0 (0%)		not shown
Peled 2014	Total: 599/2,396	28 (4.7%)	29 (1.6%)	**<0.001**
	No TPN: 422/2,219	25 (5.9%)		not shown
	TPN: 122/1,919	3 (1.7%)		not shown
	<32 weeks			
Fiaschi	Total: 70,265/ 4,757,713	602 (0.9%)	30,285 (0.6%)	not shown
	1 admission: 49,580/4,737,028	431 (0.9%)		**significant**
	Readmission: 20,685/4,708,133	171 (0.8%)		not significant
Mitsuda 2018	10,495/27,042	31 (0.3%)	110 (0.7%)	not shown
Vandraas	20,004/2,266,345	148 (0.7%)	20,830 (0.9%)	**0.002**
	<28 weeks			
Mitsuda 2018	10,495/27,042	6 (0.1%)	16 (0.1%)	0.27

* Severe HG: ≥2 of the following: (1) ≥3 hospitalizations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria

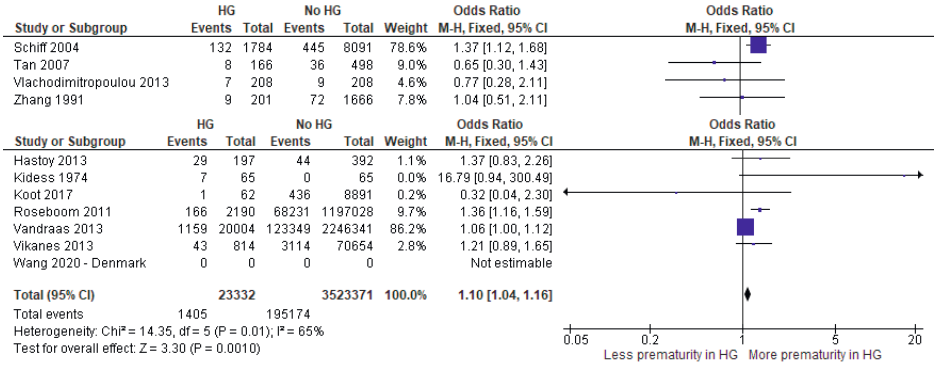
** Severe HG: ketonuria, increased BUN and hematocrit, and/or abnormal electrolytes

Appendix M. Sensitivity analysis preterm birth

M.1 Sensitivity analysis preterm birth <37 weeks only cohort studies



M.2 Sensitivity analysis preterm birth <37 weeks only case-control studies



Appendix N. Overview of results of included studies reporting on infant deaths

Study	Definition	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
		Fetal loss or fetal death			
Bailit	Fetal death after 20 weeks gestation	Total: 2,433/504,829	20 (0.8%)	2,294 (0.5%)	not shown
		Singleton: 2,270/488,505	16 (0.7%)	2,092 (0.4%)	0.05
		Multiple: 163/13,891	4 (2.5%)	202 (2.1%)	0.3
Buyukkayaci	Spontaneous abortion	50/100	1 (2%)	0 (0%)	not significant
Depue	Fetal loss ≤ 20 weeks of pregnancy	414/1,239	6 (1.4%)	25 (3%)	0.08
Klebanoff	Miscarriage between 14-19 weeks of gestation or stillbirth at 20 or more weeks of gestation	188/3,690	6 (3.3%)	184 (5.3%)	0.002
Koudijs*	Fetal loss ≤ 23 weeks of pregnancy or weighing up to 500 gram	Total: 400/2,233	30 (7.5%)	136 (7.4%)	not shown
		Mild HG: 354/2,187	28 (7.9%)		0.75
		Severe HG: 46/ 1,879	2 (4.3%)		0.44
Roseboom	Death during gestation or labor	2,190/1,199,218	4 (0.2%)	5,509 (0.5%)	0.06
Tsang	Not described	207/13,614	3 (1.4%)	241 (1.8%)	Not significant
Ustun	Spontaneous abortion	35/74	0 (0%)	0 (0%)	not shown
		Stillbirths			
Bolin	Fetal death occurring at 28 weeks of gestation or later	Total: 12,270/1,155,033	41 (0.3%)	3,628 (0.3%)	not significant
		HG 1st trimester: 10,186 /1,152,949	35 (0.3%)		not significant
		HG 2nd trimester: 2,084/1,144,847	6 (0.3%)		not significant
Chin 1989	Not described	5/366	0 (0%)	0 (0%)	not shown

Continued

Study	Definition	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
Coetzee	Fetal death occurring at 24 weeks of gestation or later	70/210	0 (0%)	1 (0.7%)	not shown
Depue	Fetal death >20 weeks gestation	414/1,239	6 (1.4%)	24 (2.9%)	0.08
Fiaschi	Not described	Total: 118,121/ 8,208,039	512 (0.4%)	39,443 (0.5%)	not shown
		1 admission: 83,632/8,173,550	355 (0.4%)		**significant**
		Readmission: 34,489/8,124,407	157(0.5%)		not significant
Gunay	Birth of an infant with no signs of life at or after 24 weeks of gestation	186/386	0 (0%)	1 (0.5%)	not shown
Koudijs*	Birth of a baby with no signs of life at or after 28 weeks of gestation	Total: 337/1,837	4 (1.2%)	12 (0.8%)	not shown
		Mild HG: 300/1,800	3 (1%)		0.74
		Severe HG: 37/ 1,537	1 (2.7%)		0.26
Kuru	Not described	72/161	0 (0%)	1 (1.1%)	0.36
Vandraas	Intrauterine death before birth, death during birth or at an unknown time	20,004/2,266,345	120 (0.6%)	13,119 (0.6%)	0.77
		Perinatal death			
Agmon**	Fetal death occurring during or before labor, or any neonatal death occurring within 1-month post labor	Total:89/180	2 (2.2%)	1 (1.1%)	0.22
		Mild HG: 46/137	0 (0%)		not shown
		Severe HG: 43/ 134	2 (4.7%)		not shown
Bashiri	Not described	164/373	3 (1.9%)	4 (1.9%)	not significant
Dodds	Not described	1,270/156,091	6 (0.5%)	939 (0.6%)	not significant

Continued

Study	Definition	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
Fejzo 2013	Fetal deaths after 20 weeks gestation to 1 week after birth	254/562	3 (1.2%)	2 (0.6%)	0.8283
Hastoy	Fetal deaths after 28 weeks gestation to 1 week after birth	197/589	3 (1.5%)	8 (2.0%)	0.66
Roseboom	Death during gestation and during labor	2,190/1,199,218	7 (0.3%)	7,182 (0.6%)	0.1
Tsang	Not described	207/13,614	10 (4.8%)	530 (3.9%)	not significant
Vandraas	Stillborn with gestational length of ≥ 22 weeks, and deaths before the end of day 7 after birth.	20,004/2,266,345	191 (1.0%)	20,165 (0.9%)	0.41
Vikanes	Death during the perinatal period (lasting from≥ 22nd gestational week until the 7th day after birth)	814/71,468	0 (0%)	1,167 (1.7%)	not shown
		Neonatal deaths			
Bailit	Not described	Total: 2,433/504,829	15 (0.6%)	1,751 (0.3%)	not shown
		Singleton: 2,270/488,505	10 (0.4%)	1,460 (0.3%)	0.07
		Multiple: 163/13,891	5 (3.1%)	291 (2.1%)	0.58
Chin 1989	Not described	5/366	0 (0%)	0 (0%)	not shown
Peled 2014	Not described	Total: 599/2,396	0 (0%)	2 (0.1%)	0.4
		No TPN: 422/2,219	0 (0%)		not shown
		TPN: 122/1,919	0 (0%)		not shown
Roseboom	Death within first week after birth	2,190/1,199,218	3 (0.1%)	1,549 (0.1%)	0.96
Vandraas	Deaths during the first 28 days of life	20,004/2,266,345	79 (0.4%)	8,559 (0.4%)	0.75

* Severe HG: >5% weight loss compared to pre pregnancy weight
** Severe HG: ≥2 of the following: (1) ≥3 hospitalizations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria

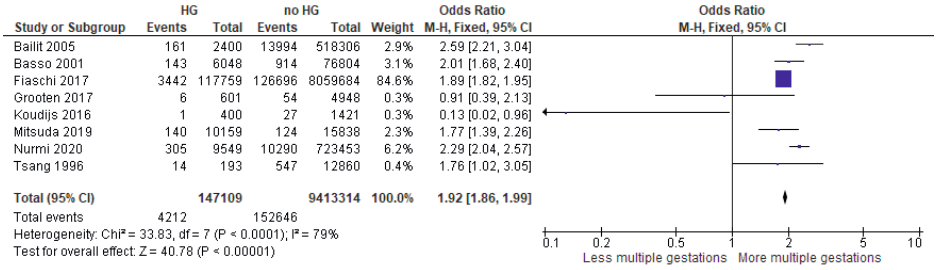
Appendix O. Overview of results of included studies reporting on multiple gestations

Study	**HG-exposed/ total sample size**	**HG-exposed** Frequencies (%)	**Non-exposed** Frequencies (%)	**P**
Bailit	2,400/520,706	161 (6.7%)	13,994 (2.7%)	<0.001
Bashiri	164/373	14 (8.5%)	9 (4.3%)	not shown
Basso	6,048/82,852	143 (2.4%)	914 (1.2%)	not shown
Fiaschi	Total: 117,759/8,177,443	3,442 (2.9%)	126,696 (1.6%)	not shown
	Total 1 admission: 83,366/8,143,050	2,279 (2.7%)		not shown
	Total readmission: 34,393/8,094,077	1,163 (3.4%)		not shown
	Total only twins: 117,759/8,177,443	3,286 (2.8%)	121,083 (1.5%)	not shown
	1 admission only twins: 83,366/8,143,050	2,184 (2.6%)		not shown
	Readmission only twins: 34,393/8,094,077	1,102 (3.2%)		not shown
	Total triplets or more: 117,759/8,177,443	156 (0.1%)	5,613 (0.1%)	not shown
	1 admission triplets or more: 83,366/8,143,050	95 (0.1%)		not shown
	Readmission triplets or more: 34,393/8,094,077	61 (0.2%)		not shown
Grooten	601/5,549	6 (10%)	54 (10.9%)	not shown
Koudijs *	Total: 400/1,821	1 (0.3%)	27 (1.9%)	not shown
	Mild HG: 354/1,775	1 (0.4%)		not shown
	Severe HG: 46/1,467	0 (0%)		not shown
Mitsuda 2019	10,159/25,997	140 (1.4%)	124 (0.8%)	not shown
Nurmi	9,549/733,002	305 (3.2%)	10,290 (1.4%)	not shown
Tsang	193/13,053	14 (7.3%)	547 (4.3%)	not shown

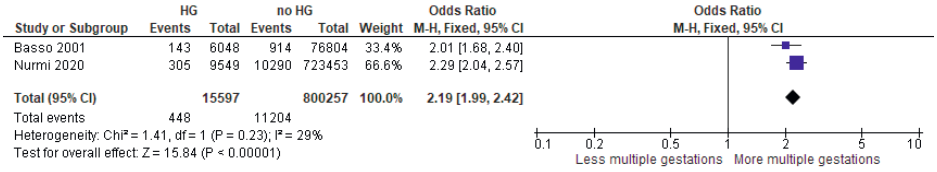
* Severe HG: >5% weight loss compared to pre pregnancy weight

Appendix P. Sensitivity analysis multiple gestations

P.1 Sensitivity analysis multiple gestations only cohort studies



P.2 Sensitivity analysis multiple gestations only European studies



Appendix Q. Overview of results of included studies reporting on postpartum neonatal complications

Study	**HG-exposed/ total sample size**	**HG-exposed** Frequencies (%)	**Non-exposed** Frequencies (%)	**P**
		Respiratory distress syndrome		
Peled 2014	Total: 599/2,396	16 (2.7%)	22 (1.2%)	**0.01**
	No TPN: 422/2,219	13 (3.1%)		not shown
	TPN: 122/1,919	3 (1.7%)		not shown
		Necrotizing enterocolitis		
Peled 2014	Total: 599/2,396	3 (0.5%)	9 (0.5%)	0.9
	No TPN: 422/2,219	3 (0.7%)		not shown
	TPN: 122/1,919	0 (0%)		not shown
		Jaundice requiring phototherapy		
Peled 2014	Total: 599/2,396	28 (4.7%)	59 3.3%)	0.1
	No TPN: 422/2,219	22 (5.2%)		not shown
	TPN: 122/1,919	6 (3.4%)		not shown
		Hypoglycemia		
Peled 2014	Total: 599/2,396	12 (2.0%)	23 (1.3%)	0.2
	No TPN: 422/2,219	7 (1.7%)		not shown
	TPN: 122/1,919	5 (2.8%)		not shown
		Resuscitation		
Coetzee	62/199	7 (11.3%)	9 (6.6%)	not sinificant
Fiaschi	Total: 61,578/4,289,145	7,751 (12.6%)	499,579 (11.8%)	not shown
	Total 1 admission: 43,461/4,271,028	5,507 (12.7%)		not shown
	Total readmission: 18,117/4,245,684	2,244 (12.4%)		not shown

Appendix R. Overview of results of included studies reporting on NICU admission and length of stay in hospital

Study	HG-exposed/ total sample size	HG-exposed Mean ± SD; Frequencies (%)	Non-exposed Mean ± SD; Frequencies (%)	P
		NICU admission		
Coetzee	57/194	3 (5.3%)	16 (11.7%)	not significant
Fiaschi	Total: 75,501/5,300,097	1,550 (2.1%)	90,491 (1.7%)	not shown
	1 admission HG: 53,547/5,278,143	1,105 (2.1%)		not shown
	Readmission HG: 21,954/5,246,550	445 (2.0%)		not shown
Hallak*	Total: 138/12,473	8 (5.8%)	974 (7.9%)	not shown
	Mild HG: 40/12,375	4 (10%)		not significant
	Severe HG: 98/12,433	4 (4.1%)		not significant
Ong	Total: 190/486	10 (5.3%)	14 (4.8%)	not shown
	Severe vomiting with hospitalization: 67/363	8 (11.9%)		not shown
	Severe vomiting with hospitalization+ term birth: 60/339	5 (8.3%)	8 (2.9%)	not shown
	Severe vomiting with hospitalization+ preterm birth: 7/24	3 (42.9%)	6 (35.3%)	not shown
	Severe vomiting: 123/419	2 (1.6%)	14 (4.8%)	not shown
	Severe vomiting+ term birth: 108/387	1 (0.9%)	8 (2.9%)	not shown
	Severe vomiting+ preterm birth: 15/32	1 (6.7%)	6 (35.3%)	not shown
Paauw	45/351	5 (11.4%)	18 (6.0%)	not significant
Peled 2014	Total: 599/2,396	43 (7.2%)	45 (2.5%)	**<0.001**
	No TPN: 422/2,219	37 (8.8%)		not shown
	TPN: 122/1,919	6 (3.4%)		not shown
Roseboom	2,186/1,193,705	26 (1.2%)	10,724 (0.9%)	0.2
		Length of stay in hospital		
Paauw	45/351	2.9 ±0.5	1.8 ±0.1	**<0.05**

* Severe HG: ketonuria, increased BUN and hematocrit, and/or abnormal electrolytes

Appendix S. Overview of results of included studies reporting on newborn measurements

Study	HG-exposed/ total sample size	HG-exposed Mean ± SD; Median (IQR); Frequencies (%)	Non-exposed Mean ± SD; Median (IQR); Frequencies (%)	P
Crown-heel length (cm)				
Coetzee	All: 61/198	51 (45-56)	52 (37-58)	not significant
	Born at term: 67/197	52 (45-56)	52 (45.5-58)	not significant
Kidess	56/130	49.3 ±0.8	49.9 ±0.8	not shown
Head circumference (cm)				
Coetzee	All: 61/198	34.5 (30.2-38)	35.5 (26.5-39)	not significant
	Born at term: 58/185	34.8 (30.2-38)	35.5 (31-39)	**0.02**
Kidess	56/130	34.4 ±0.2	34.4 ±0.2	not shown
Biparietale diameter (cm)				
Kidess	65/130	9.3 ±0.1	9.2 ±0.1	not shown
Head circumference is small for gestational age				
Coetzee	All: 61/198	10 (16.4 %)	11 (8%)	not significant
	Born at term: 58/185	10 (17.2%)	10 (7.9%)	not significant
Head circumference is large for gestational age				
Coetzee	All: 61/198	12 (19.7%)	31 (22.7%)	not significant
	Born at term: 58/185	11 (19%)	31 (24.4%)	not significant

Appendix T. Overview of results of included studies reporting on placental abruption

Study	HG-exposed/ total sample size	HG-exposed Frequencies (%)	Non-exposed Frequencies (%)	P
Agmon *	Total: 89/180	1 (1.1%)	0 (0%)	0.494
	Mild HG: 46/137	0 (0%)		not shown
	Severe HG: 43/ 134	1 (2.3%)		not shown
Bolin	Total: 12,270/1,155,033	64 (0.5%)	4,652 (0.4%)	significant
	HG 1st trimester: 10,186 /1,152,949	42 (0.4%)		not significant
	HG 2nd trimester: 2084/1,144,847	22 (1.1%)		significant
Chin 1989	5/366	0 (0%)	1 (0.6%)	not shown
Fiaschi	Total: 118,197/ 8,211,850	486 (0.4%)	29,562 (0.4%)	not shown
	1 admission: 83,679/8,177,332	340 (0.4%)		not significant
	Readmission: 34,518/8,128,171	146 (0.4%)		not significant
Gunay	186/386	2 (1.1%)	1 (0.5%)	not shown
Peled 2014	Total: 599/2,396	11 (1.8%)	20 (1.1%)	0.2
	No TPN: 422/2,219	9 (2.1%)		not shown
	TPN: 122/1,919	2 (1.1%)		not shown

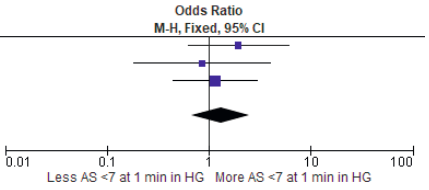
* Severe HG: ≥2 of the following: (1) ≥3 hospitalisations in the first half of pregnancy, (2) elevated liver enzymes, (3) Abnormal levels of sodium or potassium, (4) weight gain < 7 kg or (5) ketonuria

Appendix U. Subgroup analysis studies performed after 2001

U.1 Subgroup analysis studies performed after 2001 – Apgar <7 at 1 minute

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Agmon 2019	9	89	5	91	28.5%	1.94 [0.62, 6.02]
Caltekin 2021	3	52	4	60	22.5%	0.86 [0.18, 4.02]
Coetzee 2011	7	62	14	140	49.0%	1.15 [0.44, 2.99]
Total (95% CI)		**203**		**291**	**100.0%**	**1.31 [0.68, 2.51]**
Total events	19		23			

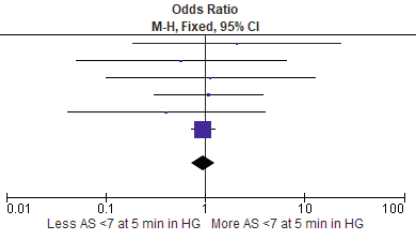
Heterogeneity: Chi² = 0.82, df = 2 (P = 0.66); I² = 0%
Test for overall effect: Z = 0.80 (P = 0.42)



U.2 Subgroup analysis studies performed after 2001 – Apgar score <7 at 5 minutes

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Agmon 2019	2	89	1	91	0.9%	2.07 [0.18, 23.23]
Caltekin 2021	1	52	2	60	1.6%	0.57 [0.05, 6.46]
Coetzee 2011	1	62	2	140	1.1%	1.13 [0.10, 12.71]
Gunay 2020	5	186	5	200	4.1%	1.08 [0.31, 3.78]
Kuru 2012	1	72	3	89	2.3%	0.40 [0.04, 3.97]
Roseboom 2011	50	2186	28596	1191519	90.0%	0.95 [0.72, 1.26]
Total (95% CI)		**2647**		**1192099**	**100.0%**	**0.95 [0.73, 1.24]**
Total events	60		28609			

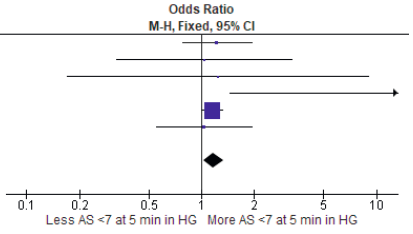
Heterogeneity: Chi² = 1.17, df = 5 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.38 (P = 0.70)



U.3 Subgroup analysis studies performed before 2001 – Apgar score <7 at 5 minutes

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Dodds 2006	19	1270	1898	154821	7.7%	1.22 [0.78, 1.93]
Hallak 1993	3	138	259	12335	1.4%	1.04 [0.33, 3.27]
Koot 2017	1	62	116	8891	0.4%	1.24 [0.17, 9.02]
Peled 2014	4	599	0	1797	0.1%	27.17 [1.46, 505.32]
Vandraas 2013	199	20004	19421	2246341	85.7%	1.15 [1.00, 1.33]
Vikanes 2013	10	814	838	70654	4.8%	1.04 [0.55, 1.94]
Total (95% CI)		**22887**		**2494839**	**100.0%**	**1.17 [1.03, 1.33]**
Total events	236		22532			

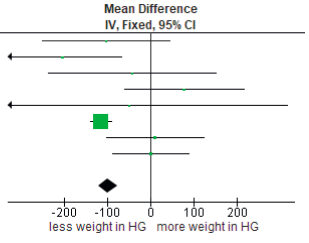
Heterogeneity: Chi² = 4.71, df = 5 (P = 0.45); I² = 0%
Test for overall effect: Z = 2.35 (P = 0.02)



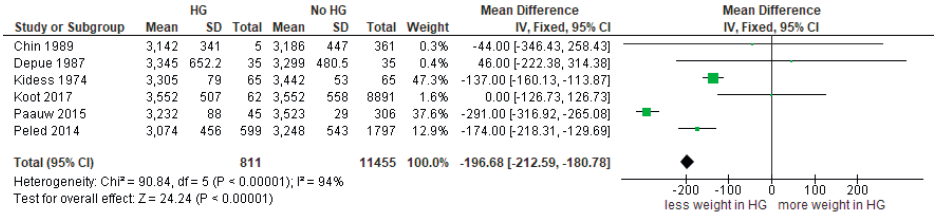
U.4 Subgroup analysis studies performed after 2001 – Birth weight

Study or Subgroup	HG Mean	HG SD	HG Total	No HG Mean	No HG SD	No HG Total	Weight	Mean Difference IV, Fixed, 95% CI
Agmon 2019	3,076.06	449.97	89	3,178.9	559.32	91	2.2%	-102.84 [-250.98, 45.30]
Caltekin 2021	3,060	321	52	3,265	414	60	2.6%	-205.00 [-341.33, -68.67]
Kuru 2012	3,121	595.4	72	3,164	664.5	89	1.3%	-43.00 [-237.87, 151.87]
Muraoka 2020	3,148.9	348	34	3,070.9	316	70	2.6%	78.00 [-60.43, 216.43]
Ozay 2019	2,850	980	46	2,900	870	54	0.4%	-50.00 [-416.12, 316.12]
Roseboom 2011	3,337	587	2190	3,453	579	1194838	80.7%	-116.00 [-140.61, -91.39]
Salunkhe 2021	2,658.8	484	76	2,648.8	455.9	959	3.9%	10.00 [-102.58, 122.58]
Tan 2007	3,000	500	166	3,000	500	498	6.3%	0.00 [-87.83, 87.83]
Total (95% CI)			**2725**			**1196659**	**100.0%**	**-99.71 [-121.82, -77.59]**

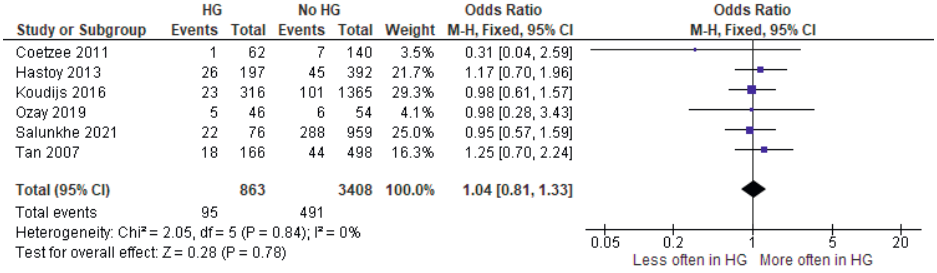
Heterogeneity: Chi² = 19.30, df = 7 (P = 0.007); I² = 64%
Test for overall effect: Z = 8.84 (P < 0.00001)



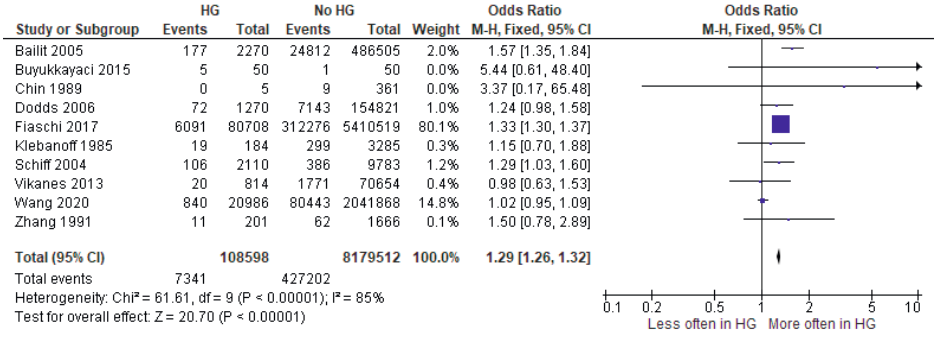
U.5 Subgroup analysis studies performed before 2001 – Birth weight



U.6 Subgroup analysis studies performed after 2001 – Birth weight <2500 grams



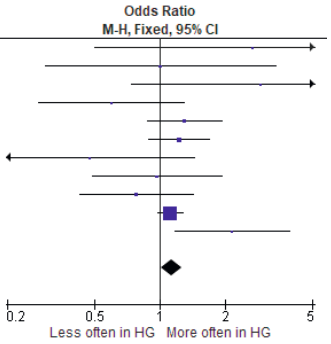
U.7 Subgroup analysis studies performed before 2001 – Birth weight <2500 grams



U.8 Subgroup analysis studies performed after 2001 – Small for gestational age

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Agmon 2019	5	89	2	91	0.3%	2.65 [0.50, 14.02]
Coetzee 2011	4	62	9	140	0.9%	1.00 [0.30, 3.39]
Fejzo 2013	7	254	3	308	0.5%	2.88 [0.74, 11.26]
Gunay 2020	11	186	19	200	2.9%	0.60 [0.28, 1.29]
Hastoy 2013	52	197	85	392	7.2%	1.30 [0.87, 1.93]
Koudijs 2016	60	307	212	1277	11.3%	1.22 [0.89, 1.68]
Kuru 2012	5	72	12	89	1.7%	0.48 [0.16, 1.43]
Morkuma 2016 HG	9	128	440	6062	2.9%	0.97 [0.49, 1.92]
Ong 2021	18	190	35	296	4.2%	0.78 [0.43, 1.42]
Roseboom 2011	237	2190	117309	1194838	65.5%	1.11 [0.97, 1.28]
Vlachodimitropoulou 2013	35	208	18	208	2.6%	2.14 [1.17, 3.91]
Total (95% CI)		**3883**		**1203901**	**100.0%**	**1.13 [1.02, 1.26]**
Total events	443		118144			

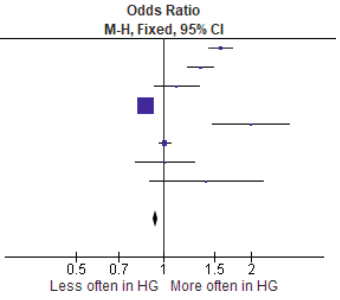
Heterogeneity: Chi² = 14.46, df = 10 (P = 0.15); I² = 31%
Test for overall effect: Z = 2.24 (P = 0.02)



U.9 Subgroup analysis studies performed before 2001 – Small for gestational age

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bailit 2005	663	2270	101193	486505	3.2%	1.57 [1.43, 1.72]
Bolin 2013	381	12270	26683	1142763	2.7%	1.34 [1.21, 1.49]
Dodds 2006	137	1270	15217	154821	1.1%	1.11 [0.93, 1.33]
Fiaschi 2017	7662	66797	704297	5424430	73.8%	0.87 [0.85, 0.89]
Peled 2014	76	599	122	1797	0.3%	2.00 [1.47, 2.70]
Vandraas 2013	2115	20004	235954	2246341	18.1%	1.01 [0.96, 1.05]
Vikanes 2013	78	814	6715	70654	0.7%	1.01 [0.80, 1.28]
Zhang 1991	25	201	153	1666	0.1%	1.40 [0.90, 2.20]
Total (95% CI)		**104225**		**9528977**	**100.0%**	**0.94 [0.92, 0.95]**
Total events	11137		1090334			

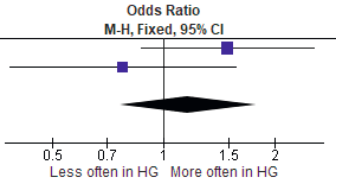
Heterogeneity: Chi² = 250.92, df = 7 (P < 0.00001); I² = 97%
Test for overall effect: Z = 6.50 (P < 0.00001)



U.10 Subgroup analysis studies performed after 2001 – Large for gestational age

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Ong 2021	29	190	32	296	54.6%	1.49 [0.87, 2.55]
Vlachodimitropoulou 2013	15	208	19	208	45.4%	0.77 [0.38, 1.57]
Total (95% CI)		**398**		**504**	**100.0%**	**1.16 [0.76, 1.78]**
Total events	44		51			

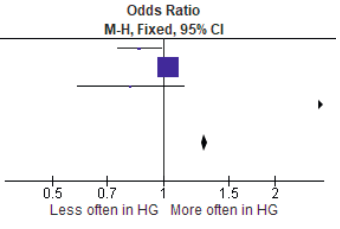
Heterogeneity: Chi² = 2.08, df = 1 (P = 0.15); I² = 52%
Test for overall effect: Z = 0.69 (P = 0.49)



U.11 Subgroup analysis studies performed before 2001 – Large for gestational age

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bailit 2005	222	2270	54498	486505	2.9%	0.86 [0.75, 0.99]
Fiaschi 2017	8881	66797	704297	5424430	95.8%	1.03 [1.00, 1.05]
Peled 2014	48	599	174	1797	0.5%	0.81 [0.58, 1.13]
Vandraas 2013	2158	20004	7497	2246341	0.8%	36.11 [34.35, 37.97]
Total (95% CI)		**89670**		**8159073**	**100.0%**	**1.29 [1.26, 1.31]**
Total events	11309		766466			

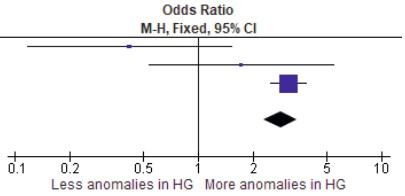
Heterogeneity: Chi² = 17434.87, df = 3 (P < 0.00001); I² = 100%
Test for overall effect: Z = 24.91 (P < 0.00001)



U.12 Subgroup analysis studies performed after 2001 – Congenital anomalies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Coetzee 2011	3	61	15	137	8.7%	0.42 [0.12, 1.51]
Fejzo 2013	7	254	5	308	4.3%	1.72 [0.54, 5.48]
Zhang 2017	258	419	574	1684	87.0%	3.10 [2.48, 3.87]
Total (95% CI)		**734**		**2129**	**100.0%**	**2.81 [2.27, 3.46]**
Total events	268		594			

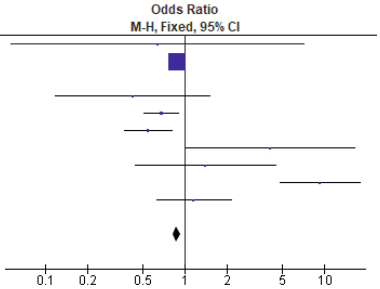
Heterogeneity: Chi² = 9.93, df = 2 (P = 0.007); I² = 80%
Test for overall effect: Z = 9.60 (P < 0.00001)



U.13 Subgroup analysis studies performed before 2001 – Congenital anomalies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bashiri 1995	1	164	2	209	0.1%	0.63 [0.06, 7.06]
Boissière-O'Neil 2021	868	17512	116992	2098069	88.9%	0.88 [0.82, 0.95]
Buyukkayaci 2015	0	50	0	50		Not estimable
Coetzee 2011	3	61	15	137	0.4%	0.42 [0.12, 1.51]
Czeizel 2003	89	217	1279	2525	5.8%	0.68 [0.51, 0.90]
Czeizel 2005	33	125	11973	30156	3.5%	0.54 [0.37, 0.81]
Depue 1987	6	413	3	833	0.1%	4.08 [1.01, 16.39]
Hallak 1993	3	138	192	12335	0.2%	1.41 [0.44, 4.45]
Hohlbein 1961	10	120	274	28000	0.1%	9.20 [4.76, 17.77]
Tsang 1996	11	207	621	13407	0.9%	1.16 [0.63, 2.13]
Total (95% CI)		**19007**		**2185721**	**100.0%**	**0.87 [0.82, 0.93]**
Total events	1024		131351			

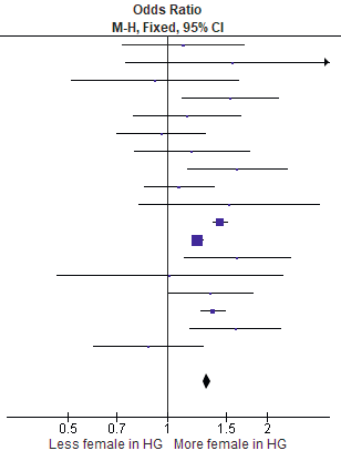
Heterogeneity: Chi² = 65.31, df = 8 (P < 0.00001); I² = 88%
Test for overall effect: Z = 4.13 (P < 0.0001)



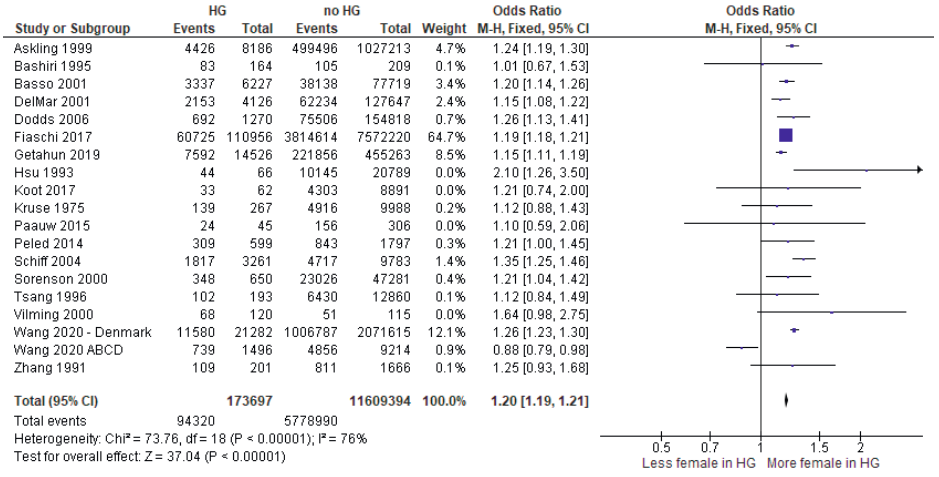
U.14 Subgroup analysis studies performed after 2001 – Fetal sex

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bayraktar 2021	49	95	416	852	0.5%	1.12 [0.73, 1.71]
Caltekin 2021	31	52	29	60	0.1%	1.58 [0.75, 3.34]
Coetzee 2011	29	70	61	140	0.3%	0.92 [0.51, 1.64]
Fejzo 2013	147	254	145	308	0.6%	1.54 [1.10, 2.16]
Fejzo 2015	171	312	87	169	0.6%	1.14 [0.79, 1.66]
Gu 2021	104	232	251	547	1.0%	0.96 [0.70, 1.30]
Gunay 2020	99	186	98	200	0.5%	1.18 [0.79, 1.77]
Hastoy 2013	121	197	194	392	0.6%	1.62 [1.15, 2.30]
Koudijs 2016	155	323	651	1415	1.5%	1.08 [0.85, 1.38]
Kuru 2012	44	72	45	89	0.2%	1.54 [0.82, 2.89]
Mitsuda 2019	5590	10299	7211	15962	30.4%	1.44 [1.37, 1.51]
Nurmi 2020	5161	9549	353331	723425	49.8%	1.23 [1.18, 1.28]
Ong 2021	112	190	139	296	0.5%	1.62 [1.12, 2.34]
Ozay 2019	24	46	28	54	0.1%	1.01 [0.46, 2.23]
Rashid 2012	103	184	4739	9771	0.9%	1.35 [1.01, 1.81]
Roseboom 2011	1176	2190	548150	1197028	10.9%	1.37 [1.26, 1.49]
Tan 2006	100	166	2311	4761	0.7%	1.61 [1.17, 2.20]
Vlachodimitropoulou 2013	103	208	110	208	0.7%	0.87 [0.59, 1.28]
Total (95% CI)		**24625**		**1955677**	**100.0%**	**1.31 [1.28, 1.35]**
Total events	13319		917996			

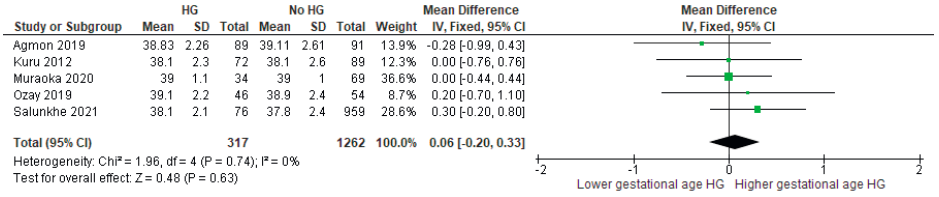
Heterogeneity: Chi² = 43.58, df = 17 (P = 0.0004); I² = 61%
Test for overall effect: Z = 18.94 (P < 0.00001)



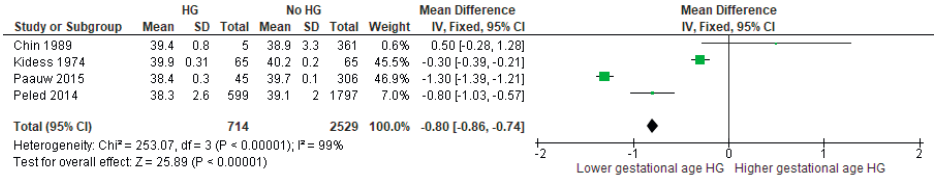
U.15 Subgroup analysis studies performed before 2001 – Fetal sex



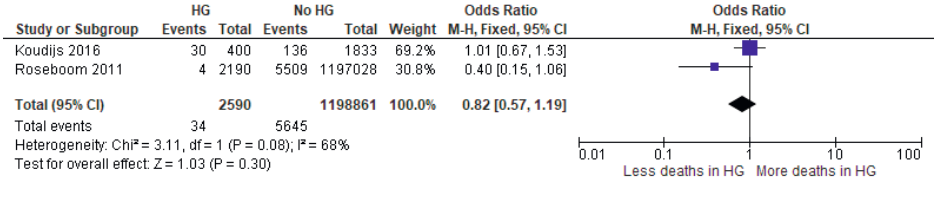
U.16 Subgroup analysis studies performed after 2001 – Gestational age at delivery



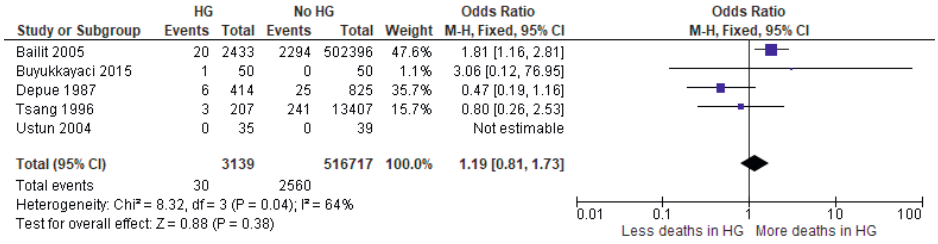
U.17 Subgroup analysis studies performed before 2001 – Gestational age at delivery



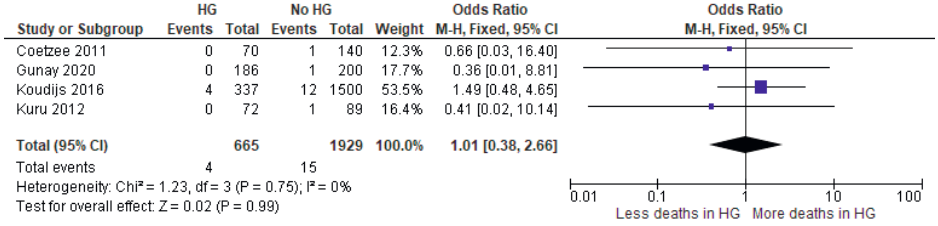
U.18 Subgroup analysis studies performed after 2001 – Fetal loss or fetal death



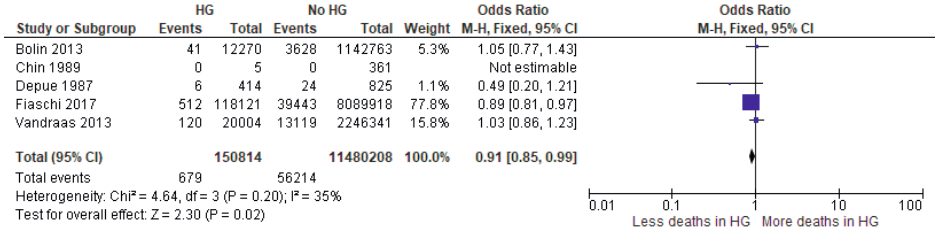
U.19 Subgroup analysis studies performed before 2001 – Fetal loss or fetal death



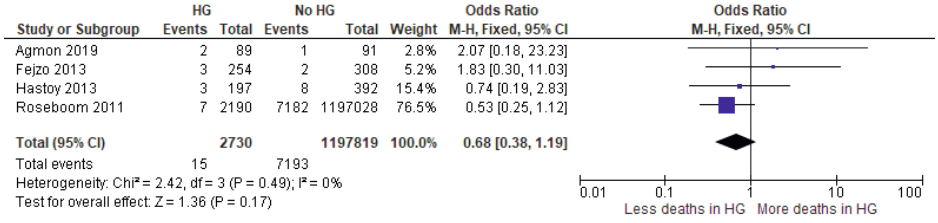
U.20 Subgroup analysis studies performed after 2001 – Stillbirth



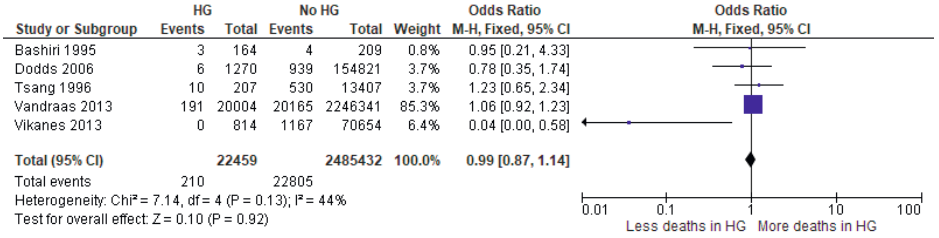
U.21 Subgroup analysis studies performed before 2001 – Stillbirth



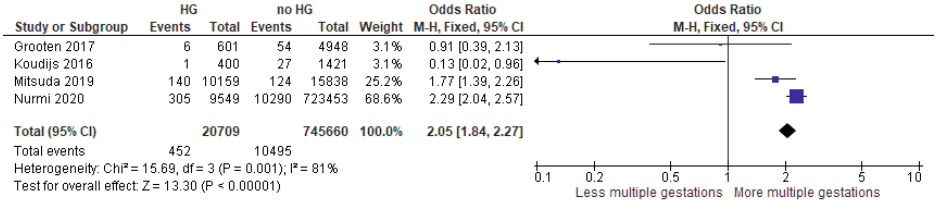
U.22 Subgroup analysis studies performed after 2001 – Perinatal death



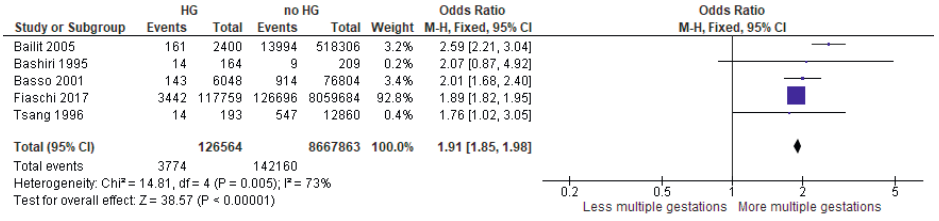
U.23 Subgroup analysis studies performed before 2001 – Perinatal death



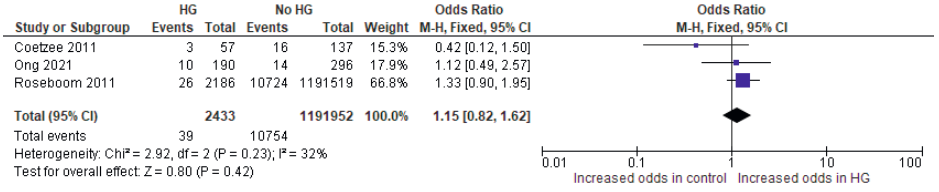
U.24 Subgroup analysis studies performed after 2001 – Multiple gestations



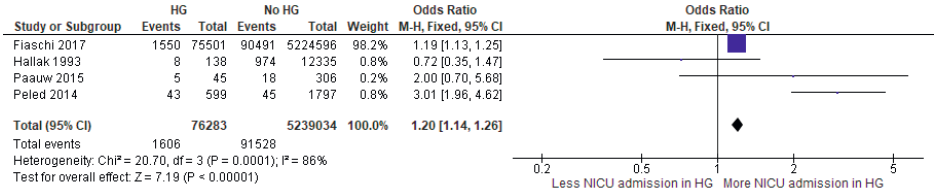
U.25 Subgroup analysis studies performed before 2001 – Multiple gestations



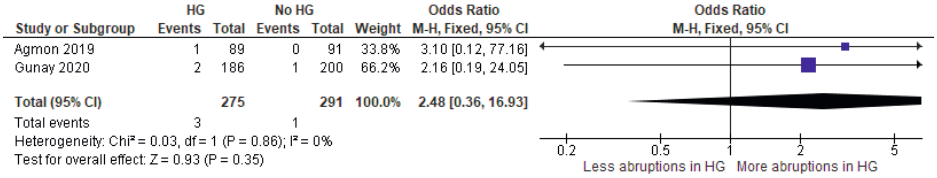
U.26 Subgroup analysis studies performed after 2001 – NICU admission



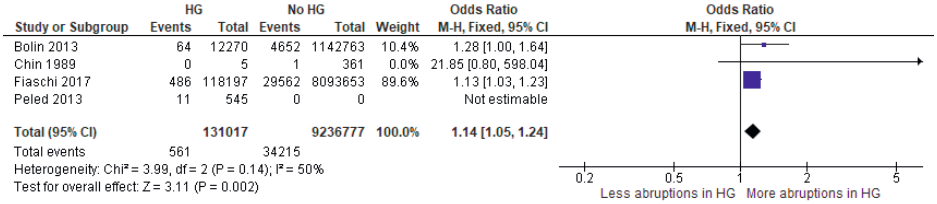
U.27 Subgroup analysis studies performed before 2001 – NICU admission



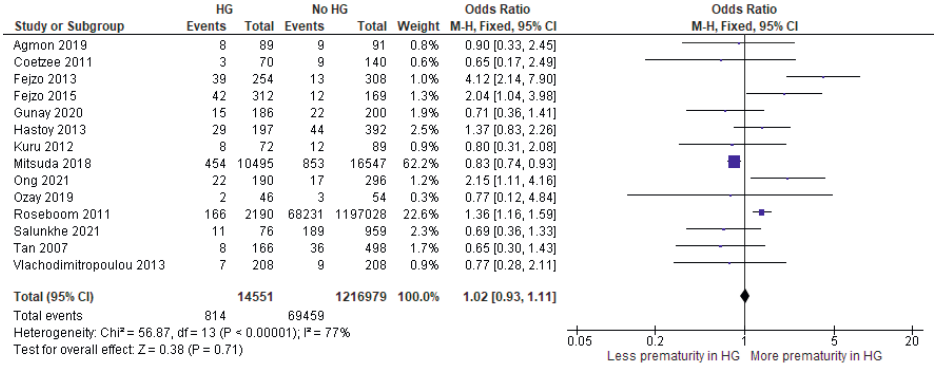
U.28 Subgroup analysis studies performed after 2001 – Placental abruption



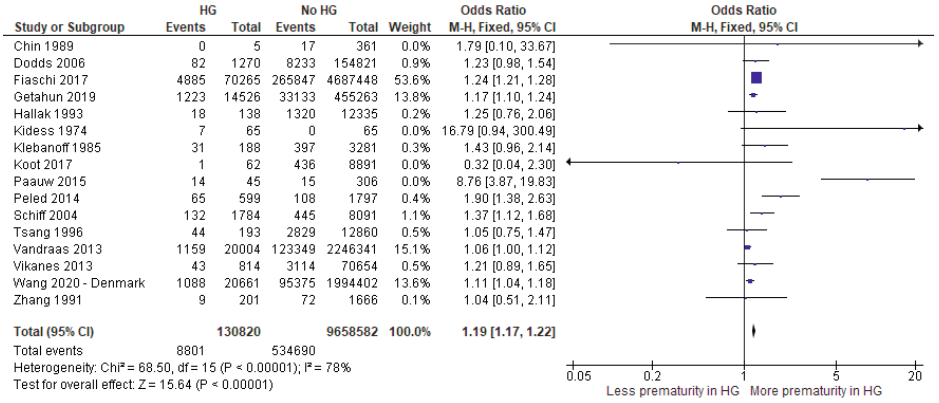
U.29 Subgroup analysis studies performed before 2001 – Placental abruption



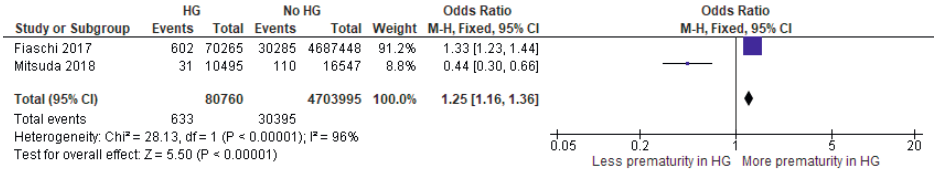
U.30 Subgroup analysis studies performed after 2001 – Preterm birth <37 weeks



U.31 Subgroup analysis studies performed before 2001 – Preterm birth <37 weeks



U.32 Subgroup analysis studies performed after 2001 – Preterm birth <32 weeks

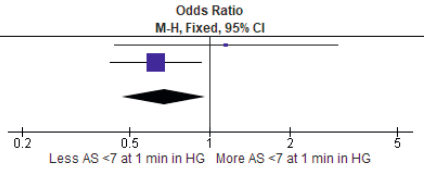


Appendix V. Subgroup analysis hospitalization

V.1 Subgroup analysis hospitalization – Apgar <7 at 1 minute

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Coetzee 2011	7	62	14	140	8.9%	1.15 [0.44, 2.99]
Vikanes 2013	26	814	3527	70654	91.1%	0.63 [0.42, 0.93]
Total (95% CI)		**876**		**70794**	**100.0%**	**0.67 [0.47, 0.97]**
Total events	33		3541			

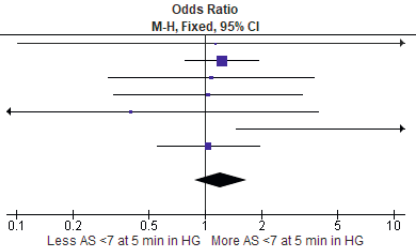
Heterogeneity: Chi² = 1.29, df = 1 (P = 0.26); I² = 23%
Test for overall effect: Z = 2.15 (P = 0.03)



V.2 Subgroup analysis hospitalization – Apgar score <7 at 5 minutes

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Coetzee 2011	1	62	2	140	1.9%	1.13 [0.10, 12.71]
Dodds 2006	19	1270	1898	154821	47.8%	1.22 [0.78, 1.93]
Gunay 2020	5	186	5	200	7.4%	1.08 [0.31, 3.78]
Hallak 1993	3	138	259	12335	8.8%	1.04 [0.33, 3.27]
Kuru 2012	1	72	3	89	4.2%	0.40 [0.04, 3.97]
Peled 2014	4	599	0	1797	0.4%	27.17 [1.46, 505.32]
Vikanes 2013	10	814	838	70654	29.6%	1.04 [0.55, 1.94]
Total (95% CI)		**3141**		**240036**	**100.0%**	**1.21 [0.88, 1.66]**
Total events	43		3005			

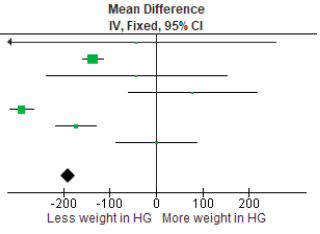
Heterogeneity: Chi² = 5.57, df = 6 (P = 0.47); I² = 0%
Test for overall effect: Z = 1.15 (P = 0.25)



V.3 Subgroup analysis hospitalization – Birth weight

Study or Subgroup	HG Mean	HG SD	HG Total	No HG Mean	No HG SD	No HG Total	Weight	Mean Difference IV, Fixed, 95% CI
Chin 1989	3,142	341	5	3,186	447	361	0.3%	-44.00 [-346.43, 258.43]
Kidess 1974	3,305	79	65	3,442	53	65	45.7%	-137.00 [-160.13, -113.87]
Kuru 2012	3,121	595.4	72	3,164	664.5	89	0.6%	-43.00 [-237.87, 151.87]
Muraoka 2020	3,148.9	348	34	3,070.9	316	70	1.3%	78.00 [-60.43, 216.43]
Paauw 2015	3,232	88	45	3,523	29	306	36.4%	-291.00 [-316.92, -265.08]
Peled 2014	3,074	456	599	3,248	543	1797	12.5%	-174.00 [-218.31, -129.69]
Tan 2007	3,000	500	166	3,000	500	498	3.2%	0.00 [-87.83, 87.83]
Total (95% CI)			**986**			**3186**	**100.0%**	**-189.77 [-205.41, -174.13]**

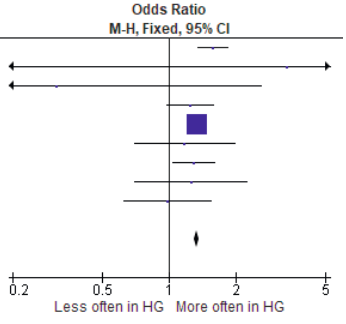
Heterogeneity: Chi² = 114.48, df = 6 (P < 0.00001); I² = 95%
Test for overall effect: Z = 23.78 (P < 0.00001)



V.4 Subgroup analysis hospitalization – Birth weight <2500 grams

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bailit 2005	177	2270	24812	486505	2.4%	1.57 [1.35, 1.84]
Chin 1989	0	5	9	361	0.0%	3.37 [0.17, 65.48]
Coetzee 2011	1	62	7	140	0.0%	0.31 [0.04, 2.59]
Dodds 2006	72	1270	7143	154821	1.2%	1.24 [0.98, 1.58]
Fiaschi 2017	6091	80708	312276	5410519	94.0%	1.33 [1.30, 1.37]
Hastoy 2013	26	197	45	392	0.3%	1.17 [0.70, 1.96]
Schiff 2004	106	2110	386	9783	1.4%	1.29 [1.03, 1.60]
Tan 2007	18	166	44	498	0.2%	1.25 [0.70, 2.24]
Vikanes 2013	20	814	1771	70654	0.4%	0.98 [0.63, 1.53]
Total (95% CI)		**87602**		**6133673**	**100.0%**	**1.33 [1.30, 1.37]**
Total events	6511		346493			

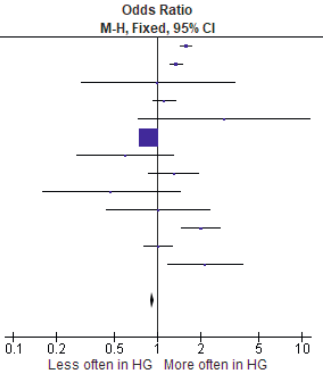
Heterogeneity: Chi² = 9.18, df = 8 (P = 0.33); I² = 13%
Test for overall effect: Z = 22.10 (P < 0.00001)



V.5 Subgroup analysis hospitalization – Small for gestational age

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Baillit 2005	663	2270	101193	486505	3.9%	1.57 [1.43, 1.72]
Bolin 2013	381	12270	26683	1142763	3.3%	1.34 [1.21, 1.49]
Coetzee 2011	4	62	9	140	0.0%	1.00 [0.30, 3.39]
Dodds 2006	137	1270	15217	154821	1.3%	1.11 [0.93, 1.33]
Fejzo 2013	7	254	3	308	0.0%	2.88 [0.74, 11.26]
Fiaschi 2017	7662	66797	704297	5424430	89.8%	0.87 [0.85, 0.89]
Gunay 2020	11	186	19	200	0.1%	0.60 [0.28, 1.29]
Hastoy 2013	52	197	85	392	0.2%	1.30 [0.87, 1.93]
Kuru 2012	5	72	12	89	0.1%	0.48 [0.16, 1.43]
Ong 2021	8	67	35	296	0.1%	1.01 [0.45, 2.29]
Peled 2014	76	599	122	1797	0.3%	2.00 [1.47, 2.70]
Vikanes 2013	78	814	6715	70654	0.8%	1.01 [0.80, 1.28]
Vlachodimitropoulou 2013	35	208	18	208	0.1%	2.14 [1.17, 3.91]
Total (95% CI)		**85066**		**7282603**	**100.0%**	**0.92 [0.90, 0.94]**
Total events	9119		854408			

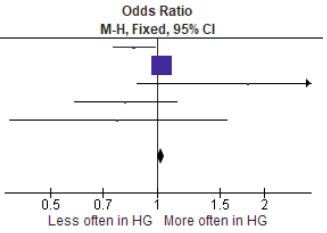
Heterogeneity: Chi² = 252.91, df = 12 (P < 0.00001); I² = 95%
Test for overall effect: Z = 7.27 (P < 0.00001)



V.6 Subgroup analysis hospitalization – Large for gestational age

Study or Subgroup	HG Events	HG Total	No HG Events	No HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Baillit 2005	222	2270	54498	486505	3.0%	0.86 [0.75, 0.99]
Fiaschi 2017	8881	66797	704297	5424430	96.3%	1.03 [1.00, 1.05]
Ong 2021	12	67	32	296	0.1%	1.80 [0.87, 3.71]
Peled 2014	48	599	174	1797	0.5%	0.81 [0.58, 1.13]
Vlachodimitropoulou 2013	15	208	19	208	0.1%	0.77 [0.38, 1.57]
Total (95% CI)		**69941**		**5913236**	**100.0%**	**1.02 [1.00, 1.04]**
Total events	9178		759020			

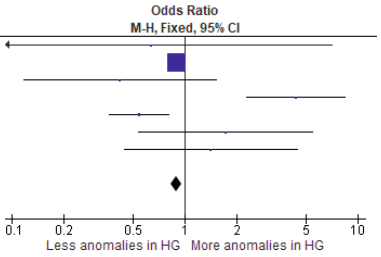
Heterogeneity: Chi² = 10.99, df = 4 (P = 0.03); I² = 64%
Test for overall effect: Z = 1.91 (P = 0.06)



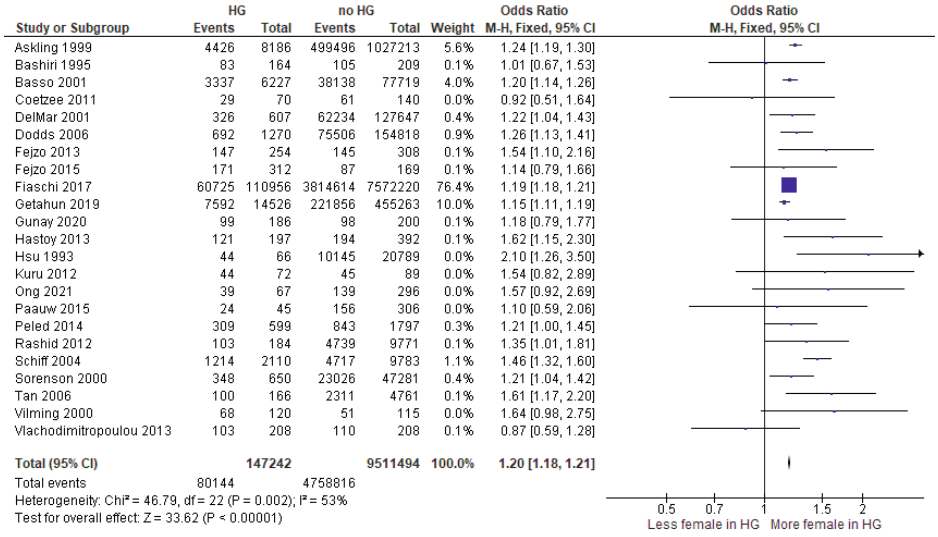
V.7 Subgroup analysis hospitalization – Congenital anomalies

Study or Subgroup	HG Events	HG Total	no HG Events	no HG Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bashiri 1995	1	164	2	209	0.1%	0.63 [0.06, 7.06]
Boissière-O'Neil 2021	797	15951	116992	2098069	94.4%	0.89 [0.83, 0.96]
Coetzee 2011	3	61	15	137	0.5%	0.42 [0.12, 1.51]
Coffey 1957	31	48	73	248	0.5%	4.37 [2.28, 8.39]
Czeizel 2005	33	125	11973	30156	4.1%	0.54 [0.37, 0.81]
Fejzo 2013	7	254	5	308	0.2%	1.72 [0.54, 5.48]
Hallak 1993	3	138	192	12335	0.2%	1.41 [0.44, 4.45]
Total (95% CI)		**16741**		**2141462**	**100.0%**	**0.89 [0.83, 0.96]**
Total events	875		129252			

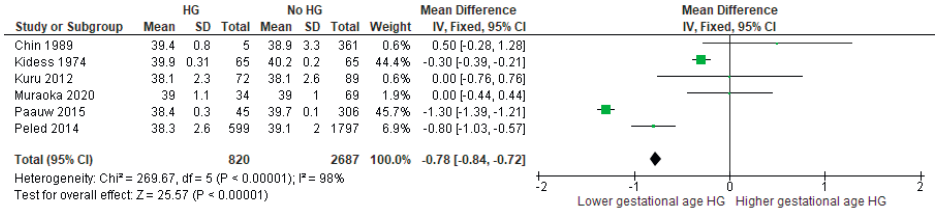
Heterogeneity: Chi² = 31.97, df = 6 (P < 0.0001); I² = 81%
Test for overall effect: Z = 3.19 (P = 0.001)



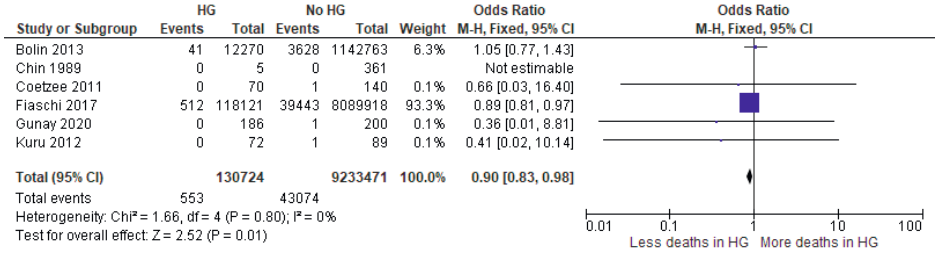
V.8 Subgroup analysis hospitalization – Fetal sex



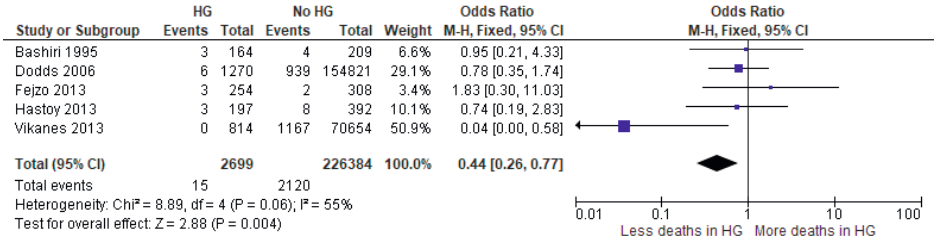
V.9 Subgroup analysis hospitalization – Gestational age at delivery



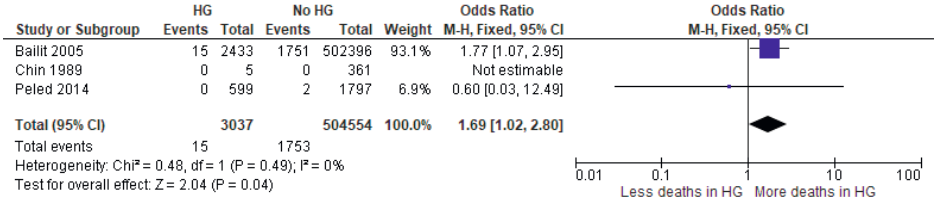
V.10 Subgroup analysis hospitalization – Stillbirth



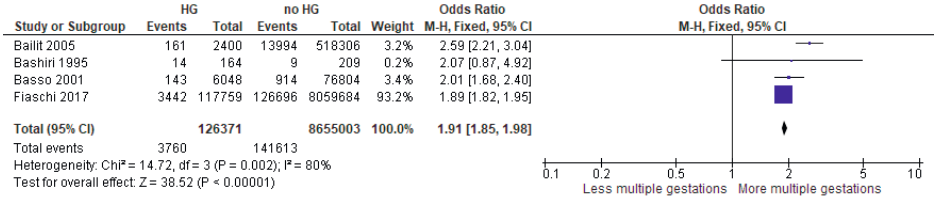
V.11 Subgroup analysis hospitalization – Perinatal death



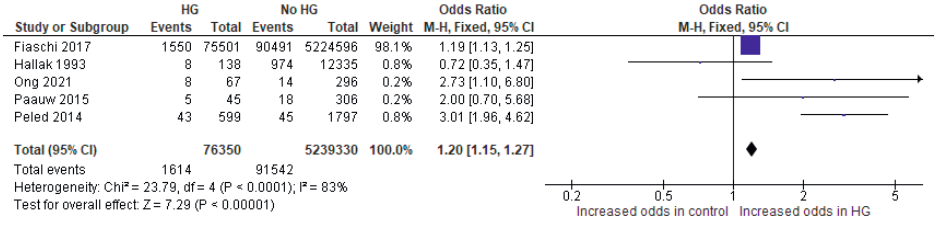
V.12 Subgroup analysis hospitalization – Neonatal death



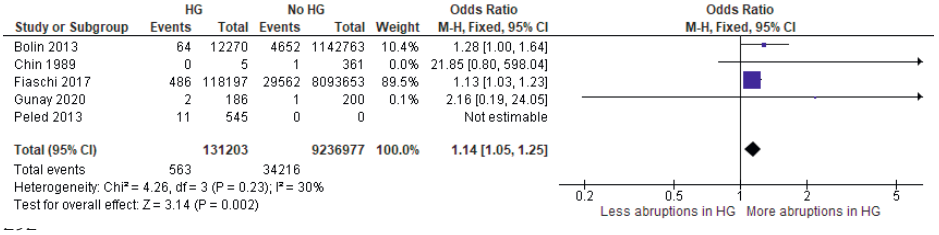
V.13 Subgroup analysis hospitalization – Multiple gestations



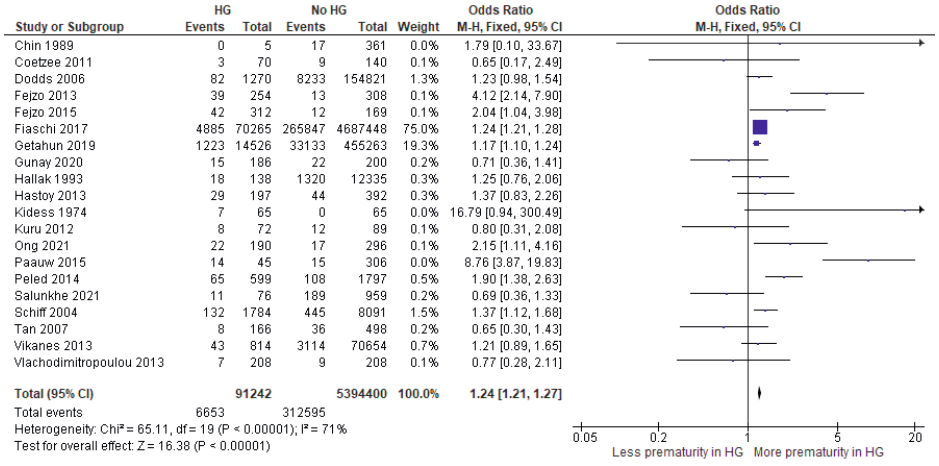
V.14 Subgroup analysis hospitalization – NICU admission



V.15 Subgroup analysis hospitalization – Placental abruption

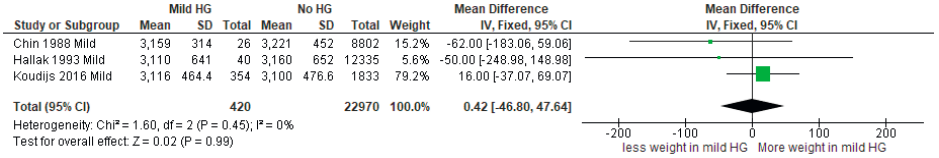


V.16 Subgroup analysis hospitalization – Prematurity <37 weeks

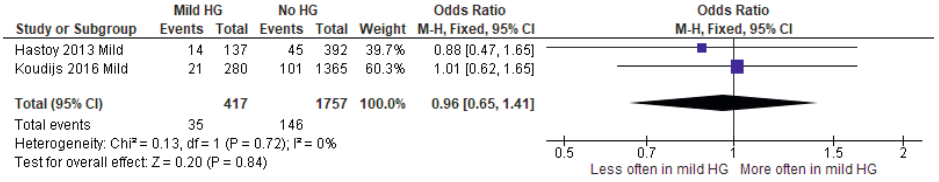


Appendix W. Subgroup analysis severity of HG

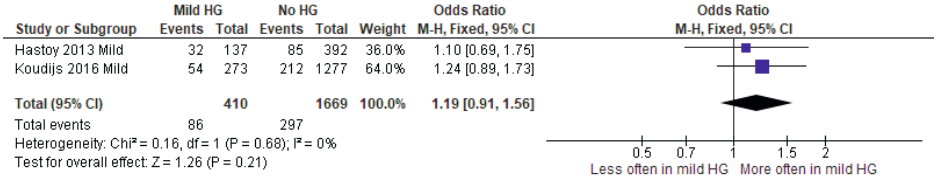
W.1 Subgroup analysis severity of HG – Birth weight – Mild HG compared with no HG



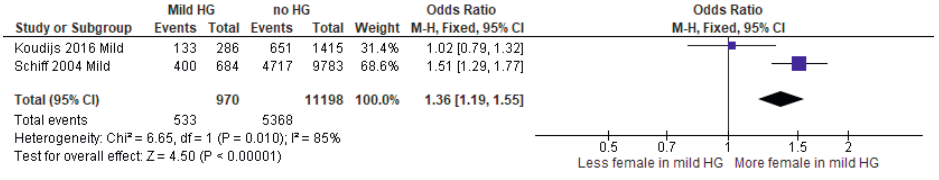
W.2 Subgroup analysis severity of HG – Birth weight <2500 grams – Mild HG compared with no HG



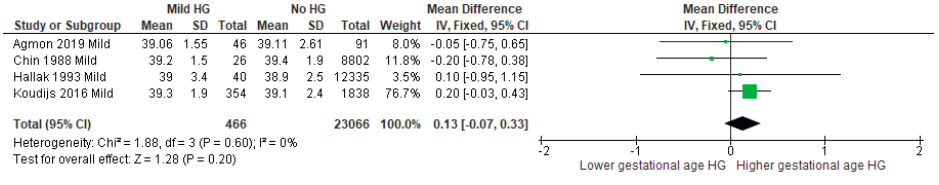
W.3 Subgroup analysis severity of HG – Small for gestational age – Mild HG compared with no HG



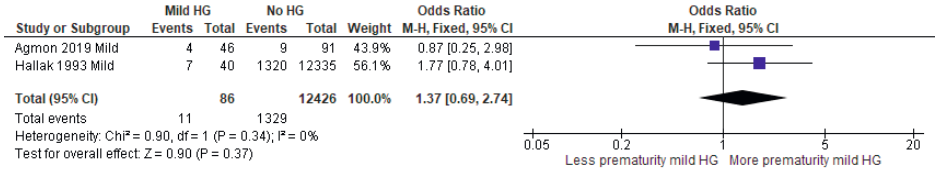
W.4 Subgroup analysis severity of HG – Fetal sex – Mild HG compared with no HG



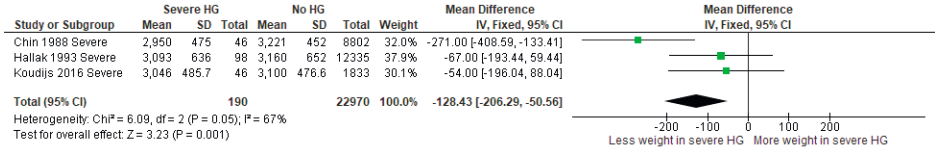
W.5 Subgroup analysis severity of HG – Gestational age at delivery – Mild HG compared with no HG



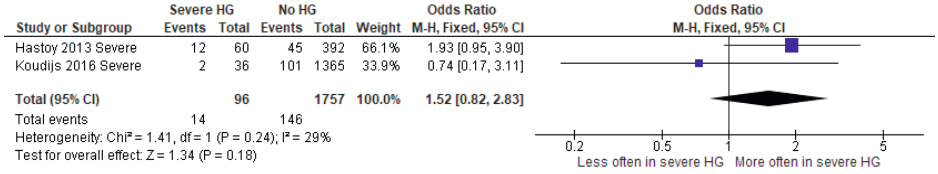
W.6 Subgroup analysis severity of HG – Prematurity <37 weeks – Mild HG compared with no HG



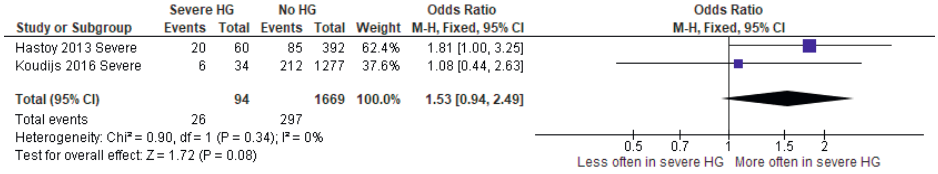
W.7 Subgroup analysis severity of HG – Birth weight – Severe HG compared with no HG



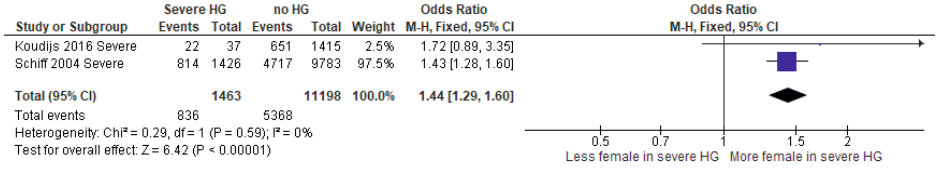
W.8 Subgroup analysis severity of HG – Birth weight <2500 grams – Severe HG compared with no HG



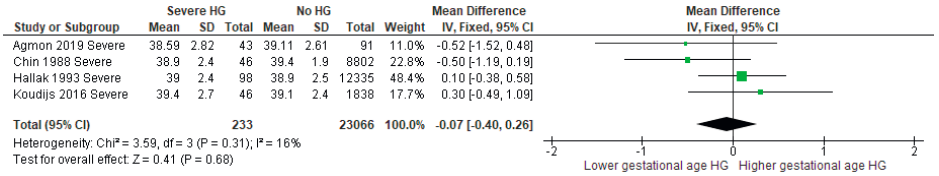
W.9 Subgroup analysis severity of HG – Small for gestational age – Severe HG compared with no HG



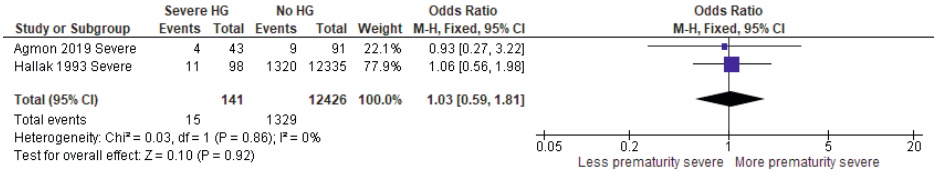
W.10 Subgroup analysis severity of HG – Fetal sex – Severe HG compared with no HG



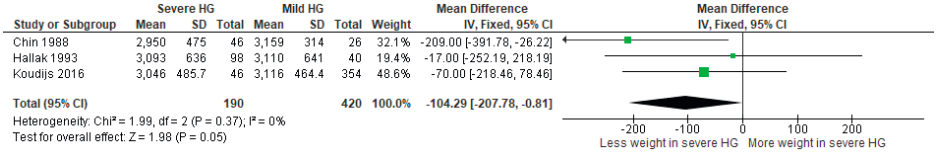
W.11 Subgroup analysis severity of HG – Gestational age at delivery – Severe HG compared with no HG



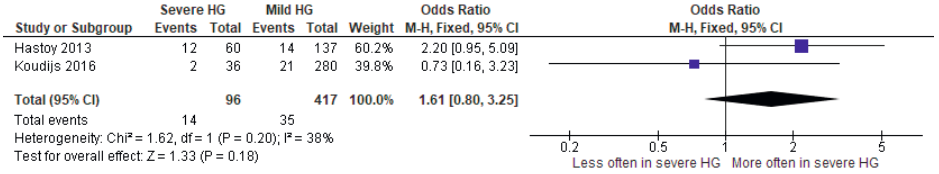
W.12 Subgroup analysis severity of HG – Prematurity <37 weeks – Severe HG compared with no HG



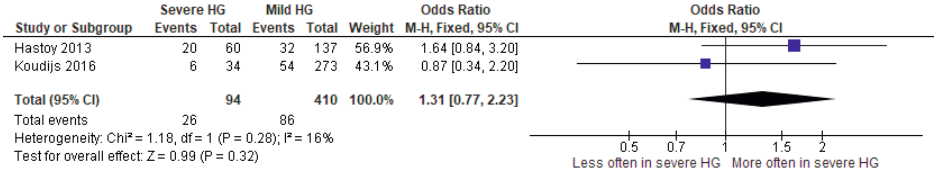
W.13 Subgroup analysis severity of HG – Birth weight – Severe HG compared with mild HG



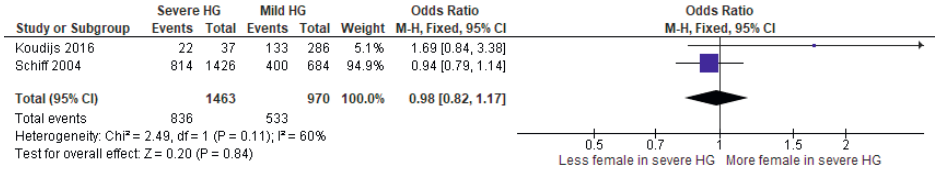
W.14 Subgroup analysis severity of HG – Birth weight <2500 grams – Severe HG compared with mild HG



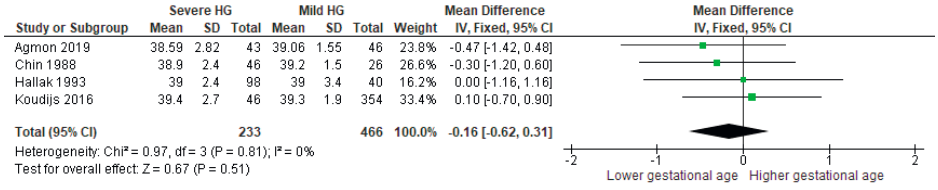
W.15 Subgroup analysis severity of HG – Small for gestational age – Severe HG compared with mild HG



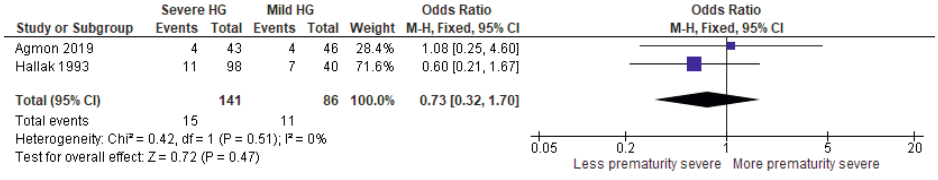
W.16 Subgroup analysis severity of HG – Fetal sex – Severe HG compared with mild HG



W.17 Subgroup analysis severity of HG – Gestational age at delivery – Severe HG compared with mild HG



W.18 Subgroup analysis severity of HG – Prematurity <37 weeks – Severe HG compared with mild HG



CHAPTER 8

Long-term health outcomes of children of mothers with hyperemesis gravidarum: a systematic review and meta-analysis

Kelly Nijsten, Larissa A.W. Jansen, Jacqueline Limpens, Martijn J.J. Finken, Marjette H. Koot, Iris J. Grooten, Tessa J. Roseboom, Rebecca C. Painter

American Journal of Obstetrics & Gynecology, 2022.

AJOG AT GLANCE

A. Why was this study conducted?

Hyperemesis gravidarum (HG) can lead to undernutrition in pregnancy. While there is evidence that HG leads to adverse perinatal effects, aggregate evidence about children's health after maternal HG is lacking at present.

B. What are the key findings?

Meta-analysis showed that children of mothers with HG had an increased chance of developing anxiety disorder (OR 1.74), sleep disorder (OR 2.94) and possibly testicular cancer (OR 1.60, signs of heterogeneity based on 95% prediction interval: 0.83-3.08). Narrative synthesis showed that maternal HG was associated with an increased risk of neurodevelopmental disorders in children, including autism spectrum disorder and attention deficit (hyperactivity) disorder. No consistent associations between HG during gestation and children's cardiometabolic outcomes were found.

C. What does this study add to what is already known?

This systematic review showed that HG is associated with a small increase in neurodevelopmental disorders, mental health disorders and possibly testicular cancer.

ABSTRACT

Objective

Hyperemesis gravidarum (HG) is characterized by severe nausea and vomiting in pregnancy, frequently resulting in severe maternal nutritional deficit. Maternal undernutrition is associated with adverse offspring health outcomes. Whether HG permanently affects offspring health remains unclear. This review aimed to evaluate effects of maternal HG on offspring health.

Data sources

Medline and Embase were searched from inception to September 6th, 2021.

Study eligibility

Studies reporting on health at any age beyond the perinatal period of children born to mothers with HG were included.

Study appraisal and synthesis methods

Two reviewers independently selected studies and extracted data. The Newcastle-Ottawa Quality Assessment Scale was used to assess risk of bias. We conducted a narrative synthesis and meta-analysis, where possible. In meta-analyses with high heterogeneity (I^2>75%), we did not provide a pooled odds ratio (OR).

Results

Nineteen studies were included in this systematic review (n=1,814,785 offspring). Meta-analysis (n=619, 2 studies: 1 among adolescents and 1 among adults) showed that HG was associated with anxiety disorder (OR 1.74, 95% CI: 1.04;2.91, I^2: 0%) and sleep problems in offspring (OR 2.94, 95% CI: 1.25;6.93, I^2: 0%). HG was associated with testicular cancer in male offspring aged up to 40 years upon meta-analysis (5 studies, n=20,930 offspring), although heterogeneity was observed based on a wide 95% prediction interval (PI) (OR 1.60, 95% CI: 1.07;2.39, I^2: 0%, 95% PI: 0.83-3.08). All six studies reporting on attention deficit (hyperactivity) disorder and autism spectrum disorder reported an increase among children of mothers with HG in comparison to children of unaffected mothers. Meta-analysis showed high heterogeneity, precluding us from reporting a pooled OR. The majority of studies reporting on cognitive and motor problems found an increase among HG-exposed children. One study investigated brain structure and found smaller cortical volumes and areas among children from HG affected pregnancies in comparison to unaffected pregnancies. Studies evaluating anthropometry and cardiometabolic disease risk of HG-exposed children had inconsistent findings.

Conclusion

Our systematic review showed that maternal HG is associated with small increases in adverse health outcomes among children, including neurodevelopmental disorders, mental health disorders and possibly testicular cancer, although evidence is based on few studies of low quality.

INTRODUCTION

Hyperemesis gravidarum (HG) is a pregnancy condition consisting of severe nausea and vomiting. Most commonly, symptoms arise in early pregnancy, commonly to improve before 20 weeks gestation, although symptoms can persist until delivery.[1, 2] HG is accompanied by poor nutritional intake and can lead to dehydration, electrolyte disturbances and weight loss.[1] Due to the lack of an available cure, treatment is symptomatic and supportive.[3]

There is a growing body of evidence linking in utero undernutrition to an increased cardiometabolic and mental health disease risk in later life.[4] In particular, maternal undernutrition in early pregnancy can have marked effects on offspring health in later life.[5-7] In the first trimester, when organogenesis takes place, specific nutrient deficiencies including folic acid or vitamin K, are of particular relevance as they can lead to congenital anomalies.[8, 9] In addition, there is evidence showing that specific maternal nutritional deficiencies during pregnancy, such as vitamin B12, folic acid and iron deficiencies, can have a negative impact on children's neurobehavioral development.[10] Since HG can lead to general undernutrition as well as specific nutrient deficiencies with an onset in early pregnancy, it is likely that HG could impact health of offspring in childhood and in adulthood.[11] This notion is at odds with existing guidance for health care providers, which emphasizes the need to reassure HG patients that 'hyperemesis gravidarum portends well for pregnancy outcome'.[12]

A previous systematic review, published in 2012, found only sparse literature on health beyond the perinatal period of children born to women with HG; only one study was included that identified excessive nausea during pregnancy as a risk factor for developing testicular cancer in male offspring.[13]

Recently, long-term health of children born to mothers with HG was placed in the top 10 of most urgent priorities in HG research by stakeholders, including patients.[14] Therefore, we aimed to update the systematic summary of the available evidence on long-term health outcomes of children born to mothers with HG.

METHODS

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A review protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) as CRD42020209560.

Search strategy

A medical information specialist (J.L.) performed a search in Medline and Embase from inception to September 6th, 2021. We used respectively MeSH- or EMTREE-terms and text words for the concepts: 1. hyperemesis gravidarum and 2. child, offspring, pregnancy outcomes or long term effects. Animal studies, reviews, case reports, editorials and conference abstracts (EMBASE) were excluded. No other limitations, including date and language, were applied. For the complete search strategies see **Appendix A1**. The search also included perinatal outcomes, which will be discussed in a separate systematic review that is currently in progress. All references were imported in ENDNOTE (X9.3.3) and duplicates were removed. Reference lists and citing articles of identified relevant papers were checked for additional relevant studies using Web of Science.

Protocol deviations

For practical reasons, a number of deviations from the published PROSPERO protocol were necessary. We did not search the Cochrane Central Register of Controlled Trials (CENTRAL), which only contains controlled clinical trials and mainly has added value as a source for grey literature. Since we expected to find evidence from observational studies only and we excluded conference abstracts and other grey literature, we limited our search to Medline and Embase. Additionally, none of the relevant studies that were identified by cited reference searching before constructing the actual search were found in CENTRAL. By repeating cited reference searching after completion of the search we aim to ensure that the searches are comprehensive. Furthermore, we decided not to use the Risk Of Bias in Non-randomized Studies - of Exposure (ROBINS-E) tool, since this tool unfortunately had remained 'under development' during the entire review process.

Study selection

Two reviewers (K.N. and L.J.) independently screened titles and abstracts using Rayyan,[15] after which potentially eligible studies were obtained in full text. Two reviewers (K.N. and L.J.) independently performed a second eligibility check for studies in the full text. Any disagreements were discussed until consensus was reached or, if necessary, a third reviewer was consulted (R.P.).

Eligibility criteria were:

- Studies reporting on long-term health outcomes of offspring born to mothers with severe nausea and vomiting in pregnancy (NVP) or HG, as reported by the authors.
- Long term health outcomes included: general health, growth development, cardiometabolic outcomes, cognitive development, behavioral development, neurodevelopment, mental health and cancer.

Exclusion criteria were:

- Case reports, case series, letters to the editor, conference abstracts and reviews
- Studies not reporting a control group, unless offspring's long-term health among the disease spectrum of severity of HG was assessed.

Data extraction

A piloted data extraction form was used to extract data by one reviewer (K.N.), which was critically appraised by a second reviewer (L.J.). Any disagreements were solved by consensus and in case of persistent disagreement, a third reviewer was consulted (R.P.). Authors were contacted by email if data was unclear or missing.

Assessment of risk of bias

A quality assessment was performed independently by two reviewers (K.N. and L.J.). Any disagreements were discussed until consensus was reached. The Newcastle-Ottawa Scale (NOS) was used, which consists of 8 questions with a maximum score of 9.[16] Studies scoring ≥7 were considered as good quality, ≥5 as fair quality and ≤4 as poor quality. Low quality upon assessment was not a reason for exclusion.

Data synthesis

Findings were described by meta-analysis, where sufficient data were available, or otherwise described narratively. In the meta-analyses, data were presented as odds ratios (OR) with corresponding 95% confidence intervals (95% CI). Random effects models according the Mantel-Haenszel method were used, based on anticipated heterogeneity. We used I^2 statistics to assess heterogeneity with I^2 values >75% considered as high heterogeneity. In those cases, we did not provide a pooled OR and we performed a sensitivity analysis, if possible. We also assessed heterogeneity by calculating 95% prediction intervals (PIs) of pooled ORs of meta-analyses that included at least three studies, to give an estimate of an interval in which 95% of effects which might be found in future, comparable studies.[17] We assessed publication bias by analyzing funnel plots if at least 10 studies reported on the same outcome, according

to the Cochrane Handbook for Systematic Reviews of Interventions.[18] *P*-values below 0.05 were considered statistically significant. Review Manager (RevMan, version 5.4, The Cochrane Collaboration, 2020) was used to conduct meta-analyses.

Strength of the evidence

Strength of the evidence of all meta-analyses was assessed by the GRADE method according to the GRADE handbook.[19] Two reviewers (K.N. and L.J.) independently graded evidence with use of the GRADEpro (Evidence Prime Inc, Ontario, Canada). Since only observational studies were included, the initial quality of evidence started at very low and was upgraded if there was: 1. A large magnitude of an effect (one level up in case of a RR <0.5 or >2; two levels up in case of a RR <0.2 or >5), 2. Signs of a dose response relationship or 3. In case of plausible residual confounding. Evidence could be rated as very low, low, moderate or high quality.

RESULTS

Study selection

Our search identified 1360 unique studies. Nineteen studies were considered eligible and included in this systematic review, as shown in **Figure 1**.[2, 20-37]

Study characteristics

Characteristics of included studies are summarized in **Table 1.** Two studies of Fejzo *et al.*[22, 23] were both follow-up survey studies of children of the same population at the age of 8 and 12 respectively. Both studies were included in this systematic review, but only the latter of the two was included in meta-analysis, based on our ability to produce 2 × 2 tables and because of a longer follow-up period.[22] Furthermore, Wang *et al.*[34] validated their results of an American cohort in a different, Danish cohort and therefore, these two study populations were described separately in this systematic review.

Of the studies included, 12 were cohort studies and 8 case-control studies. HG diagnosis was based on self-reports in seven studies,[2, 22, 23, 25, 28, 30, 37] based on interviews in seven studies,[21, 27, 29, 31, 32, 34, 36] based on ICD codes in five studies,[24, 26, 33-35] and derived from medical records in one study.[20] In total, 1,814,785 children were included in this systematic review of whom 36,546 children were born to women who experienced HG.

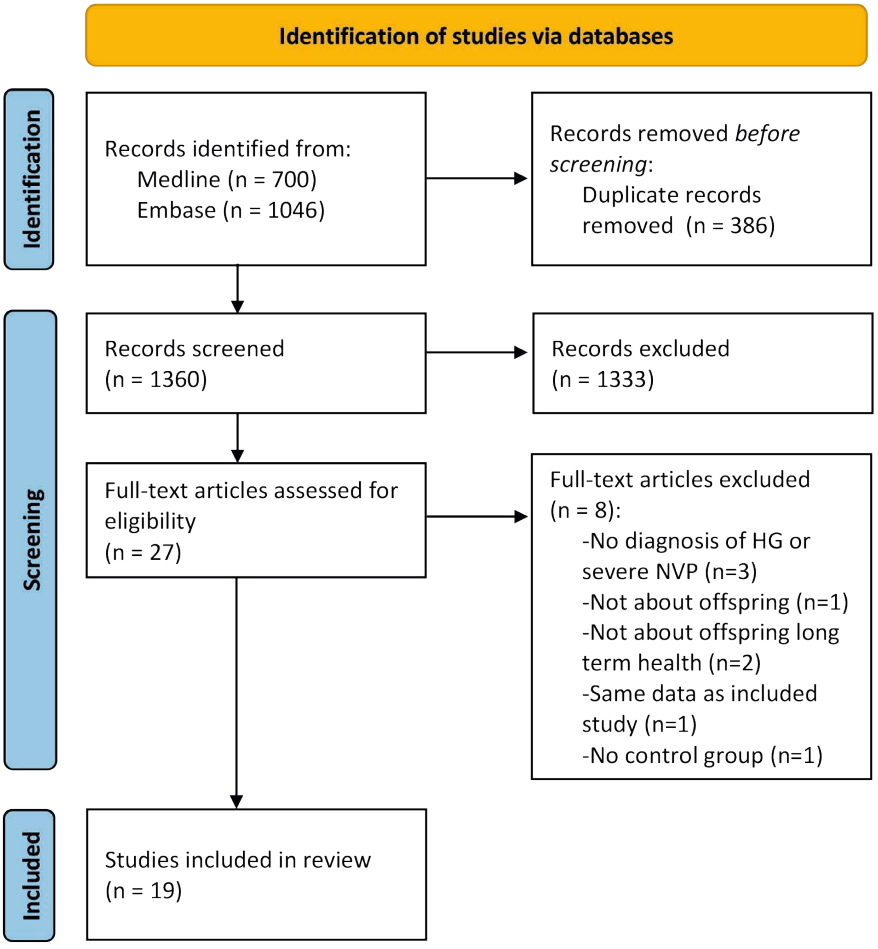


Figure 1. PRISMA flow diagram: selection process of articles

Risk of bias of included studies

Four studies, three cohort and one case-control studies, were rated to be of poor quality, as shown in **Table 2**.[2, 22, 23, 28] These studies scored low on either the *selection* domain, because of being self-selected survey studies and/or on the *comparability* domain, because of not adjusting for confounders in statistical analyses. Three studies were rated as fair quality (two cohort and one case-control study).[27, 31, 37] Twelve studies, including six cohort and six case-control studies, were rated as good quality.[20, 21, 24-26, 29, 30, 32-36] We were not able to conduct funnel plots in order to assess publication bias due to the low number of studies included in meta-analyses.

Table 1. Study characteristics

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Ayyavoo *et al.* (2013)	New Zealand	Prospective case-control	2011	*From medical records:* Admitted with severe HG with electrolyte abnormalities	36/78	Mean age 8.8 (±1.9) years	**Anthropometry:** Height, weight, body composition **Cardiometabolic health:** Insulin metabolism (sensitivity, glucose, insulin), HDL, LDL, total cholesterol, IGF-I, IGF-II, IGFBP- 1, IGFBP-3, cortisol, leptin, adiponectin, androstenedione, dehydroepiandrosterone and sulphate.	Assessed during clinical visit, including: - Whole-body dual x-ray absorptiometry - Fasting venous blood samples and a 90 minute frequent sampling intravenous glucose tolerance test
Depue *et al.* (1983)	USA	Case-control	1973-1979	*Interviews:* Treated nausea during index pregnancy (or during index pregnancy for first birth)	35/176*	Between 16 and 30 years at time of testicular cancer diagnosis	**Cancer risk:** Risk factors for testicular cancer	Population-based cancer registry (Cancer surveillance program Los Angeles County)
Fejzo *et al.* (2009)	USA	Self-selected online survey cohort	2003-2005	*Self-reported:* NVP with significant weight loss and debility, typically requiring medications or IV fluids	819 *(all HG exposed)*	Mean age 32.2 (±5.4) years	**Neurodevelopment and mental health:** Autism, behavioral disorder, colic, emotional disorder, gastroesophageal reflux disease, learning disorder, sensory disorder	Self-reported - non-validated questionnaire

Table 1. *Continued*

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Fejzo *et al.* (2015)	USA	Self-selected online survey cohort	2007-2011 (follow-up in 2012)	*Self-reported:* HG diagnosis with treatment with IV fluids and/or TPN/ nasogastric feeding tube	312/418	Mean age between 8 and 9 years	**Neurodevelopment and mental health:** ADD/ADHD, learning difficulties delays, sensory integration disorder or sensory processing disorder, social development delay or social anxiety, speech or language impairment/ delay	Self-reported - non-validated questionnaire
Fejzo *et al.* (2019)	USA	Self-selected online survey cohort	2007-2017 (follow-up in 2018)	*Self-reported:* HG diagnosis with treatment with IV fluids and/or TPN/ nasogastric feeding tube	267/360	Mean age between 11 and 13 years	**Neurodevelopment and mental health:** ADD/ADHD, anxiety, ASD, sensory integration disorder or sensory processing disorder, sleep difficulty, social developmental delay or anxiety	Self-reported - non-validated questionnaire
Getahun *et al.* (2019)	USA	Retrospective cohort	1991-2014	ICD-9	14,526/ 469,789	2 to 17 years	**Neurodevelopment and mental health:** ASD	At least one documented DSM-IV-TR code for ASD at two separate visits during the follow-up period
Gu *et al.* (2021)	China	Prospective cohort	2011-2018	*Interviews:* Severe NVP	232/1,942	At 1, 3, 6, 12, 18 and 24 months	**Anthropometry:** Height, weight	
Henderson *et al.* (1979)	USA	Case-control	1972-1974	*Self-reported:* Excessive nausea during pregnancy	12/156*	Between 15 and 40 years	**Cancer risk:** Risk factors for testicular cancer	Population-based cancer registry (Cancer Surveillance Program Los Angeles County)

Table 1. *Continued*

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Hisle-Gorman *et al.* (2018)	USA	Retrospective case-control	2000-2013	ICD-9	2,459/35,040	Between 2 and 8 years	**Neurodevelopment and mental health:** ASD	ICD-9 code for ASD at two separate visits during data collection period
Koot *et al.* (2017)	The Netherlands	Prospective cohort	1985-1986 (follow-up in 2002)	ICD-8	42/6,462	16 years	**Anthropometry:** Height; weight; BMI; waist/hip ratio; mean SBP; mean DBP **Cardiometabolic health:** Insulin metabolism (insulin, glucose, HOMA-IR) and lipid profile (Apo-A, Apo-B, HDL, LDL, total cholesterol, triglycerides)	Assessed during clinical visit, including: - Fasting venous blood samples
Koren *et al.* (2018)	Canada/ Israel	Review of prospective cohort	2006-2012	*Interviews:* Women hospitalized for severe NVP symptoms	22/241	Between 3.5 and 7.0 years	**Neurodevelopment and mental health:** Verbal IQ, performance IQ, full scale IQ	Assessed during clinical visit: - Revised Wechsler Intelligence Scale for Children (WISC-R)
Mullin *et al.* (2011)	USA	Prospective case-control	Not reported	*Self-reported:* HG diagnosis with treatment with IV fluids and/or TPN/ nasogastric feeding tube	87/259	Median between 33 and 34 years	**Neurodevelopment and mental health:** ADD, ADHD, alcoholism/ drug addiction, anxiety, Asperger's, autism, bipolar disorder, delayed sleep phase syndrome, depression, dyslexia, emotional disorder, learning disorder, obsessive compulsive disorder, Rett syndrome, schizophrenia, speech delay, Tourette syndrome	Self-reported - non-validated questionnaire

Table 1. *Continued*

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Ong *et al.* (2021) †	Singapore	Prospective cohort	2009-2010	*Self-reported:* Severe vomiting: regular vomiting with inability to retain meals	190/1,172	At 3 weeks and at 3, 6, 9, 12, 15, 18, 24, 36, 48, 54, 60, 66 and 72 months	**Anthropometry:** Height, weight, BMI	
Petridou *et al.* (1997)	Greece	Case-control	1993-1994	*Interviews:* Severe nausea during pregnancy	31/295	Not reported	**Cancer risk:** Risk factors for testicular cancer	Medical records of 4 hospitals in the Greater Athens area. Testicular cancer was histologically confirmed in all cases
Poeran *et al.* (2019)	The Netherlands	Prospective cohort	2001-2005	*Self-reported:* Daily vomiting	463/4,769	6 years	**Anthropometry:** Height, weight, total body fat mass, android/gynoid fat mass ratio, preperitoneal fat mass, SBP, DBP **Cardiometabolic health:** Insulin metabolism (insulin, c-peptide) and lipid profile (HDL, LDL, total cholesterol, triglycerides)	Assessed during clinical visit, including: - Dual-energy X-ray absorptiometry scanner - Fasting venous blood samples
Swerdlow *et al.* (1982)	United Kingdom	Case-control	1953-1973	*Interviews:* Hyperemesis, not further described	413/10,215	Up to 16 years	**Cancer risk:** Risk factors for children who died from testicular cancer	Oxford survey of childhood cancer registry

Table 1. *Continued*

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Syn *et al.* (2020) †	Singapore	Prospective cohort	2009-2010	*Interviews:* Severe vomiting with inability to retain meals	190/1,172	1 to 4.5 years	**Neurodevelopment and mental health:** **A)** Social and emotional problems **B)** ASD **C)** Cognitive development **D)** Emotional and behavioral problems **E)** Intelligence test	Assessed during clinical visits: **A)** 1-year Infant-Toddler Social and Emotional Assessment (ITSAE) **B)** 1.5-year Quantitative Checklist for Autism in Toddler (Q-CHAT) **C)** 2-year Bayley Scales of Infant and Toddler Development (Bayley) **D)** 2- and 4-year Child Behavior Checklist (CBCL) **E)** 4.5-year Kaufman Brief Intelligence Test (KBIT)
Vandraas *et al.* (2015)	Norway	Nested retrospective case-control	1967-2009	ICD-8 and ICD-10	915/162,514	Up to 21 years	**Cancer risk** Childhood cancer (including leukemia, lymphoma, cancer of the central nervous system, testis, bone, ovary, breast, adrenal and thyroid gland, nephroblastoma, hepatoblastoma and retinoblastoma)	National cancer registries of Denmark, Norway and Sweden (population-based database with mandatory reporting of all incident tumors)

Table 1. *Continued*

Study (year)	Country	Study design	Data collection period	HG ascertainment and definition	HG exposed/ Total sample size	Offspring's age at assessment	Outcomes	Outcomes ascertainment
Wang *et al.* (2020) **Part 1. USA cohort**	China *(USA population)*	Prospective cohort	2016-2018	*Interviews:* Severe NVP with either weight loss or prolonged disease course (>6 months gestation)	1,496/10,710	9 to 11 years	**Neurodevelopment and mental health:** **A)** Brain morphology **B)** Cognitive development **C)** Emotional and psychiatric problems, including anxiety, depression, ADHD, conduct disorders, oppositional defiant disorders and somatic complaints.	Assessed during clinical visits: **A)** Structured neuroimaging processing (MRI) **B)** NIH Toolbox Cognition Battery **C)** Child behavioral Checklist (CBCL)
Wang *et al.* (2020) **Part 2. Validation in Danish cohort**	China *(Danish population)*	Retrospective cohort	1995-2012‡ (follow-up until 2016)	ICD-10	14,189/ 1,109,370‡	Up to 18 years	**Neurodevelopment and mental health:** ADD/ADHD, ASD, childhood autism, conduct or oppositional defiant disorders, developmental disorders (including language, learning and motor skills disorders), emotional disorders	Clinical diagnosis identified from the Danish National Patient Register and Danish Psychiatric Central Research Register (ICD-10). ADHD was also diagnosed if offspring had ≥2 redeemed prescriptions for ADHD-specific medication from the National Prescription Register

* Total study sample size based on study participants having information on maternal HG available. † Ong *et al.* and Syn *et al.* evaluated the same study population, but reported on different offspring long term health outcomes. ‡ Numbers displayed are from 1995-2012, since data about offspring's long term health was only available for this time period and not from the complete original Danish cohort of 1978-2012 including 2,092,897 offspring in total. **Abbreviations:** ADD/ADHD: attention deficit (hyperactivity) disorder. Apo-A1/B: apolipoprotein A1 or B. ASD: autism spectrum disorder. DBP/SBP: diastolic or systolic blood pressure. HDL: high density lipoprotein. HG: hyperemesis gravidarum. ICD: International classification of disease. IGF(BP): Insulin-like growth factor (binding protein). LDL: low density lipoprotein. NVP: nausea and vomiting in pregnancy.

Table 2. Risk of bias assessment using the Newcastle-Ottawa Quality Assessment Scale (NOS) of included cohort and case-control studies

Cohort studies					
Studies	**Selection**	**Comparability**	**Outcome**	**Total score**	**Quality score**
Fejzo *et al*, 2009	*		*	2	Poor
Fejzo *et al*, 2015	*		**	3	Poor
Fejzo *et al*, 2019	*		*	2	Poor
Getahun *et al*, 2019	***	**	***	8	Good
Gu *et al*, 2021	***	**	**	7	Good
Koot *et al*, 2017	***	**	***	8	Good
Koren *et al*, 2018	**	*	**	5	Fair
Ong *et al*, 2021	**	**	**	6	Fair
Poeran *et al*, 2019	**	**	***	7	Good
Syn *et al*, 2020	***	**	***	8	Good
Wang *et al*, 2020	***	**	***	8	Good

Case-control studies					
Studies	**Selection**	**Comparability**	**Exposure**	**Total score**	**Quality score**
Ayyavoo *et al*, 2013	****	**	**	8	Good
Depue *et al*, 1983	****	**	*	7	Good
Henderson *et al*, 1979	****	**	*	7	Good
Hisle-Gorman *et al*, 2018	**	**	***	7	Good
Mullin *et al*, 2011	**		**	4	Poor
Petridou *et al*, 1997	****	**	*	7	Good
Swerdlow *et al*, 1982	***		***	6	Fair
Vandraas *et al*, 2015	***	**	**	7	Good

The NOS risk of bias assessment tool consisted of 8 items with a total maximum score of 9. A score ≥7 was considered as good quality, a score ≥5 as fair quality and a score ≤4 as poor quality.

Anthropometry

Five studies, including 14,423 children, reported on anthropometry measures, as shown in **Supplement Table S1**.[20, 26, 30, 36, 37] Four studies reported on height.[20, 26, 36, 37] Two studies reported on boys and girls separately (n=3,114 children) and found contradictory findings: the first study found that girls, but not boys from mothers with HG were taller at age 12, 18 and 24 months,[36] while the second study found that boys, but not girls of mothers with HG were taller at 72 months (adjusted β 0.64 SDs, 95% CI: 0.23;1.04).[37] Two studies (n=6,540 children) that analyzed boys and girls together did not find any differences in height at age 4-11 and 16.[20, 26]

Three studies reported on weight growth.[30, 36, 37] The two studies that evaluated boys and girls separately (n=3,114 children) again showed conflicting findings: the first study showed that girls, but not boys exposed to HG were *heavier* at 12, 18 and 24 months,[36] while the other study found

that exposed girls were *lighter* (adjusted β -0.53 SDs, 95% CI: -1.03;-0.03) and boys were heavier at 5 years (adjusted β 0.57 SDs, 95% CI: 0.05;1.08).[37] The third study that evaluated weight growth among both boys and girls (n=4,760 children) found that HG-exposed children weighted more compared to non-exposed children at age 2 (9,838±1,712 compared to 9,581±1,440 g).[30]

Four studies reported on children's BMI. One study (n=4,760 children) found higher BMI among 6 year old HG-exposed children compared to non-exposed children (male and female: adjusted β 0.08, 95% CI: 0.00;0.17),[30] while another study (n=1,172 children) found lower BMI among 66 months old female offspring exposed to HG (adjusted β -0.57 SDs, 95% CI: -1.09;-0.05), but no differences among male offspring.[37] The two other studies did not find differences in BMI among HG-exposed and non-exposed children at 4-11 and 16 years (n=6,540 children).[20, 26]

Lastly, one study, including 4,760 children, found that, at age 6, those born to women with HG had higher total body fat mass (adjusted β 0.12, 95% CI: 0.03;0.20), higher android/gynoid fat mass ratio (adjusted β 0.11, 95% CI 0.02;0.21) and higher abdominal preperitoneal fat mass area (adjusted β 0.10, 95% CI: 0.00;0.20) compared to those born to women without HG,[30] while two other studies (n=6,540 children) did not find any differences in waist/hip ratio, android/gynoid fat mass ratio and total body fat percentage at ages 4-11 and 16.[20, 26]

Cardiometabolic health

Blood pressure

Two studies, including 10,832 children, evaluating blood pressure had conflicting findings.[26, 30] One study (n=4,370 children) found that those born to women with HG had significantly higher diastolic (61.4±7.3 vs 60.5±6.7 mmHg) and higher systolic blood pressures (103.8±8.8 vs 102.4±8.1 mmHg) at the age of 6, which was not sustained after adjustments for confounders,[30] whereas the other study that included 6,462 adolescents of 16 years old, did not find such differences.[26] The fact that studies differed in children's ages at assessment, prohibited meta-analyses.

Cardiometabolic laboratory measures

Three studies, including 8,847 children, reported on cardiometabolic laboratory measures. None of the studies found any differences in lipid profile measures between HG-exposed and non-exposed children.[20, 26, 30] Two out of three studies (n=8,747 children) found no differences in fasting insulin, glucose, c-peptide or HOMA-IR levels at the age of 6 and 16.[26, 30] The third study (n=78 children, aged 4 to 11) found that those born to mothers with HG had significantly higher fasting insulin levels than those of control pregnancies (6.88 vs 5.04 mIU/L) after adjustments

for confounders, including ethnicity, birth weight, birth order, age, sex and BMI, were made.[20] Due to age-specific laboratory reference intervals, these results could not be pooled for meta-analyses. Insulin sensitivity, as assessed from a 90 minute frequent sampling intravenous glucose tolerance test, was also reported to be 20% lower in HG-exposed children than in non-exposed children.[20] Lastly, they reported that children of women with HG had significantly higher early-morning cortisol (256 vs 210 nmol/L), lower IGF binding protein 1 (11.8 vs 19.0 ng/mL) and lower IGF binding protein 3 levels (1955 vs 3435 ng/mL).

A narrative summary of results of included studies reporting on anthropometry and cardiometabolic health outcomes is displayed in **Supplement Table S1.**

Neurodevelopment and mental health

Cognitive and motor development

Two studies (n=9,696 children) found that HG-exposed children scored significantly lower on cognitive development scores than non-exposed children at age 2,[32] and 9 to 11 years respectively.[34] Two studies assessed intelligence by measuring IQ scores: the first study (n= 469 children), did not find any differences in IQ scores at the age of 4,[32] while another study (n=241 children) found that HG-exposed children scored significantly lower at verbal, performance and full scale IQ scores than non-exposed children aged 3.5 to 7.[27] Only the first study adjusted for confounders.[32]

Three studies reported on learning difficulties and speech or language impairment/delay.[23, 28, 32] One of them found that families with at least one HG-exposed child more often reported learning delays (12.3%) and speech or language impairment/delays (24.1%) compared to families where none of the children were exposed to HG (3.4% and 11.2% respectively), without adjusting for possible confounders.[23] Another study (n= 482 children at age 2) did not find any differences in language skills, measured by the Bayley-III language scale, between HG-exposed and non-exposed children in multivariable regression analysis.[32] The third study (n= 259 adults) did not assess associations per separate outcome, but only stated that 38% of HG-exposed offspring reported having a psychological and/or behavioral disorder (including learning disorders and speech delay) compared to 15% of non-exposed offspring (unadjusted OR 3.57, 95% CI: 1.87;6.90, P<0.001).[28]

An association between HG and neurosensory disorders in children was found in two studies of Fejzo *et al.*,[22, 23] including 292 families and 360 children respectively at the age of 8-9 and 11-13. No adjustments were made for confounders and, as described earlier, it is likely that

these studies partially reported on the same children. One study (N=1,109,370 children, aged up to 18) reported on developmental problems, including language, learning, and motor skills problems, and found that HG-exposed children had an increased chance compared to non-exposed children (adjusted HR, 1.18; 95% CI, 1.05–1.34),[34] while another study (n=475 children, aged 2) did not find any differences in Bayley-III scores on the subdomain motor difficulties.[32]

A narrative summary of results of included studies reporting on cognitive and motor development is displayed in **Supplement Table S2.**

Mental health

All eight included studies reporting on mental health found an increased risk for HG-exposed children compared to non-exposed children, as narratively summarized in **Supplement Table S3**. A meta-analysis including two studies (n=1,109,629 offspring, including children aged up to 18 and adults) that reported on emotional problems found no significant association (OR 1.19, 95% CI: 0.89;1.61, I^2: 0%), as shown in **Figure 2** (GRADE level of evidence: very low quality, see **Supplement Table S4**).[28, 34] One study assessed affective problems as a subdomain of the Child Behavioral Checklist (CBCL) and found higher scores among HG-exposed than non-exposed children at 2 and 4 years old after adjustments for confounding factors.[32]

Four studies (n=10,490 offspring varying in age from 4 to adulthood) found that HG-exposed offspring were at increased risk of anxiety compared to non-exposed offspring, based on CBCL subdomains,[32, 34] as well as on self-reported symptoms.[22, 28] When pooling these last two studies in meta-analysis, we found that HG-exposure was significantly associated with anxiety disorder in offspring (OR 1.74, 95% CI: 1.04;2.91, I^2: 0%) **(Figure 2)** (GRADE level of evidence: very low quality, see **Supplement Table S4**). Two studies (n=9,473 offspring) found a significant association between HG-exposure and CBCL depression scores at age 9-11,[34] and depression rates in adulthood (unadjusted OR 6.35).[28] The latter of the two also found that bipolar disorder more often occurred in HG-exposed offspring (unadjusted OR 4.90).[28]

Two studies (n=778 children, aged 8-9 and 11-13), which are likely to partially include the same children because of being a follow-up of the same original study population, found higher rates of social developmental delay or social anxiety in HG-exposed children compared to non-exposed children (unadjusted OR 5.02 and 3.58).[22, 23]

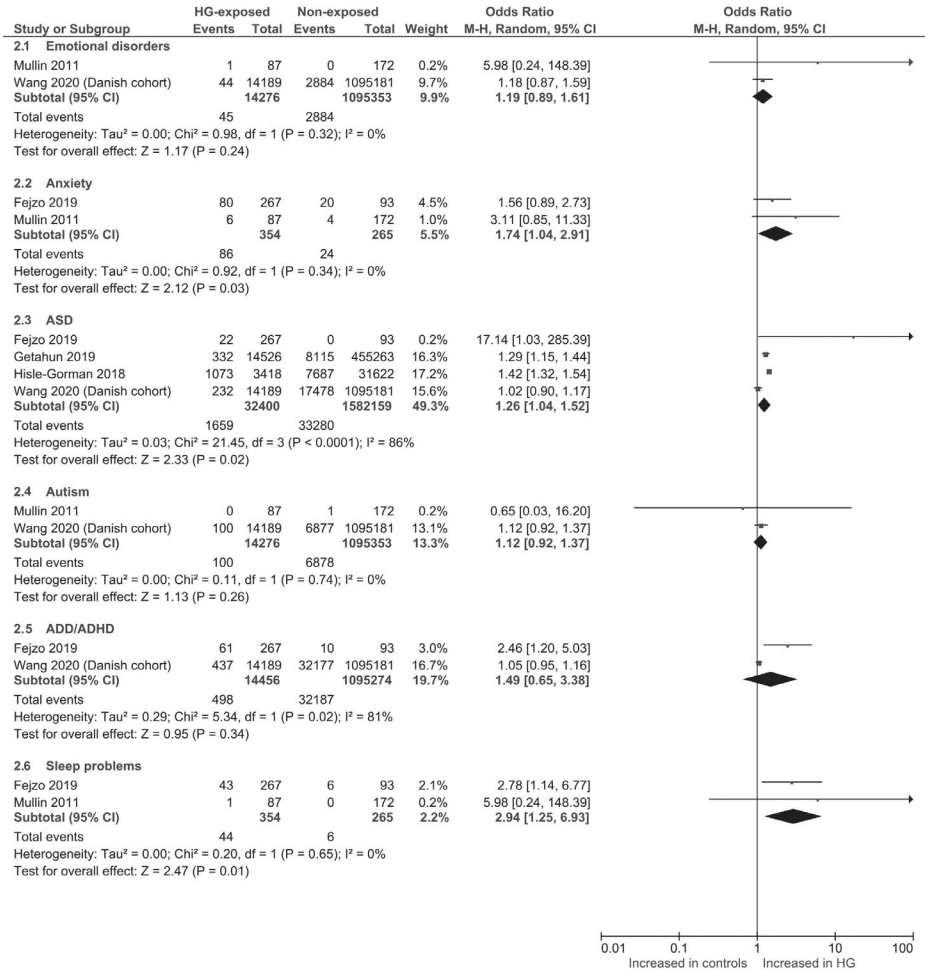


Figure 2. Meta-analyses neurodevelopment and mental health outcomes

Neurobehavioral development

As shown in a narrative summary in **Supplement Table S3,** all studies reporting on neurobehavioral development found an increase of neurobehavioral developmental disorders among HG-exposed offspring. Six studies (n=1,600,520 offspring, aged 1.5 to adulthood) reporting on Autism Spectrum disorders (ASD) or on autism alone, found that this occurred more often in HG-exposed than in non-exposed offspring.[22, 24, 28, 32, 34, 35] A meta-analysis including four studies (n= 1,614,559 children, aged 2-18) that reported on the association between HG exposure and ASD in children showed high heterogeneity based on an $I^2$of 86% **(Figure 2)** and a 95% PI of 0.54-2.95 (GRADE level of evidence: low quality, see **Supplement Table S4**).[22,

[24, 34, 35] Sensitivity analysis showed a significant association between HG exposure and ASD when including only the three studies from the USA (OR 1.37, 95% CI: 1.19;1.57, I^2: 62%, 95% PI: 0.29-6.44) **(Supplementary Figure S1)**, but did not reveal differences when excluding studies with a poorer study design (i.e. survey and case-control study) **(Supplementary Figure S2)**. No significant association was found between HG-exposure and autism alone in meta-analysis including two studies (OR 1.12, 95% CI: 0.92;1.37, I^2: 0%) (**Figure 2)** (GRADE level of evidence: very low quality, see **Supplement Table S4**).[28, 34]

A significant association between HG and attention deficit disorder (ADD) or attention deficit hyperactivity disorder (ADHD) in offspring, combined or as separate diagnoses, was found in all six studies reporting on this topic (n=1,120,014 offspring, aged 2 to adulthood). Pooling two of these studies (n= 1,109,730 children, aged up to 18) in meta-analysis that reported on ADD/ADHD combined showed substantial heterogeneity with an I^2 of 81% **(Figure 2)** (GRADE level of evidence: very low quality, see **Supplement Table S4**).[22, 34] Since only two studies were included, we were not able to perform a sensitivity analysis.

One study, reporting on two different study populations (n=1,118,584 children, aged up to 18), found that conduct and oppositional defiant disorders more often occurred in HG-exposed children and that they had higher scores on corresponding CBCL subdomains than non-exposed children.[34] Three studies reported on sleep problems. One study (n= 530 children) found increased scores on the sleep problem subdomains of the Infant-Toddler Social and Emotional Assessment and CBCL at respectively 1 and 2 years old in HG-exposed children in multivariable regression analysis.[32] A meta-analysis including the other two studies (n= 619 offspring, at the age 11-13 and adults) showed increased sleep problems among HG-exposed offspring (OR 2.94, 95% CI: 1.25;6.93, I^2: 0%) **(Figure 2)** (GRADE level of evidence: very low quality, see **Supplement Table S4**).[22, 28]

The effect of HG severity and treatment on neurodevelopment and mental health in offspring

Few studies also assessed differences in children's neurobehavioral outcomes among the disease spectrum of HG. One study (n= 819 offspring with a mean age of 32) found no significant differences in autism, behavioral, emotional, sensory and learning disorders between offspring of HG patients with and without severe weight loss (>15% of prepregnancy weight).[2] The second study (n= 418 children, aged 8-9) found that the presence of early HG symptoms (below 5 weeks gestation) was significantly associated with neurodevelopmental delay in children, but found no differences between in- or outpatient care or between different medications and

treatments.[23] Another study (n= 1,172 children, aged 1 to 4.5) found that severe NVP *without* admission was more often associated with reduced neurobehavioral development in children than severe NVP *with* admission.[32]

Two studies assessed associations between ASD rates and HG severity and treatments. The first study (n= 360 children, aged 11-13) did not find any differences between different treatments (anti-emetics or tube feeding), in- or outpatient care and early onset of symptoms (below 5 weeks gestation).[22] The second study (n= 469,789 children, aged 2-17) found a higher ASD incidence rate among children of women diagnosed with HG in the first (adjusted HR 1.58; 95% CI: 1.40;1.79, *P*<0.001) or second trimester (adjusted HR 1.36; 95% CI: 1.05;1.75, *P* 0.02) and children of HG patients with metabolic disturbances that required rehydration or tube feeding (adjusted HR 1.41, 95% CI: 1.05;1.90; *P* 0.02).[24] No differences were found in ASD rates between early and late maternal hospitalization for HG.[24]

Brain morphology and its association with cognitive and psychological symptoms in children

One study (n= 10,710 children, aged 9-11) assessed brain morphology in children and measured the cortical area, volume and thickness of each brain region.[34] They found that the total cortical volume and area were significantly smaller in children exposed to severe NVP and that these smaller volumes and area's mediated associations between severe NVP exposure and cognitive and psychological symptoms.

Cancer risk

Data on offspring's cancer risk was available in 5 studies (n=173,502 offspring, aged up to 40).[21, 25, 29, 31, 33] Four studies reported solely on testicular cancer and assessed whether HG during pregnancy was a risk factor for the disease.[21, 25, 29, 31] The fifth study reported whether HG was a risk factor for developing multiple types of childhood cancer, including testicular cancer.[33] Meta-analysis including these 5 studies showed that HG was significantly associated with testicular cancer in male offspring (OR 1.60, 95% CI: 1.07-2.39, I^2 = 0%, 95% PI: 0.83-3.08) as shown in **Figure 3** (GRADE level of evidence: very low quality, see **Supplement Table S4**). Sensitivity analyses showed that a significant association between HG exposure and testicular cancer in male offspring only remained after omitting the study of Vandraas *et al.*[33] **(Supplement Figures S3 to S7)**, which means that this study has a large impact on the results. The study of Vandraas *et al.*[33] is a large, case-control study (n=162,514 offspring) that did not find a significant association between HG exposure and testicular cancer in offspring.

No association was found between HG and multiple childhood cancers in one study including 162,514 offspring in total.[33] When dividing their study population in offspring aged 1 to 10 years old and 10 to 20 years old however, HG was significantly associated with lymphoma in offspring aged 10 to 20 years (adjusted RR 2.08, 95% CI: 1.11-3.90).[33]

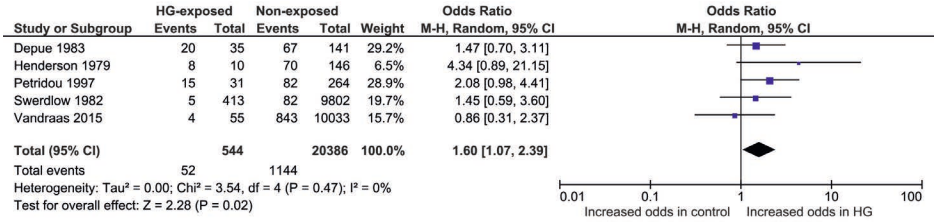


Figure 3. Meta-analysis testicular cancer

DISCUSSION

Main findings

Our systematic review identified 19 studies reporting on children's long-term health after HG-exposure in utero. We were able to conduct meta-analyses for seven outcomes and found that maternal HG was significantly associated with anxiety disorder and sleep problems in offspring, while no associations were found with emotional disorders and autism. Maternal HG was also significant associated with testicular cancer in male offspring in meta-analyses, although there were signs of heterogeneity based on the 95% PI, that is considerably wider than the CI and which crosses 1. All studies reporting on ADD/ADHD and ASD found an increased risk among children of mothers with HG, but showed high heterogeneity when pooling these results in meta-analysis. Narrative synthesis showed that a majority of the studies showed an increase in cognitive and motor problems in children of mothers with HG. One study showed that children's brain morphology was affected, with children exposed to severe NVP having smaller total cortical volume and area compared to non-exposed children. Inconsistent associations were found between HG exposure and anthropometry and cardiometabolic disease risk markers. Studies reporting on the effect of HG disease severity or treatment on children's long term health had inconsistent findings. All evidence included in this systematic review is based on a small number of studies with evidence of meta-analyses graded to be of very low to low quality.

Strengths and limitations

This systematic review has several strengths. A research protocol was published online, before conducting this review. We were able to include studies on multiple long-term health domains, including cardiometabolic health, neurodevelopment, mental health and cancer risk. We also managed to collect additional data from corresponding authors of included studies in order to extend our meta-analysis. Moreover, two thirds of included studies were rated as to be of fair to good quality and we included a relatively large group of offspring that was exposed to HG.

Limitations of this review include heterogeneity in reported outcomes as well as in HG diagnosis. Heterogeneity in HG diagnoses is a problem that has been previously identified in a systematic review that concluded that only 2 criteria (nausea and vomiting) were commonly used in diagnosing HG.[38] Heterogeneity led to variances in HG rates among the included studies in our review: cohort studies diagnosing HG based on self-reports or interviews reported rates between 9.7-21.9%, whereas cohort studies including HG patients based on ICD codes reported much lower rates varying from 0.6-3.1%. While ICD codes have been shown to be only valid for diagnosing women with a mild HG,[39] ICD codes were used to diagnose HG patients in the 4 largest studies included in this review.[24, 33-35] The fact that we included mainly case-control studies, which are prone for recall bias, taken together with different methods of disease ascertainment, which may have led to over- and underestimation of HG or disease severity, this could have influenced results.

Additionally, studies reported on many different long-term health outcomes and used different methods to ascertain these. Therefore, we were not able to aggregate results in meta-analysis for a majority of the outcomes and were our results mostly presented by use of a narrative synthesis. Moreover, evidence of outcomes that were assessed by meta-analyses was graded as very low or low quality. Furthermore, the Diagnostic and Statistical Manual of Mental Disorders was updated in 2013, combining several disorders, including autism disorder, into one diagnosis: ASD.[40] This hampered us in conducting meta-analysis for ASD and autism combined, since both outcomes are used interchangeably in current literature. Lastly, some of the studies included self-selected participants that self-reported offspring outcomes, plausibly leading to higher rates of adverse health outcomes and possible recall bias.[22, 23, 28]

Comparison with existing literature

Our systematic review and meta-analysis showed that HG-exposure is associated with neurodevelopmental and mental health problems, along with a decrease in brain volume, in offspring. So far, only few studies have evaluated the effect of maternal undernutrition

during pregnancy on children's neurodevelopment and mental health. Research from the Dutch famine found that in utero exposure to undernutrition led to an increased risk for developing schizoid or antisocial personality disorders,[41, 42] while animal studies found that maternal protein restriction during pregnancy had a negative effect on anxiety and cognitive behavior in offspring.[43] Different mechanisms can be hypothesized to be the underlying cause, for example when interpreting our results in the context of ASD research. A recent systematic review, including 36 studies, found that an appropriate intake of folic acid and Vitamin D could protect against ASD in offspring.[44] Conversely, this could mean that a reduced availability of these nutrients, potentially caused by HG, may lead to an increased ASD risk in children. Another recent systematic review and meta-analysis demonstrated that SGA neonates are at increased risk of developing ASD.[45] Since previous research showed that HG children are more often born SGA, an alternative explanation could be that the risk of HG-exposed children developing ASD originates in fetal growth restriction.[13] Genetics also play an important role, whereas a large multi-country cohort study found a high heritability rate for ASD of 80%.[46] No evidence of genetic factors being involved however, was found in a self-reported survey study comparing HG-exposed offspring to unexposed siblings and to offspring with no family history of HG.[47]

We did not find evidence of a consistent effect of HG on offspring anthropometry and cardiometabolic disease risk markers. These findings are at odds with the large body of evidence on effects of maternal undernutrition in pregnancy. Animal experiments have linked intra-uterine undernutrition to adverse health outcomes in offspring in later life.[48, 49] Additionally, studies about the Dutch, Chinese and Nigerian Famine showed that adults exposed to maternal undernutrition in utero more often had type II diabetes mellitus, hypertension, overweight and coronary heart diseases in later life.[5, 7, 50, 51] This may be explained by the fact that the studies in our systematic review included children of young ages, which precludes any definitive statement about the possible effects HG may have on adult offspring. Alternatively, it could be that women included in the studies had relatively mild HG and that only women with severe HG are comparable with pregnant women during the Famine in terms of maternal undernutrition. Finally, it is possible that HG, despite incurring maternal undernutrition, simply does not have effects on cardiometabolic health in the next generation. Moreover, due to the presence of age-specific reference intervals for anthropometry and cardiometabolic laboratory measures, we were not able to pool results in meta-analysis. More research is warranted in order to draw firm conclusions about possible adverse cardiometabolic disease risks for HG-exposed children.

HG was associated with testicular cancer in male offspring with an OR of 1.60, which has been previously hypothesized to be caused by increased estrogen levels during pregnancies complicated by HG.[21, 25] Nonetheless, these results should be interpreted with caution. While there were no signs of heterogeneity based on the I^2 of zero, the 95% PI suggested otherwise, estimating that in 95% of future, subsequent studies the true effect estimate could range from 0.83 to 3.08. Additionally, only a small number of HG cases were included in most studies and in four out of five studies HG was assessed by questionnaires or interviews long after their pregnancy occurred, which could have led to recall bias.[21, 25, 29, 31] Importantly, the fifth and most recent, large study diagnosed HG based on ICD codes and did not find an association between HG and testicular cancer in offspring.[33]

We were unable to assess whether maternal treatment interventions can have a preventive role in avoiding health sequelae among the offspring. Only three studies that assessed neurodevelopment and mental health,[22-24] but none of the studies on other long term health conditions, evaluated the association with treatment interventions and showed variations in results. Sensitivity analysis showed that HG was associated with more offspring ASD only in studies from the USA, but not in Denmark. This might indicate that differences in treatment or accessibility to health care between the USA and Denmark alter offspring's ASD disease risk, which is an important matter to be addressed in future research, as it has direct implications for health care policy choices. Lastly, yet importantly, HG was not associated with more offspring ASD when excluding studies with a poorer study design.

Conclusions and implications

This systematic review and meta-analysis showed that there is an increased adverse long-term health risk for children exposed to HG during gestation, in terms of neurodevelopmental problems and mental wellbeing, while the impact on cardiometabolic disease risk and testicular cancer remains unclear. Conclusions are based, however, on a small number of studies, with evidence from meta-analyses being graded as very low to low quality. Altogether, long-term health research in HG is still in its infancy and more research with long-term follow-up is needed to determine the pathophysiology of these associations and to further explore the role of (early) treatment in order to reduce adverse long-term health effects in offspring.

Acknowledgements

We thank Rik van Eekelen for helping us calculating and interpreting 95% prediction intervals.

Funding

This systematic review was funded by the Amsterdam Reproduction & Development (AR&D) research institute (grant number 23346) and the Dutch Heart Foundation (grant number 2013T085).

Contribution to authorship

KN, LAWJ, RCP and TJR conceived and designed the study. JL performed the electronic search. KN and LAWJ screened titles and abstracts for eligibility and performed data extraction. KN performed all statistical analyses and drafted the manuscript. LAWJ, JL, MHK, IJG, RCP and TJR contributed in interpreting the results and revising the manuscript. All authors approved the final draft of this manuscript.

REFERENCES

1. Niebyl JR. Clinical practice. Nausea and vomiting in pregnancy. The New England journal of medicine. 2010;363(16):1544-50.
2. Fejzo MS, Poursharif B, Korst LM, Munch S, MacGibbon KW, Romero R, et al. Symptoms and pregnancy outcomes associated with extreme weight loss among women with hyperemesis gravidarum. J Womens Health (Larchmt). 2009;18(12):1981-7.
3. Boelig RC, Barton SJ, Saccone G, Kelly AJ, Edwards SJ, Berghella V. Interventions for treating hyperemesis gravidarum. The Cochrane database of systematic reviews. 2016(5):Cd010607.
4. Bateson P, Barker D, Clutton-Brock T, Deb D, D'Udine B, Foley RA, et al. Developmental plasticity and human health. Nature. 2004;430(6998):419-21.
5. de Rooij SR, Painter RC, Roseboom TJ, Phillips DI, Osmond C, Barker DJ, et al. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. Diabetologia. 2006;49(4):637-43.
6. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Bleker OP. Plasma lipid profiles in adults after prenatal exposure to the Dutch famine. The American journal of clinical nutrition. 2000;72(5):1101-6.
7. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Schroeder-Tanka JM, et al. Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. Heart (British Cardiac Society). 2000;84(6):595-8.
8. Viswanathan M, Treiman KA, Kish-Doto J, Middleton JC, Coker-Schwimmer EJ, Nicholson WK. Folic Acid Supplementation for the Prevention of Neural Tube Defects: An Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. Jama. 2017;317(2):190-203.
9. Howe AM, Webster WS. Vitamin K--its essential role in craniofacial development. A review of the literature regarding vitamin K and craniofacial development. Aust Dent J. 1994;39(2):88-92.
10. Nyaradi A, Li J, Hickling S, Foster J, Oddy WH. The role of nutrition in children's neurocognitive development, from pregnancy through childhood. Front Hum Neurosci. 2013;7:97-.
11. Maslin K, Shaw V, Brown A, Dean C, Shawe J. What is known about the nutritional intake of women with Hyperemesis Gravidarum?: A scoping review. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2021;257:76-83.
12. ACOG Practice Bulletin No. 189: Nausea And Vomiting Of Pregnancy. Obstetrics and gynecology. 2018;131(1):e15-e30.
13. Veenendaal MV, van Abeelen AF, Painter RC, van der Post JA, Roseboom TJ. Consequences of hyperemesis gravidarum for offspring: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2011;118(11):1302-13.
14. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient–clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ Open. 2021;11(1):e041254.
15. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
16. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. 2014 [Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
17. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247.

18. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). 2020. Available from: www.training.cochrane.org/handbook.

19. GRADE Handbook 2013 [Available from: https://gdt.gradepro.org/app/handbook/handbook.html.

20. Ayyavoo A, Derraik JG, Hofman PL, Biggs J, Bloomfield FH, Cormack BE, et al. Severe hyperemesis gravidarum is associated with reduced insulin sensitivity in the offspring in childhood. The Journal of clinical endocrinology and metabolism. 2013;98(8):3263-8.

21. Depue RH, Pike MC, Henderson BE. Estrogen exposure during gestation and risk of testicular cancer. J Natl Cancer Inst. 1983;71(6):1151-5.

22. Fejzo M, Kam A, Laguna A, MacGibbon K, Mullin P. Analysis of neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum reveals increased reporting of autism spectrum disorder. Reprod Toxicol. 2019;84:59-64.

23. Fejzo MS, Magtira A, Schoenberg FP, Macgibbon K, Mullin PM. Neurodevelopmental delay in children exposed in utero to hyperemesis gravidarum. Eur J Obstet Gynecol Reprod Biol. 2015;189:79-84.

24. Getahun D, Fassett MJ, Jacobsen SJ, Xiang AH, Takhar HS, Wing DA, et al. Autism Spectrum Disorders in Children Exposed in Utero to Hyperemesis Gravidarum. Am J Perinatol. 2019.

25. Henderson BE, Benton B, Jing J, Yu MC, Pike MC. Risk factors for cancer of the testis in young men. Int J Cancer. 1979;23(5):598-602.

26. Koot MH, Grooten IJ, Sebert S, Koiranen M, Jarvelin MR, Kajantie E, et al. Hyperemesis gravidarum and cardiometabolic risk factors in adolescents: a follow-up of the Northern Finland Birth Cohort 1986. BJOG : an international journal of obstetrics and gynaecology. 2017;124:1107-14.

27. Koren G, Cohen R. Effect of hyperemesis gravidarum on child neurodevelopment. Australasian Medical Journal. 2018;11(10):492-6.

28. Mullin PM, Bray A, Schoenberg F, MacGibbon KW, Romero R, Goodwin TM, et al. Prenatal exposure to hyperemesis gravidarum linked to increased risk of psychological and behavioral disorders in adulthood. J Dev Orig Health Dis. 2011;2(4):200-4.

29. Petridou E, Roukas KI, Dessypris N, Aravantinos G, Bafaloukos D, Efraimidis A, et al. Baldness and other correlates of sex hormones in relation to testicular cancer. Int J Cancer. 1997;71(6):982-5.

30. Poeran-Bahadoer S, Jaddoe VWV, Gishti O, Grooten IJ, Franco OH, Hofman A, et al. Maternal vomiting during early pregnancy and cardiovascular risk factors at school age: the Generation R Study. Journal of Developmental Origins of Health and Disease. 2020;11(2):118-26.

31. Swerdlow AJ, Stiller CA, Wilson LM. Prenatal factors in the aetiology of testicular cancer: an epidemiological study of childhood testicular cancer deaths in Great Britain, 1953-73. J Epidemiol Community Health. 1982;36(2):96-101.

32. Syn NL, Chan SY, Chia EWY, Ong WX, Phua D, Cai S, et al. Severity of nausea and vomiting in pregnancy and early childhood neurobehavioural outcomes: The Growing Up in Singapore Towards Healthy Outcomes study. Paediatr Perinat Epidemiol. 2020;23:23.

33. Vandraas KF, Vikanes AV, Stoer NC, Troisi R, Stephansson O, Sorensen HT, et al. Hyperemesis gravidarum and risk of cancer in offspring, a Scandinavian registry-based nested case-control study. BMC Cancer. 2015;15:398.

34. Wang H, Rolls ET, Du X, Du J, Yang D, Li J, et al. Severe nausea and vomiting in pregnancy: psychiatric and cognitive problems and brain structure in children. BMC Med. 2020;18(1):228.

35. Hisle-Gorman E, Susi A, Stokes T, Gorman G, Erdie-Lalena C, Nylund CM. Prenatal, perinatal, and neonatal risk factors of autism spectrum disorder. Pediatr Res. 2018;84(2):190-8.

36. Gu L, Mo M, Si S, Luo W, Shao B, Xin X, et al. Association of nausea and vomiting of pregnancy with infant growth in the first 24 months of life. Arch Gynecol Obstet. 2021;304(2):429-38.

37. Ong J, Sadananthan SA, Soh SE, Ng S, Yuan WL, Aris IM, et al. Increasing nausea and vomiting of pregnancy is associated with sex-dependent differences in early childhood growth: the GUSTO mother-offspring cohort study. BMC Pregnancy Childbirth. 2021;21(1):578.

38. Koot MH, Boelig RC, van't Hooft J, Limpens J, Roseboom TJ, Painter RC, et al. Variation in hyperemesis gravidarum definition and outcome reporting in randomised clinical trials: a systematic review. BJOG: An International Journal of Obstetrics & Gynaecology. 2018;125(12):1514-21.

39. Vikanes Å, Magnus P, Vangen S, Lomsdal S, Grjibovski AM. Hyperemesis gravidarum in the Medical Birth Registry of Norway – a validity study. BMC Pregnancy and Childbirth. 2012;12(1):115.

40. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). Washington. 2013.

41. Hoek HW, Susser E, Buck KA, Lumey LH, Lin SP, Gorman JM. Schizoid personality disorder after prenatal exposure to famine. Am J Psychiatry. 1996;153(12):1637-9.

42. Neugebauer R, Hoek HW, Susser E. Prenatal Exposure to Wartime Famine and Development of Antisocial Personality Disorder in Early Adulthood. JAMA. 1999;282(5):455-62.

43. Reyes-Castro LA, Rodriguez JS, Charco R, Bautista CJ, Larrea F, Nathanielsz PW, et al. Maternal protein restriction in the rat during pregnancy and/or lactation alters cognitive and anxiety behaviors of female offspring. Int J Dev Neurosci. 2012;30(1):39-45.

44. Zhong C, Tessing J, Lee BK, Lyall K. Maternal Dietary Factors and the Risk of Autism Spectrum Disorders: A Systematic Review of Existing Evidence. Autism Res. 2020;13(10):1634-58.

45. Jenabi E, Bashirian S, Asali Z, Seyedi M. Association between small for gestational age and risk of autism spectrum disorders: a meta-analysis. Clin Exp Pediatr. 2021;64(10):538-42.

46. Bai D, Yip BHK, Windham GC, Sourander A, Francis R, Yoffe R, et al. Association of Genetic and Environmental Factors With Autism in a 5-Country Cohort. JAMA Psychiatry. 2019;76(10):1035-43.

47. Mullin PM, Bray A, Vu V, Schoenberg-Paik F, MacGibbon K, Romero R, et al. No increased risk of psychological/behavioral disorders in siblings of women with hyperemesis gravidarum (HG) unless their mother had HG. J Dev Orig Health Dis. 2012;3(5):375-9.

48. Jones AP, Friedman MI. Obesity and adipocyte abnormalities in offspring of rats undernourished during pregnancy. Science. 1982;215(4539):1518-9.

49. Edwards LJ, McMillen IC. Periconceptional nutrition programs development of the cardiovascular system in the fetal sheep. Am J Physiol Regul Integr Comp Physiol. 2002;283(3):R669-79.

50. Liu H, Chen X, Shi T, Qu G, Zhao T, Xuan K, et al. Association of famine exposure with the risk of type 2 diabetes: A meta-analysis. Clin Nutr. 2020;39(6):1717-23.

51. Hult M, Tornhammar P, Ueda P, Chima C, Bonamy AK, Ozumba B, et al. Hypertension, diabetes and overweight: looming legacies of the Biafran famine. PLoS One. 2010;5(10):e13582.

Appendix A1. Search strategy

Search in Medline (1946 to September 6th, 2021)

#	Searches	Results
1	hyperemesis gravidarum/	1686
2	((*morning sickness/ and (*vomiting/ or *nausea/)) or (*vomiting/ and *nausea/)) and pregnancy/	472
3	(hypereme* adj15 (pregnanc* or pregnant or gestat* or gravidi* or gravidar* or trimester* or maternal or prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or in-utero)).tw,kf.	1804
4	((pernicious* or serious* or sever* or excessiv*) adj2 (vomiting or nause*) adj9 (gravidar* or gravidit* or gestat* or pregnanc* or pregnant or trimester*)).tw,kf.	313
5	(nausea adj2 vomit* adj3 pregnan*).tw,kf.	877
6	**or/1-5 [hyperemesis gravidarum]**	**3066**
7	exp pregnancy outcome/ [incl stillbirth, live birth, spontaneous abortion]	78350
8	prenatal exposure delayed effects/ or maternal exposure/	38290
9	fetus/ or exp fetal heart/ or exp child/ or exp infant/ or puberty/ or schools/ or pediatrics/ [child /fetus]	2706830
10	child mortality/ or fetal mortality/ or exp infant mortality/	32885
11	embryo loss/ or fetal diseases/ or fetal macrosomia/ or fetal growth retardation/ or fetal hypoxia/ or fetal nutrition disorders/ or exp fetal death/ or exp fetal membranes, premature rupture/ or exp obstetric labor, premature/ or oligohydramnios/ or hydrops fetalis/ or perinatal death/ or placenta diseases/ or abruptio placentae/ or placental insufficiency/ or infant, newborn, diseases/ or asphyxia neonatorum/ or exp infant, premature, diseases/ or neonatal sepsis/ or jaundice neonatal/ or exp infant death/ [pregnancy complications/infant death]	210822
12	exp birth weight/ or fetal weight/ or cephalometry/ or crown-rump length/ or fetal distress/ or apgar score/	80637
13	child nutritional physiological phenomena/ or infant nutritional physiological phenomena/ or prenatal nutritional physiological phenomena/ or "growth and development"/ or exp human development/ or "embryonic and fetal development"/ or embryonic development/ or fetal development/ or fetal movement/ or fetal organ maturity/ or fetal viability/ or sex determination processes/ or sex differentiation/ or sexual development/ or language development/ or psychology, developmental/ or psychology, educational/ or exp education, special/ or exp child behavior/ or behavioral symptoms/ or neurobehavioral manifestations/	201303
14	adolescent health/ or child health/ or infant health/	5867
15	neurodevelopmental disorders/ or exp "attention deficit and disruptive behavior disorders"/ or exp autism spectrum disorder/ or obsessive-compulsive disorder/ or exp tic disorders/ or exp psychomotor performance/ or motor skills disorders/ or child behavior disorders/	223188
16	exp aptitude tests/ or behavior rating scale/ or neuropsychological tests/ or language tests/ or exp "memory and learning tests"/ or stroop test/ or trail making test/	124448
17	sex factors/ or sex ratio/ or sex determination analysis/	286865

Continued

#	Searches	Results
18	exp testicular diseases/	39645
19	exp musculoskeletal system/ab or exp heart/ab or exp nervous system/ab or genitalia/ab or abdominal wall/ab or urinary tract/ab or kidney/ab or urinary bladder/ab or kidney diseases/cn	87954
20	exp *congenital abnormalities/ or exp congenital abnormalities/et, ep	516968
21	abnormalities, severe teratoid/ or exp cardiovascular abnormalities/ or exp nervous system malformations/ or hydrocephalus/ or exp musculoskeletal abnormalities/ or exp bone diseases, developmental/ or cleft lip/ or exp digestive system abnormalities/ or exp respiratory system abnormalities/ or exp urogenital abnormalities/ or exp hydronephrosis/	513447
22	((pregnancy or gestat*) adj outcom*).tw,kf.	27889
23	((perinat* or peri-nat* or birth*1 or childbirth* or deliver* or labo?r* or obstetric*) adj3 outcome*).tw,kf.	34700
24	((perinat* or peri-nat*) adj3 (complicat* or health or morbidit* or cancer* or malignan* or neoplas*)).tw,kf.	9811
25	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj18 expos*).tw,kf.	29816
26	(maternal adj2 expos*).tw,kf.	7898
27	((prenat* or pre-nat* or antenat* or ante-nat* or in-utero or intra-uterine or intrauterine) adj3 (factor* or variabl* or parameter* or circumstanc* or condition*)).tw,kf.	5808
28	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj life).tw,kf.	1940
29	(DOHAD* or FOAD* or (early adj3 origin*)).tw,kf. or (development* adj3 origin* adj4 (health* or diseas* or adult)).tw,kf,jw.	5837
30	(offspring* or progeny or (born adj2 mother*)).tw,kf.	116107
31	(f?etal or f?etus* or neonat* or neo-nat* or new*born* or new-born* or child or child*1 or children* or schoolchild* or childhood or infant* or infanc* or toddler* or prekindergarten* or kindergarten* or preschool* or school-age* or schoolage* or high-school* or highschool* or elementary school* or graders or puber* or teens or teenager* or youth or juvenil* or adolescence or adulthood or young adult* or adult life or older age* or "early life" or later-life or "later in life").tw,kf. [child filter]	2682682
32	(((perinat* or peri-nat* or intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (mortalit* or death* or demise)) or stillbirth* or stillborn* or asphyx* or miscarriag* or IUFD or (spontan* adj3 abort*) or ((embry* or pregnancy) adj2 loss*) or liveborn* or (live adj3 (birth* or born*))).tw,kf.	108628
33	(((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (growth* or develop* or brain or movement*)) or ((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj12 growth adj2 (restrict* or retard*)) or FGR* or IUGR* or SFGR* or SIUGR*).tw,kf.	30590
34	(placent* adj3 (insufficien* or d*sfunct* or inflammat* or abruptio*)).tw,kf.	8432

Continued

#	Searches	Results
35	((PROM and ruptur* and (membran* or amnio*)) or PPROM* or EPPROM*1 or ((prematur* or pre-matur* or i?matur* or preterm* or pre-term* or pre-labo?r or prelabo?r) adj6 ruptur* adj4 (amnio* or membran*)) or chorioamn* or amnionit* or intraamnio* or funisit*).tw,kf. or (((ruptur* adj2 (amnio* or membran*)) or ROM).tw,kf. and (pregnan* or gestat* or gravidit* or trimester* or intrauterine* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*).mp.)	14299
36	(prematurity or ((preterm* or pre-term* or prematur* or pre-matur*) adj3 (labo?r or deliver* or birth* or childbirth*)) or PTB or PTBs or TPTB* or SPTB* or VPTB* or EPTB* or PTL or TPTL* or PTD*).tw,kf.	85358
37	(((small* or large* or deliver* or labo?r* or birth or childbirth) adj4 gestat* adj2 (age or ages)) or (gestat* adj ("at birth" or "at deliver*")) or birth age* or SGA or LGA).tw,kf.	34043
38	(((birth or births or childbirth* or born or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or peri-nat* or intrauterine or intra-uterine or in-utero) adj2 (underweight* or weight* or overweight* or siz* or length*)) or birthweight* or LBW* or VLBW* or ELBW*).tw,kf.	88258
39	((((head or cephal* or body or arm or arms or leg*) adj4 (circumfer* or measur* or siz* or small* or larg*)) or cephalometr* or anthropometr* or body mass* or BMI) adj9 (birth or births or childbirth* or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or per-nat* or intrauterine or intra-uterine or in-utero)).tw,kf.	10653
40	(neurodevelop* or ((brain or neurol* or neural or sex* or grow*) adj2 develop*) or "ages and stages" or DDST or (developmental adj4 (outcome* or test* or quotient* or index or indices or scor* or scale*)) or ((behav* or neurocognit* or cognit* or neurobehav* or neuropsychomot* or psychomotor* or psycho-motor* or neuromotor* or neuro-motor* or sensor?motor* or sensory motor* or visuomotor or visual motor or neuro-sensory or neurosensory) adj3 (abilit* or outcom* or problem* or develop*)) or ((executive or motor) adj3 (function* or d#sfunc* or deficit* or problem*)) or interference control or psychointellect* or intellect* or intelligen* or IQ or DQ or psycholinguist* or linguist*).tw,kf.	483193
41	(((language or learning or speech or reading or memory) adj3 (skill* or test* or scale* or scor* or task* or cognitiv*)) or verbal skill* or wording or naming or (numeric* adj3 memory) or 5-digit or digit-span or letter-digit).tw,kf.	114783
42	(Binet* or Wechsler* or WAIS or WASI* or WIAT* or WDD or WDR or WRR or WISC or (WISC* not Wisconsin*) or WPPSI* or WRIT or CANTAB or complex figure or RCF or RCOF or RCFOSS or GIT-2 or FSIQ or VIQ or PIQ or (assess* adj2 batter*) or ((mobile or assesment*) adj3 ABC) or m-ABC or mABC or kABC or CELF or (mental development adj3 (index or indic* or scor*)) or MDI or WJ-R or Ba?ley* or BSID* or NEPSY or Beery or Basic Concept Scale or BBCS* or CMS or continuous performance* or Serial Addition or PDI or RBMT or Stroop or CNT or PPVT* or Everyday Attention or TOMI or MPC or CNT or Trail Making or Brunet or LMT or LMTs or TMT or TMTs or TMTa or Achievement Test or WJ-SAT or WJ IV or WRAT* or sensory profil* or ITSP or SSP or SPNL or CBCL).tw,kf.	114014
43	(((female or male or women or men or males or ratio* or distribut* or proportion* or factor*) adj3 (sex or gender*)) or ((male or males) adj1 female*) or girls or boys or ((deliver* or parturit* or birth or born*) adj2 (girl* or boy* or femal* or male or males))).tw,kf.	317406

Continued

#	Searches	Results
44	((testi* adj3 (cancer* or neoplas* or malignan* or tumo?r* or undescen* or descen*)) or cryptorch*).tw,kf.	35203
45	(congenit* or anomal* or malformat* or deformit* or d#smorph* or aplas* or d#splas* or hypoplas* or atres* or agenes*).tw,kf.	653144
46	((prenat* or pre-nat* or antenat* or ante-nat* or perinat or peri-nat* or birth or anatomic* or morphological* or isolated or chromosom* or nonchromosom* or cardiac or noncardiac or extracard* or cardio* or heart or outflow tract or OFT or conotrunc* or cono-trunc* or septal or septum or endocard* cushion* or atrioventric* or atrio-ventr* or AV or musc*skelet* or skelet* or bone or bones or osseous or spine or spinal or limb or limbs or extremit* or foot or feet or hand or hands or cranio* or orofacial or facial* or palat*2 or mouth or lip or lips or (digest* adj2 (system or tract*)) or GI or intestin* or duoden* or esophag* or oesophag* or trach*esophag* or abdominal or respirator* or pulmonar* or lung or diaphragm* or hemidiaphragm* or sex or sexual or genit* or urogenit* or kidney* or uret* or renal or bladder or neural-tube* or nervous system or CNS or brain) adj3 (abnormalit* or defect*1)).tw,kf.	201082
47	(Down* syndrome or CHD or Fallot* or Ebstein* or coarct* or (aort* adj1 arch*) or double outlet* or DORV or HLHS or HLV or HRHS or univentricular or uni-ventricular or single ventricle* or ((common arterial or arterios*) adj2 (trunk or truncus)) or VSD or (common adj3 (septum or septal) adj3 canal*) or AVSD or CAVC or scimitar* or TAPVC* or PAPVC* or encephaloc?el* or cephaloc?el* or meningoencephaloc?el* or notoencephaloc?ele or craniac?el* or ((cereb* or mening*) adj2 hernia*) or anencephal* or acrani* or aprosencephal* or ((spin* or cranium or crania) adj2 (bifid* or open)) or d#sraphi* or rachischis* or crani*-schis* or crani*schis* or myeloc?ele* or mening*myeloc?el* or hydrocephal* or hydro-cephal* or ventricul*-megal* or ventricul*megal* or holoprosencephal* or holo-prosencephal* or arhinencephal* or achondroplasi* or thanatophor* or osteochondrod#splas* or osteod#splast* or osteod#d#stroph* or chondrod#splas* or chondrod#stroph* or ((limb or limbs) adj2 reduct*) or talipes or clubfoot or club-foot or cleft or clefts or gastro?chis* or gastro-schis* or (umbilic* adj2 hernia*) or omphaloc?el* or exomphal* or ((cystic or polycyst* or multicyst*) adj2 (kidney* or (renal adj2 (diseas* or disorder*)))) or PKD or MCKD or megacystis or hydronephro* or Smith-Lemli-Opitz or micromelia or ectromeli* or (hydrops adj3 f?etalis) or hypospad* or hip dislocat*).tw,kf.	250725
48	**or/7-47 [perinatal and long term offspring outcomes]**	**5716019**
49	**6 and 48 [HG + perinatal and long term offspring outcome]**	**1191**
50	(editorial or "systematic review").pt. or (editorial or reply or (case-report not case-report-survey) or two-cases or three-cases or four-cases or five-cases or 2-cases or 3-cases or 4-cases or 5-cases).ti. or cochrane.jw. or ((review.pt. or case reports/ or case report*.jw. or (review or overview).ti. or (search* adj15 (literatur* or ((electronic* or medical or biomedical) adj3 database*) or medline or pubmed or embase or psyc?info or exhaustiv* or systematic*)).tw,kf,kw.) not (cohort studies/ or longitudinal studies/ or prospective studies/ or retrospective studies/ or cross-sectional studies/ or case-control studies/ or (case-control* or cohort* or retrospectiv* or prospectiv* or crosssection* or cross-section* or population-based or ((chart* or record* or retrospectiv*) adj3 review*)).tw,kf,kw.)) or (exp animals/ not exp humans/) or animal.jw. or (rodent* or rabbit* or mice or mouse or murine or rat or rats or (animal* adj3 (experiment* or model))).ti. [filter for original human studies]	10692231

Continued

#	Searches	Results
51	**49 not 50 [HG + perinatal and long term offspring outcome - original human studies]**	**701**
52	**remove duplicates from 51 [HG + perinatal and long term offspring outcome -original human studies - duplicates removed]**	**700**

Search in Embase Classic and Embase (1947 to September 6th, 2021)

#	Searches	Results
1	hyperemesis gravidarum/	3255
2	((*"nausea and vomiting"/ or hyperemesis.dj. or (*vomiting/ and *nausea/)) and (pregnancy/ or pregnancy complication/ or prenatal period/ or prenatal exposure/)) or ((*vomiting/ or *nausea/) and (prenatal exposure/ or prenatal period/))	828
3	(hypereme* adj15 (pregnanc* or pregnant or gestat* or gravidi* or gravidar* or trimester* or maternal or prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or in-utero)).tw,kw.	2498
4	((pernicious* or serious* or sever* or excessiv*) adj2 (vomiting or nause*) adj9 (gravidar* or gravidit* or gestat* or pregnanc* or pregnant or trimester*)).tw,kw.	424
5	(nausea adj2 vomit* adj3 pregnan*).tw,kw.	1162
6	**or/1-5 [HG]**	**4971**
7	apgar score/ or exp birth weight/ or crown rump length/ or cephalometry/ or exp fetus maturity/ or fetus outcome/ or fetus weight/ or live birth/ or exp perinatal morbidity/ or placenta weight/ or pregnancy outcome/ ["parameters concerning the fetus, newborn and pregnancy"]	273732
8	maternal exposure/ or prenatal exposure/	28799
9	fetus/ or fetus brain/ or fetus heart/ or child/ or juvenile/ or exp infant/ or preschool child/ or school child/ or toddler/ or adolescence/ or adulthood/ or exp childhood/ or high school/ or kindergarten/ or middle school/ or primary school/ or pediatrics/ or progeny/ [fetus/child]	3547791
10	childhood mortality/ or embryo mortality/ or fetus mortality/ or infant mortality/ or exp perinatal mortality/ or prenatal mortality/ [child/fetus mortality]	69986
11	exp "immature and premature labor"/ or fetus disease/ or prenatal disorder/ or chorioamnionitis/ or dysmaturity/ or fetal malnutrition/ or fetus distress/ or fetus hypoxia/ or fetus malformation/ or exp hydramnios/ or exp intrauterine growth retardation/ or macrosomia/ or exp oligohydramnios/ or premature fetus membrane rupture/ or fetus wastage/ or spontaneous abortion/ or exp child death/ or embryo death/ or exp fetus death/ or perinatal death/ or placenta disorder/ or placenta insufficiency/ or solutio placentae/ or infant disease/ or newborn disease/ or dysmaturity/ or immaturity/ or large for gestational age/ or neonatal respiratory distress syndrome/ or neonatal stress/ or newborn apnea/ or newborn hypoxia/ or newborn infection/ or newborn sepsis/ or newborn vomiting/ or perinatal asphyxia/ or perinatal stress/ or prematurity/ or retrolental fibroplasia/ or newborn assessment/ or newborn intensive care/ or lung dysplasia/ or encephalomalacia/	453074

Continued

#	Searches	Results
12	"growth, development and aging"/ or nerve cell differentiation/ or human development/ or adolescent development/ or language development/ or psychosocial development/ or speech development/ or exp postnatal development/ or prenatal development/ or embryo development/ or fetus development/ or fetal well being/ or fetus lung maturation/ or exp postnatal development/ or special education/ or exp sexual development/ or child behavior/ or neurobehavioral manifestations/ or neurodevelopment* outcome*.dq.	706313
13	adolescent health/ or child health/	39074
14	behavior disorder/ or attention deficit disorder/ or congenital behavior disorder/ or exp disruptive behavior/ or oppositional defiant disorder/ or exp autism/ or exp tic/ or obsessive compulsive disorder/ or exp learning disorder/ or exp "disorders of higher cerebral function"/ or attention disturbance/ or developmental coordination disorder/ or exp intellectual impairment/ or language disability/ or exp developmental language disorder/ or developmental disorder/ or developmental delay/	1113944
15	behavior assessment/ or aptitude test/ or learning test/ or exp neuropsychological test/ or exp cognition assessment/ or developmental screening/	144565
16	"gender and sex"/ or gender/ or sex/ or sex difference/ or sex factor/ or sex ratio/ or exp sex determination/	833481
17	exp testis disease/	79045
18	exp *congenital disorder/ or exp congenital disorder/et, ep	973317
19	congenital malformation/ or exp "head and neck malformation"/ or exp limb malformation/ or exp cardiovascular malformation/ or severe teratoid abnormality/ or exp nervous system malformation/ or exp digestive system malformation/ or exp male genital tract malformation/ or exp musculoskeletal system malformation/ or hydronephrosis/	797608
20	((pregnancy or gestat*) adj outcom*).tw,kw.	41438
21	((perinat* or peri-nat* or birth*1 or childbirth* or deliver* or labo?r* or obstetric*) adj3 outcome*).tw,kw.	52218
22	((perinat* or peri-nat*) adj3 (complicat* or health or morbidit* or cancer* or malignan* or neoplas*)).tw,kw.	14447
23	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj18 expos*).tw,kw.	39656
24	(maternal adj2 expos*).tw,kw.	10434
25	((prenat* or pre-nat* or antenat* or ante-nat* or in-utero or intra-uterine or intrauterine) adj3 (factor* or variabl* or parameter* or circumstanc* or condition*)).tw,kw.	8096
26	((prenat* or pre-nat* or antenat* or ante-nat* or intrauterine or intra-uterine or utero) adj life).tw,kw.	3478
27	(DOHAD* or FOAD* or (early adj3 origin*)).tw,kw. or (development* adj3 origin* adj4 (health* or diseas* or adult)).tw,kw,jw.	6960
28	(offspring* or progeny or (born adj2 mother*)).tw,kw.	141617
29	(f?etal or f?etus* or neonat* or neo-nat* or new*born* or new-born* or child or child*1 or children* or schoolchild* or childhood or infant* or infanc* or toddler* or prekindergarten* or kindergarten* or preschool* or school-age* or schoolage* or high-school* or highschool* or elementary school* or graders or puber* or teens or teenager* or youth or juvenil* or adolescence or adulthood or young adult* or adult life or older age* or "early life" or later-life or "later in life").tw,kw. [child filter]	3582830

Continued

#	Searches	Results
30	(((perinat* or peri-nat* or intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (mortalit* or death* or demise)) or stillbirth* or stillborn* or asphyx* or miscarriag* or IUFD or (spontan* adj3 abort*) or ((embry* or pregnancy) adj2 loss*) or liveborn* or (live adj3 (birth* or born*))).tw,kw.	162175
31	(((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj3 (growth* or develop* or brain or movement*)) or ((intrauterin* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*) adj12 growth adj2 (restrict* or retard*)) or FGR* or IUGR* or SFGR* or SIUGR*).tw,kw.	45264
32	(placent* adj3 (insufficien* or d*sfunct* or inflammat* or abruptio*)).tw,kw.	13730
33	((PROM and ruptur* and (membran* or amnio*)) or PPROM* or EPPROM*1 or ((prematur* or pre-matur* or i?matur* or preterm* or pre-term* or pre-labo?r or prelabo?r) adj6 ruptur* adj4 (amnio* or membran*)) or chorioamn* or amnionit* or intraamnio* or funisit*).tw,kw. or (((ruptur* adj2 (amnio* or membran*)) or ROM).tw,kw. and (pregnan* or gestat* or gravidit* or trimester* or intrauterine* or intra-uterin* or in-utero or prenat* or pre-nat* or antenat* or ante-nat*).mp.)	21576
34	(prematurity or ((preterm* or pre-term* or prematur* or pre-matur*) adj3 (labo?r or deliver* or birth* or childbirth*)) or PTB or PTBs or TPTB* or SPTB* or VPTB* or EPTB* or PTL or TPTL* or PTD*).tw,kw.	126669
35	(((small* or large* or deliver* or labo?r* or birth or childbirth) adj4 gestat* adj2 (age or ages)) or (gestat* adj ("at birth" or "at deliver*")) or birth age* or SGA or LGA).tw,kw.	51397
36	(((birth or births or childbirth* or born or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or peri-nat* or intrauterine or intra-uterine or in-utero) adj2 (underweight* or weight* or overweight* or siz* or length*)) or birthweight* or LBW* or VLBW* or ELBW*).tw,kw.	123200
37	((((head or cephal* or body or arm or arms or leg*) adj4 (circumfer* or measur* or siz* or small* or larg*)) or cephalometr* or anthropometr* or body mass* or BMI) adj9 (birth or births or childbirth* or parturit* or delivery or baby or babies or postnat* or post-nat* or perinat* or per-nat* or intrauterine or intra-uterine or in-utero)).tw,kw.	16456
38	(neurodevelop* or ((brain or neurol* or neural or sex* or grow*) adj2 develop*) or "ages and stages" or DDST or (developmental adj4 (outcome* or test* or quotient* or index or indices or scor* or scale*)) or ((behav* or neurocognit* or cognit* or neurobehav* or neuropsychomot* or psychomotor* or psycho-motor* or neuromotor* or neuro-motor* or sensor?motor* or sensory motor* or visuomotor or visual motor or neuro-sensory or neurosensory) adj3 (abilit* or outcom* or problem* or develop*)) or ((executive or motor) adj3 (function* or d#sfunc* or deficit* or problem*)) or interference control or psychointellect* or intellect* or intelligen* or IQ or DQ or psycholinguist* or linguist*).tw,kw.	645254
39	(((language or learning or speech or reading or memory) adj3 (skill* or test* or scale* or scor* or task* or cognitiv*)) or verbal skill* or wording or naming or (numeric* adj3 memory) or 5-digit or digit-span or letter-digit).tw,kw.	149827

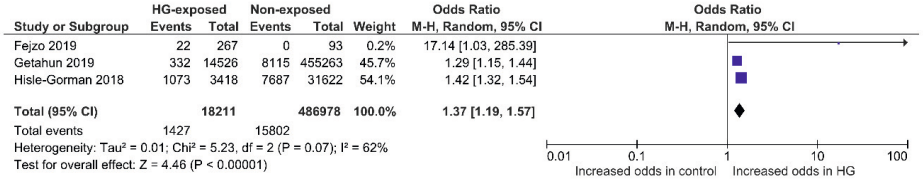
Continued

#	Searches	Results
40	(Binet* or Wechsler* or WAIS or WASI* or WIAT* or WDD or WDR or WRR or WISC or (WISC* not Wisconsin*) or WPPSI* or WRIT or CANTAB or complex figure or RCF or RCOF or RCFOSS or GIT-2 or FSIQ or VIQ or PIQ or (assess* adj2 batter*) or ((mobile or assesment*) adj3 ABC) or m-ABC or mABC or kABC or CELF or (mental development adj3 (index or indic* or scor*)) or MDI or WJ-R or Ba?ley* or BSID* or NEPSY or Beery or Basic Concept Scale or BBCS* or CMS or continuous performance* or Serial Addition or PDI or RBMT or Stroop or CNT or PPVT* or Everyday Attention or TOMI or MPC or CNT or Trail Making or Brunet or LMT or LMTs or TMT or TMTs or TMTa or Achievement Test or WJ-SAT or WJ IV or WRAT* or sensory profil* or ITSP or SSP or SPNL or CBCL).tw,kw.	154869
41	(((female or male or women or men or males or ratio* or distribut* or proportion* or factor*) adj3 (sex or gender*)) or ((male or males) adj1 female*) or girls or boys or ((deliver* or parturit* or birth or born*) adj2 (girl* or boy* or femal* or male or males))).tw,kw.	476728
42	((testi* adj3 (cancer* or neoplas* or malignan* or tumo?r* or undescen* or descen*)) or cryptorch*).tw,kw.	50422
43	(congenit* or anomal* or malformat* or deformit* or d#smorph* or aplas* or d#splas* or hypoplas* or atres* or agenes*).tw,kw.	918817
44	((prenat* or pre-nat* or antenat* or ante-nat* or perinat or peri-nat* or birth or anatomic* or morphological* or isolated or chromosom* or nonchromosom* or cardiac or noncardiac or extracard* or cardio* or heart or outflow tract or OFT or conotrunc* or cono-trunc* or septal or septum or endocard* cushion* or atrioventric* or atrio-ventr* or AV or musc*skelet* or skelet* or bone or bones or osseous or spine or spinal or limb or limbs or extremit* or foot or feet or hand or hands or cranio* or orofacial or facial* or palat*2 or mouth or lip or lips or (digest* adj2 (system or tract*)) or GI or intestin* or duoden* or esophag* or oesophag* or trach*esophag* or abdominal or respirator* or pulmonar* or lung or diaphragm* or hemidiaphragm* or sex or sexual or genit* or urogenit* or kidney* or uret* or renal or bladder or neural-tube* or nervous system or CNS or brain) adj3 (abnormalit* or defect*1)).tw,kw.	273256
45	(Down* syndrome or CHD or Fallot* or Ebstein* or coarct* or (aort* adj1 arch*) or double outlet* or DORV or HLHS or HLV or HRHS or univentricular or uni-ventricular or single ventricle* or ((common arterial or arterios*) adj2 (trunk or truncus)) or VSD or (common adj3 (septum or septal) adj3 canal*) or AVSD or CAVC or scimitar* or TAPVC* or PAPVC* or encephaloc?el* or cephaloc?el* or meningoencephaloc?el* or notoencephaloc?ele or craniac?el* or ((cereb* or mening*) adj2 hernia*) or anencephal* or acrani* or aprosencephal* or ((spin* or cranium or crania) adj2 (bifid* or open)) or d#sraphi* or rachischis* or crani*-schis* or crani*schis* or myeloc?ele* or mening*myeloc?el* or hydrocephal* or hydro-cephal* or ventricul*-megal* or ventricul*megal* or holoprosencephal* or holo-prosencephal* or arhinencephal* or achondroplasi* or thanatophor* or osteochondrod#splas* or osteod#splast* or osteod#d#stroph* or chondrod#splas* or chondrod#stroph* or ((limb or limbs) adj2 reduct*) or talipes or clubfoot or club-foot or cleft or clefts or gastro?chis* or gastro-schis* or (umbilic* adj2 hernia*) or omphaloc?el* or exomphal* or ((cystic or polycyst* or multicyst*) adj2 (kidney* or (renal adj2 (diseas* or disorder*)))) or PKD or MCKD or megacystis or hydronephro* or Smith-Lemli-Opitz or micromelia or ectromeli* or (hydrops adj3 f?etalis) or hypospad* or hip dislocat*).tw,kw.	352321
46	**or/7-45 [perinatal and long term offspring outcomes]**	**8528197**

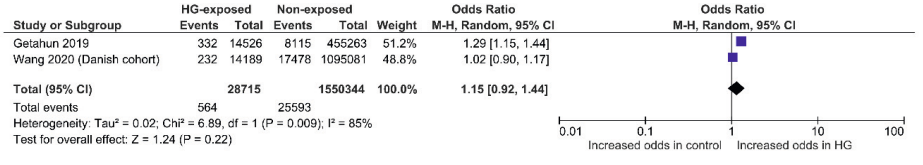
Continued

#	Searches	Results
47	**6 and 46 [HG + perinatal and long term offspring outcome]**	**2586**
48	editorial/ or "systematic review"/ or (editorial or conference abstract or conference review).pt. or (editorial or reply or (case-report not case-report-survey) or two-cases or three-cases or four-cases or five-cases or 2-cases or 3-cases or 4-cases or 5-cases).ti. or cochrane.jw. or ((review.pt. or review/ or case report/ or case report*.jw. or (review or overview).ti. or (search* adj15 (literatur* or ((electronic* or medical or biomedical) adj3 database*) or medline or pubmed or embase or psyc?info or exhaustiv* or systematic*)).tw,kw.) not (cohort analysis/ or longitudinal study/ or prospective study/ or retrospective study/ or exp case control study/ or cross-sectional study/ or (case-control* or cohort* or retrospectiv* or prospectiv* or crosssection* or cross-section* or population-based or ((chart* or record* or retrospectiv*) adj3 review*)).tw,kw.)) or ((exp animal/ or animal experiment/ or exp animal model/ or nonhuman/ or exp female animal/) not human/) or exp veterinary medicine/ or animal*.jw. or (rodent* or rabbit* or mice or mouse or murine or rat or rats or (animal* adj3 (experiment* or model))).ti. [filter for original human studies]	17746737
49	**47 not 48 [HG + perinatal and long term offspring outcome - original human studies]**	**1248**
50	**Remove duplicates from 49 [HG + perinatal and long term offspring outcome -original human studies - duplicates removed]**	**1231**
51	**50 not medline.cr. [HG + perinatal and long term offspring outcome - original human studies - duplicates removed - embase records only]**	**1046**

Supplement Figure S1. Sensitivity analysis – meta-analysis including studies from the USA



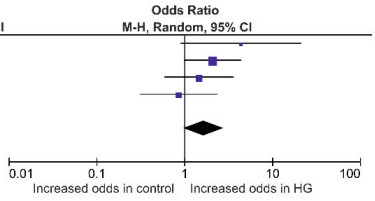
Supplement Figure S2. Sensitivity analysis – meta-analysis including cohort studies



Supplement Figure S3. Sensitivity analysis – meta-analysis testicular cancer (minus study Depue)

Study or Subgroup	HG-exposed Events	Total	Non-exposed Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Henderson 1979	8	10	70	146	10.4%	4.34 [0.89, 21.15]
Petridou 1997	15	31	82	264	38.4%	2.08 [0.98, 4.41]
Swerdlow 1982	5	413	82	9802	28.2%	1.45 [0.59, 3.60]
Vandraas 2015	4	55	843	10033	23.1%	0.86 [0.31, 2.37]
Total (95% CI)		**509**		**20245**	**100.0%**	**1.65 [0.98, 2.80]**
Total events	32		1077			

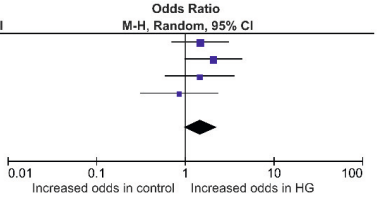
Heterogeneity: Tau² = 0.04; Chi² = 3.47, df = 3 (P = 0.32); I² = 14%
Test for overall effect: Z = 1.87 (P = 0.06)



Supplement Figure S4. Sensitivity analysis – meta-analysis testicular cancer (minus study Henderson)

Study or Subgroup	HG-exposed Events	Total	Non-exposed Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Depue 1983	20	35	67	141	31.3%	1.47 [0.70, 3.11]
Petridou 1997	15	31	82	264	30.9%	2.08 [0.98, 4.41]
Swerdlow 1982	5	413	82	9802	21.1%	1.45 [0.59, 3.60]
Vandraas 2015	4	55	843	10033	16.7%	0.86 [0.31, 2.37]
Total (95% CI)		**534**		**20240**	**100.0%**	**1.49 [0.98, 2.26]**
Total events	44		1074			

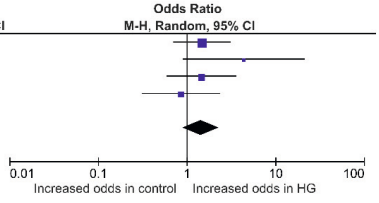
Heterogeneity: Tau² = 0.00; Chi² = 1.91, df = 3 (P = 0.59); I² = 0%
Test for overall effect: Z = 1.88 (P = 0.06)



Supplement Figure S5. Sensitivity analysis – meta-analysis testicular cancer (minus study Petridou)

Study or Subgroup	HG-exposed Events	Total	Non-exposed Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Depue 1983	20	35	67	141	41.1%	1.47 [0.70, 3.11]
Henderson 1979	8	10	70	146	9.1%	4.34 [0.89, 21.15]
Swerdlow 1982	5	413	82	9802	27.8%	1.45 [0.59, 3.60]
Vandraas 2015	4	55	843	10033	22.0%	0.86 [0.31, 2.37]
Total (95% CI)		**513**		**20122**	**100.0%**	**1.44 [0.89, 2.32]**
Total events	37		1062			

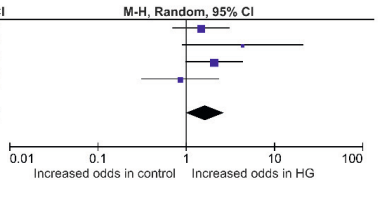
Heterogeneity: Tau² = 0.00; Chi² = 2.88, df = 3 (P = 0.41); I² = 0%
Test for overall effect: Z = 1.48 (P = 0.14)



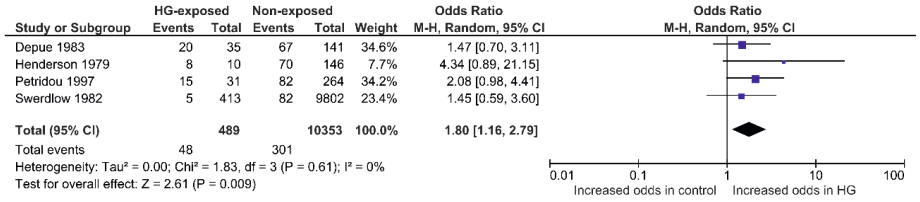
Supplement Figure S6. Sensitivity analysis – meta-analysis testicular cancer (minus study Swerdlow)

Study or Subgroup	HG-exposed Events	Total	Non-exposed Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Depue 1983	20	35	67	141	35.1%	1.47 [0.70, 3.11]
Henderson 1979	8	10	70	146	9.3%	4.34 [0.89, 21.15]
Petridou 1997	15	31	82	264	34.8%	2.08 [0.98, 4.41]
Vandraas 2015	4	55	843	10033	20.8%	0.86 [0.31, 2.37]
Total (95% CI)		**131**		**10584**	**100.0%**	**1.64 [1.00, 2.69]**
Total events	47		1062			

Heterogeneity: Tau² = 0.04; Chi² = 3.49, df = 3 (P = 0.32); I² = 14%
Test for overall effect: Z = 1.96 (P = 0.05)



Supplement Figure S7. Sensitivity analysis – meta-analysis testicular cancer (minus study Vandraas)



Supplement Table S1. Overview of results of included studies reporting on anthropometry and cardiometabolic health

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Adjusted results (β (95% CI))
		Anthropometry			
Height					
Ayyavoo (SD)	36/78	0.61 (0.31-0.91)*	0.36 (0.04-0.68)*	0.09	-
Gu – females (SD)	104/893	-	-	**<0.05**	-
Gu – males (SD)	128/1049	-	-	NS	-
Koot (cm)	42/6,462	169.4 ± 10.4	169.3 ± 8.4	0.99	-0.22 (-2.06;1.63)
Ong – females (SD)	112/563	-	-	NS	-
Ong – males (SD)	78/609	-	-	**<0.05**	0.64 (0.23;1.04)
Weight					
Gu – females (SD)	104/893	-	-	**<0.05**	-
Gu – males (SD)	128/1049	-	-	NS	-
Ong – females (SD)	112/563	-	-	**<0.05**	-0.53 (-1.03;-0.03)
Ong – males (SD)	78/609	-	-	**<0.05**	0.57 (0.05;1.08)
Poeran (g)	462/4,760	9,838 ± 1712	9,581 ± 1440	**<0.05**	Not performed
BMI					
Ayyavoo (SD)	36/78	0.22 (-0.19 to 0.64)*	0.40 (-0.01 to 0.82)*	0.47	-
Koot (kg/m^2)	42/6,462	20.5 (19.4-24.4)	20.5 (18.9-22.6)	0.36	2.5 (-2.0;7.3)
Ong – females (SD)	112/563	-	-	**<0.05**	-0.57 (-1.09;-0.05)
Ong – males (SD)	78/609	-	-	NS	
Poeran (kg/m^2)	462/4,760	16.8 ± 2.4	16.2 ± 1.8	**<0.05**	**0.08 (0.00;0.17)**
Absolute fat mass †					
Ayyavoo (%)	36/78	21.9 (18.8-25.0)*	21.5 (18.3-24.6)*	0.82	-
Poeran (%)	450/4,627	0.27 ± 0.06	0.25 ± 0.06	**<0.05**	**0.12 (0.03;0.20)**

Supplement Table S1. Overview of results of included studies reporting on anthropometry and cardiometabolic health. *Continued*

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Adjusted results (β (95% CI))
Android/gynoid fat mass ratio					
Ayyavoo (%)	36/78	0.60 (0.53-0.68)*	0.64 (0.56-0.72)*	0.48	-
Poeran (%)	450/4,627	0.27 ± 0.08	0.25 ± 0.06	**<0.05**	**0.11 (0.02;0.21)**
Waist/hip ratio					
Koot	42/6,462	0.80 ± 0.1	0.80 ± 0.1	0.60	0.00 (-0.01;0.01)
Preperitoneal fat mass area ‡					
Poeran (cm^2)	450/4,627	0.54 (0.51-0.58)	0.45 (0.44-0.46)	**<0.05**	**0.10 (0.00;0.20)**
Blood pressure					
Systolic blood pressure					
Koot (mmHg)	42/6,462	118 ± 14	116 ± 13	0.26	1.44 (-1.89;4.78)
Poeran (mmHg)	433/4,370	103.8 ± 8.8	102.4 ± 8.1	**<0.05**	0.02 (-0.08;0.12)
Diastolic blood pressure					
Koot (mmHg)	42/6,462	69 ± 7	68 ± 8	0.39	0.68 (-1.62;2.97)
Poeran (mmHg)	433/4,370	61.4 ± 7.3	60.5 ± 6.7	**<0.05**	0.03 (-0.08;0.13)
Laboratory measures					
Apolipoprotein A1 and B					
Koot - Apo A1 (mmol/L)	36/5,612	1.38 ± 0.2	1.36 ± 0.2	0.57	0.03 (-0.03;0.09)
Koot - Apo B (mmol/L)	36/5,612	0.71 ± 0.2	0.67 ± 0.2	0.11	0.04 (-0.01;0.10)
HDL					
Ayyavoo (mmol/L)	36/78	1.27 (1.15-1.40)*	1.35 (1.21-1.48)*	0.35	-
Koot (mmol/L)	36/5,612	1.40 ± 0.2	1.41 ± 0.3	0.79	0.01 (-0.10;0.10)
Poeran (mmol/L)	297/3,157	1.4 ± 0.3	1.3 ± 0.3	0.13	Not performed
LDL					
Ayyavoo (mmol/L)	36/78	2.34 (2.08-2.60)*	2.15 (1.87-2.43)*	0.25	-
Koot (mmol/L)	36/5,612	2.38 ± 0.5	2.25 ± 0.6	0.17	0.12 (-0.06;0.30)
Poeran (mmol/L)	298/3,154	2.4 ± 0.5	2.4 ± 0.6	0.61	Not performed
Total cholesterol					
Ayyavoo (mmol/L)	36/78	3.90 (3.62-4.18)*	3.89 (3.60-4.19)*	0.96	-
Koot (mmol/L)	36/5,612	4.38 ± 0.7	4.26 ± 0.8	0.37	0.12 (-0.06;0.30)
Poeran (mmol/L)	298/3,152	4.2 ± 0.6	4.2 ± 0.6	0.69	-0.04 (-0.17;0.09)
Triglycerides					
Koot (mmol/L)	36/5,612	0.71 (0.6-1.0)	0.70 (0.6-0.9)	0.80	-2.3 (-12.9;16.2)
Poeran (mmol/L)	295/3,143	1.1 (1.0-1.1)	1.1 (1.0-1.1)	0.71	0.02 (-0.11;0.15)

Supplement Table S1. Overview of results of included studies reporting on anthropometry and cardiometabolic health. *Continued*

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Adjusted results (β (95% CI))
Insulin					
Ayyavoo (fasting; mIU/L)	36/78	6.88 (5.56-8.50)*	5.04 (4.04-6.28)*	**0.02**	-
Koot (fasting; mU/L)	36/5,612	9.40 (7.0-12.1)	8.80 (6.9-11.4)	0.28	7.1 (-5.2;21.2)
Poeran (pmol/L)	293/3,126	146.5 (134.1-158.9)	137.9 (134.2-141.5)	0.16	0.09 (-0.04;0.21)
Glucose					
Ayyavoo (fasting; mg/dL)	36/78	4.76 (4.63-5.88)*	4.71 (4.58-4.84)*	0.55	-
Koot (mmol/L)	36/5,612	5.30 (5.1-5.5)	5.20 (4.9-5.5)	0.06	2.3 (-0.6;5.3)
C-peptide					
Poeran (nmol/L)	296/3,135	1.0 (1.0-1.1)	1.0 (1.0-1.1)	0.93	Not performed
HOMA-IR					
Koot	36/5,612	2.28 (1.7-3.2)	2.06 (1.6-2.7)	0.24	13.0 (-0.6;31.0)
Insulin sensitivity §					
Ayyavoo	36/78	8.49×10^{-4}min^{-1}*	10.60×10^{-4}min^{-1}*	**0.01**	-
Other (all Ayyavoo)					
-IGF-I (ng/mL)	36/78	181 (156-207)*	183 (157-209)*	0.88	-
-IGF-II (ng/mL)	36/78	651 (610-693)*	668 (624-711)*	0.54	-
-IGFBP-1 (ng/mL)	36/78	11.8 (7.9-15.6)*	19.0 (15.1-22.8)*	**<0.01**	-
-IGFBP-3 (ng/mL)	36/78	2955 (2657-3254)*	3435 (3122-3749)*	**0.01**	-
-Baseline Cortisol (nmol/L)	36/78	256 (224-292)*	210 (184-241)*	**0.02**	-

Statistically significant results are marked in **bold** (*P*<0.05). *Presented as adjusted means with 95% CI, due to missing descriptive statistics. † DEXA derived. ‡ Derived by abdominal ultrasound. § Assessed by a 90 minute frequent sampling iv glucose tolerance test. **Abbreviations:** β: beta regression coefficient. 95% CI: 95% confidence interval. HDL: high density lipoprotein. IGF (BP): Insulin-like growth factor (binding protein). LDL: low density lipoprotein. NS: not significant.

Supplement Table S2. Overview of results of included studies reporting on cognitive and motor development

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Other results (mean difference (95% CI); OR (95% CI); HR (95% CI))
Cognitive scores					
Wang – USA cohort *(NIH toolbox)*	1,496/9,214	83.7 ± 9.12	86.9 ± 8.94	**<0.001**	-
Syn *(Bayley-III)*	69/482	100.2 (96.5;103.9)†	105.6 (102.6;108.5)†	-	**MD -5.4 (-10.1;-0.6)***
Total IQ score					
Syn *(KBIT)*	65/469	92.5 (88.2;96.7)†	93.5 (90.3;96.8)†	-	MD -1.0 (-6.4;4.4)*
Koren *(WISC-R)*	22/241	108.7	114.2	**0.05**	-
Learning disorders					
Fejzo 2015	203/292‡	12.3%	3.4%	**0.03**	**OR 4.03 (1.36;17.24)**
Mullin	87/259	3 (3.4%)	3 (1.7%)	-	-
Speech/language problems					
Fejzo 2015 *(speech/ language)*	203/292‡	24.1%	11.2%	**0.02**	**OR 2.51 (1.43;31.83)**
Mullin *(speech)*	87/259	1 (1.1%)	0 (0%)	-	-
Syn *(Bayley-III, language)*	69/482	95.7 (91.6;99.9)†	97.2 (93.8;100.5)†	-	MD -1.4 (-6.7;3.9)*
Sensory disorder					
Fejzo 2015	203/292‡	19.7%	9.0%	**0.04**	**OR 2.51 (1.43;31.83)**
Fejzo 2019	267/360	20.2%	8.6%	**0.01**	**OR 2.69 (1.23;5.90)**
Motor score					
Syn *(Bayley-III)*	69/475	103.2 (99.3;107.1)†	108.0 (104.8;111.2)†	-	MD -4.8 (-9.8;0.3)*
Developmental disorders §					
Wang – Danish cohort	14,189/ 1,109,370	55 (0.4%)	3362 (0.3%)	-	**HR 1.33 (1.02-1.75)***

Statistically significant results are marked in **bold** (*P*<0.05). * Adjusted results. † Presented as unadjusted means with 95% CI, due to missing descriptive statistics. Severe NVP and no NVP group as control are shown (mild to moderate NVP group is not presented). ‡ Frequencies displayed are families with children exposed to HG/total number of families included in study. § Developmental disorders includes language, learning, and motor skills disorders. **Abbreviations:** Bayley-III: Bayley Scales of Infant and Toddler Development, Third Edition. HR: hazard ratio. KBIT: Kaufman Brief Intelligence Test. MD: mean difference. NIH Toolbox: NIH Toolbox Cognition Battery. OR: odds ratio. WISC-R: Wechsler Intelligence Scale for Children-Revised.

Supplement Table S3. Overview of results of included studies reporting on mental health and neurobehavioral development

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Other results (β (95% CI); mean difference (95% CI); OR (95% CI); HR (95% CI))
Mental health					
Emotional disorder					
Mullin	87/259	1 (1.1%)	0 (0%)	-	-
Wang – Danish cohort	14,189/ 1,109,370	44 (0.3%)	2884 (0.3%)	-	**HR 1.33 (0.98;1.89)***
Affective disorder					
Syn *(CBCL, 2 yr)* †	55/393	3.2 (2.4;4.1)‡	2.1 (1.4;2.7)‡	**<0.05**	**MD 1.2 (0.1;2.2)***
Syn *(CBCL, 4 yr)* † ¶	111/657	2.8 (2.3;3.4)‡	2.1 (1.6;2.6)‡	**<0.05**	**MD 0.8 (0.0;1.5)***
Anxiety					
Fejzo 2019	267/360	30.0%	21.5%	**0.01**	**OR 2.23 (1.21;4.10)**
Mullin	87/259	6 (6.9%)	4 (2.3%)	-	-
Wang – USA cohort *(CBCL)* †	1,496/9,214	2.63 ± 2.69	1.95 ± 2.36	**<0.001**	-
Syn *(CBCL, 4 yr)* † ¶	111/657	3.7 (3.1;4.3)‡	2.9 (2.4;3.3)‡	**<0.05**	**MD 0.6 (0.0;1.2)***
Depression					
Mullin	87/259	14 (16.1%)	5 (2.9%)	-	**OR 6.35**
Wang – USA cohort *(CBCL)* †	1,496/9,214	1.65 ± 2.41	1.18 ± 1.91	**<0.001**	-
Bipolar disorder					
Mullin	87/259	7 (8.0%)	3 (1.7%)	**-**	**OR 4.90**
Social developmental delay/ social anxiety					
Fejzo 2015	203/292 §	10.3%	2.3%	**0.03**	**OR 5.02 (1.43;31.83)**
Fejzo 2019	267/360	13.9%	4.30%	**0.02**	**OR 3.58 (1.24;10.33)**
Neurobehavioral development					
ASD (symptoms)					
Fejzo 2019	267/360	8.2%	0.0%	**0.048**	-
Getahun	14,526/ 455,263	332 (2.3%)	8115 (1.8%)	-	**HR 1.53 (1.37;1.70)***
Hisle-Gorman	2,459/35,040	712 (29.0%)	8,399 (25.8%)	-	**OR 1.10 (1.02–1.18)***
Wang – Danish cohort	14,189/ 1,109,370	232 (1.6%)	17,478 (1.6%)	**-**	**HR 1.19 (1.05;1.36)***
Syn *(Q-CHAT)*	24/228	38.8 (35.5;42.0)‡	34.7 (32.4;37.0)‡	**<0.05**	**MD 4.1 (0.1;8.0)***
Autism					
Mullin	87/259	0 (0%)	1 (0.6%)	-	-
Wang – Danish cohort	14,189/ 1,109,370	100 (0.7%)	6877 (0.6%)	-	**HR 1.19 (0.97;1.45)***

Supplement Table S3. Overview of results of included studies reporting on mental health and neurobehavioral development. *Continued*

Study	HG-exposed/ total sample size	HG-exposed (mean ± SD; median (IQR); frequencies (%))	Non-exposed (mean ± SD; median (IQR); frequencies (%))	*P*	Other results (β (95% CI); mean difference (95% CI); OR (95% CI); HR (95% CI))
ADD/ADHD					
Fejzo 2015	203/292 §	18.7%	5.6%	**0.01**	**OR 3.87 (1.56;11.55)**
Fejzo 2019	267/360	22.9%	11%	**0.01**	**OR 2.46 (1.20;5.03)**
Wang – Danish cohort	14,189/ 1,109,370			-	**HR 1.16 (1.06;1.28)***
ADHD (symptoms)					
Mullin	87/259	3 (3.4%)	4 (2.3%)	-	-
Syn *(CBCL, 2 yr)* †	55/393	5.8 (4.9;6.7)‡	4.3 (3.6;5.0)‡	**<0.05**	**MD 1.5 (0.4;2.6)***
Wang – USA cohort *(CBCL)* †	1,496/9,214	3.10 ± 3.21	2.45 ± 2.86	**<0.001**	-
ADD					
Mullin	87/259	1 (1.1%)	4 (2.3%)	-	-
Conduct disorders/ Oppositional defiant disorders					
Wang – Danish cohort	14,189/ 1,109,370	29 (0.2%)	2024 (0.2%)	**-**	**HR 1.06 (1.71;1.57)***
Conduct score					
Wang – USA cohort *(CBCL)* †	1,496/9,214	1.69±2.79	1.15±2.17	**<0.001**	-
Oppositional defiant disorders					
Wang – USA cohort *(CBCL)* †	1,496/9,214	2.12±2.20	1.67±1.97	**<0.001**	-
Sleep problems					
Fejzo 2019	267/360	16.1%	6.5%	**0.02**	**OR 2.78 (1.14;6.77)**
Mullin	87/259	1 (1.1%)	0 (0%)	**-**	-
Syn *(ITSEA, 1 yr)*	85/530	3.3 (2.7;3.8)‡	2.4 (1.9;2.8)‡	**<0.05**	**MD 0.9 (0.2;1.6)***
Syn *(CBCL, 2 yr)* †	55/393	3.7 (2.9;4.5)‡	2.6 (2.0;3.3)‡	**<0.05**	**MD 1.1 (0.1;2.1)***
Somatic problems					
Wang – USA cohort *(CBCL)* †	1,496/9,214	1.34±1.68	1.05±1.47	**<0.001**	-
CBCL (total score)					
Wang – USA cohort	1,496/9,214	22.74 ± 20.8	17.02 ± 16.9	**<0.001**	-

Statistically significant results are marked in **bold** (*P*<0.05). * Adjusted results. † Only reported DSM subcategories of CBCL due to extensiveness of questionnaire (20 subdomains). ‡ Presented as unadjusted means with 95% CI, due to missing descriptive statistics. Severe NVP and no NVP group as control are shown (mild to moderate NVP group is not presented). ¶ Only present in maternal, not paternal report. § Frequencies displayed are families with children exposed to HG/total number of families included in study. **Abbreviations:** ADD: Attention Deficit Disorder. ADHD: Attention Deficit Hyperactivity Disorder. ASD: Autism Spectrum Disorder. CBCL: Child Behaviour Checklist. HR: hazard ratio. ITSEA: Infant-Toddler Social and Emotional Assessment. MD: mean difference. OR: odds ratio. Q-CHAT: Quantitative Checklist for Autism in Toddlers.

Supplement Table S4. Grading evidence of meta-analyses

Certainty assessment							Summary of findings				
							Study event rates (%)			Anticipated absolute effects	
Participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	With non-exposure	With HG-exposure	Relative effect (95% CI)	Risk with non-exposure	Risk difference with HG-exposure
Emotional disorders											
1,112,558 (2 observational studies)	serious[a]	serious[b]	not serious	serious[c]	all plausible residual confounding would reduce the demonstrated effect	⨁◯◯◯ Very low	2,884/1,098,237 (0.3%)	45/14,321 (0.3%)	**OR 1.19** (0.89 to 1.61)	**Low** 0 per 1,000	**0 fewer per 1,000** (from 0 fewer to 0 fewer)
Anxiety											
729 (2 observational studies)	very serious[d]	not serious	not serious	very serious[e]	all plausible residual confounding would reduce the demonstrated effect	⨁◯◯◯ Very low	24/289 (8.3%)	86/440 (19.5%)	not pooled	**Low** not pooled	not pooled
ASD											
1,649498 (4 observational studies)	not serious	very serious[f]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect dose response gradient	⨁⨁◯◯ Low	33,,280/1,615,439 (2.1%)	1,659/34,059 (4.9%)	not pooled	not pooled	not pooled
Autism											
1,116,607 (2 observational studies)	serious[g]	very serious[b]	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect dose response gradient	⨁◯◯◯ Very low	6,878/1,102,231 (0.6%)	100/14,376 (0.7%)	**OR 1.12** (0.92 to 1.37)	6 per 1,000	**1 more per 1,000** (from 0 fewer to 2 more)

Supplement Table S4. *Continued*

Certainty assessment							Summary of findings				
Participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With non-exposure	With HG-exposure		Risk with non-exposure	Risk difference with HG-exposure
Sleep problems											
669 (2 observational studies)	very serious[d]	not serious	not serious	very serious[c]	strong association all plausible residual confounding would reduce the demonstrated effect	⨁◯◯◯ Very low	6/271 (2.2%)	44/398 (11.1%)	**OR 2.94** (1.25 to 6.93)	22 per 1,000	**40 more per 1,000** (from 5 more to 113 more)
Testicular cancer											
22,126 (5 observational studies)	not serious	not serious	not serious	serious[c]	none	⨁◯◯◯ Very low	1,144/21,530 (5.3%)	52/596 (8.7%)	**OR 1.60** (1.07 to 2.39)	53 per 1,000	**29 more per 1,000** (from 4 more to 65 more)

ADD/ADHD: attention deficit (hyperactivity) disorder; **ASD:** autism spectrum disorders; **CI:** confidence interval; **OR:** odds ratio

Explanations

a. Wang et al was rated to be of good quality, but Mullin et al however, was rated to be of poor quality.
b. Differences in estimates of effect.
c. Wide confidence interval.
d. Both Fejzo and Mullin et al were rated to be of poor quality.
e. Low number of events and a wide confidence interval.
f. Differences in estimates of effect and high heterogeneity.
g. Wang et al was rated to be of good quality, but Fejzo et al however, was rated to be of poor quality.

CHAPTER 9

Hyperemesis gravidarum severity, enteral tube feeding and cardiometabolic markers in offspring cord blood

Kelly Nijsten, Marjette H. Koot, Joke M.J. Bais, Carrie Ris-Stalpers, Rik van Eekelen, Henk A. Bremer, David P. van der Ham, Wieteke M. Heidema, Anjoke Huisjes, Gunilla Kleiverda, Hinke Kruizenga, Simone M. Kuppens, Judith O.E.H. van Laar, Josje Langenveld, Flip van der Made, Dimitri Papatsonis, Marie-José Pelinck, Paula J. Pernet, Leonie van Rheenen-Flach, Robbert J. Rijnders, Hubertina C.J. Scheepers, Tatjana Vogelvang, Ben W. Mol, Iris J. Grooten, Tessa J. Roseboom, Rebecca C. Painter

British Journal of Nutrition, 2022; 1-11.

ABSTRACT

The present study aimed to investigate the association between hyperemesis gravidarum (HG) severity and the effect of early enteral tube feeding on cardiometabolic markers in offspring cord blood.

We included women admitted for HG, who participated in the MOTHER randomized controlled trial (RCT) and observational cohort. The MOTHER RCT showed that early enteral tube feeding in addition to standard care did not affect symptoms or birth outcomes. Among RCT and cohort participants, we assessed how HG severity affected lipid, c-peptide, glucose and free thyroxine cord blood levels. HG severity measures were: severity of vomiting at inclusion and three weeks after inclusion, pregnancy weight gain and 24-hour energy intake at inclusion, readmissions and duration of hospital admissions. Cord blood measures were also compared between RCT participants allocated to enteral tube feeding and those receiving standard care.

Between 2013-2016, 215 women were included: 115 RCT and 100 cohort participants. Eighty-one cord blood samples were available. Univariable, not multivariable regression analysis showed that lower maternal weight gain was associated with higher cord blood glucose levels (β:-0.08, 95% CI:-0.16;-0.00). Lower maternal weight gain was associated with higher apolipoprotein-B cord blood levels in multivariable regression analysis (β:-0.01, 95% CI:-0.02;-0.01). No associations were found between other HG severity measures or allocation to enteral tube feeding and cord blood cardiometabolic markers.

In conclusion, while lower maternal weight gain was associated with higher apolipoprotein-B cord blood levels, no other HG severity measures were linked with cord blood cardiometabolic markers, nor were these markers affected by enteral tube feeding.

INTRODUCTION

Hyperemesis gravidarum (HG) is a severe form of nausea and vomiting in pregnancy that affects up to 3.6% of pregnancies.[1] HG can lead to poor nutritional intake and significant weight loss.[2,3].

Maternal undernutrition during pregnancy has been demonstrated to lead to detrimental effects on offspring cardiometabolic disease risk in later life, for example among adults exposed to maternal starvation while they were *in utero* during the Dutch Famine.[4-6] Early cardiometabolic disease markers are known to track through childhood into adulthood, enabling estimates of cardiometabolic diseases to be made in childhood, or even in infancy.[7] Cord blood lipid profile, glucose insulin metabolism and thyroid function have been demonstrated to provide such a window into the future.[8-10]

The nutritional intake of pregnant women with HG, which studies estimate to be as low as 450-1000 calories per day, falls drastically short of the recommendations.[11] Effects of HG on offspring's health in later life therefore seem likely, but have sparsely been studied.[12] In the limited literature available on the topic of long term effects of HG on offspring, a wide range of HG definitions and outcomes were employed, and little information was collected on the role of maternal nutrition.[13-16] This is of particular relevance as nutritional supplementation is not generally included in HG clinical management plans.[17] The role of maternal nutrition in HG, or the role of HG disease severity on offspring health in later life has not been investigated, nor has the ability of nutritional supplementation to amend such effects.

Therefore, in this study, we aimed to investigate whether HG severity could affect cardiometabolic markers in offspring cord blood, as well investigating the effect of early enteral tube feeding on these markers.

METHODS

Study design

For this study, we used data from the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) study.[18] The MOTHER trial, a multicentre open-label randomized controlled trial (RCT), aimed to evaluate the effect of early enteral tube feeding in addition to standard care for women hospitalized with HG. Beside this RCT, an observational cohort was assembled of eligible women who declined participation in the RCT, but did consent to cohort participation. For more detailed information about the MOTHER study methodology, we refer to the corresponding study protocol and earlier published work.[18, 19] Participants of both RCT and cohort were asked to provide informed consent for collection of maternal and umbilical cord blood for biobanking. The MOTHER study was registered at www.trialregister.nl (NTR4197) and approved by the research ethics committee of the Amsterdam UMC, location AMC (NL41011.018.12).

Population

Women admitted to hospital for HG between 5 and 20 weeks gestation were included in 19 different hospitals in the Netherlands: three university medical centres and sixteen general hospitals that collaborated within the Dutch Consortium for studies in Obstetrics and Gynaecology. In our study, HG was defined as excessive vomiting necessitating hospital admission before 20 weeks gestation, and in the absence of any other causes of vomiting. Exclusion criteria were: molar or non-vital pregnancies, maternal HIV, any contra-indication for enteral tube feeding and age <18 years. Eligible participants were identified by local staff of participating hospitals. Multiple pregnancies were excluded for the present cord blood study, because of the possibility that this could have independently affected offspring cord blood measures.

Data collection

Medical information was obtained from medical files by trained research staff and reported in a case report form. This case report form included detailed information about demographic characteristics as age, parity, gestational age and medical history as well as information about pregnancy and delivery. Preterm birth was defined as birth before the 37th week of gestation. Neonates were considered small for gestational age (SGA) if their birth weight was below the 10th percentile according to the Dutch reference curves for birth weight by gestational age.[20] Pre-pregnancy weight, height, ethnicity and education level at baseline were self-reported. Pre-pregnancy BMI was calculated. Ethnicity was based on the country of birth of the participant's

mother and for this study defined as Western (including participants from Organisation for Economic Co-operation and Development (OECD) countries) or non-Western.[21] Education level was defined as the highest completed education: primary or secondary school versus higher education. Weight (in kg) at inclusion was measured by hospital staff. Participants also recorded a weekly diary until 20 weeks gestation including symptom severity, maternal weight and a comprehensive dietary intake list. Symptom severity was measured by a validated questionnaire: the Pregnancy Unique Quantification of Emesis and nausea (PUQE-24) score; higher PUQE-24 scores indicate more severe symptoms and have previously been associated with a lower nutritional intake.[3, 22]

Data collection on nutritional intake

MOTHER participants kept a weekly self-reported, 24-hour food diary, as shown in **Supplement Figure A1**. The food diary consisted of a list of pre-specified food products, with the ability to specify non-listed food products among 'others'. Intake of food products could be filled in in millilitres as well as in pieces at six different moments. In collaboration with a dietician, databases of the Dutch National Institute for Public Health and the Environment (RIVM) were used to determine standard portion sizes and the corresponding energy content (kcal) of reported food products **(Supplement Table A1)**.[23-25] Eventually, a 24-hour energy intake was calculated per women by adding up caloric quantities of the consumed food products. More detailed information about how dietary intake was collected and energy intake was calculated is shown in **Appendix A**. According to the Dutch guidelines, a 24-hour energy intake of at least the recommended 1870 kcal was considered normal for pregnant women.[26] This is based on 85% of the average daily energy intake in Dutch pregnant women, which is advised to be used as cut-off for the recommended daily intake by the Netherlands Nutrition Centre, since this has previously been shown to lead to a sufficient intake of essential recommended micronutrients, fibres and fatty acids.[26, 27]

Maternal and umbilical cord blood samples

Random maternal blood samples were drawn during admission after study entry. Umbilical cord blood was taken after delivery, collected in a 20 millilitres (ml) syringe and then divided over five different blood aliquots of 4 ml each: one EDTA tube, one heparin tube, one sodium fluoride tube and two serum tubes. Both maternal and cord blood samples were then transported and stored in one central laboratory at -80 degrees until assayed. For this study, frozen stored samples were analysed in August 2019, 3 to 6 years after blood collection. We analysed apolipoprotein A1 (Apo-A1), apolipoprotein B (Apo-B), Low Density Lipoprotein (LDL), High Density lipoprotein (HDL), total cholesterol, triglycerides, c-peptide and free thyroxine

(FT4) in serum cord blood samples. Cord blood glucose was analysed in sodium fluoride blood samples to prevent rapid glycolysis and therefore an underestimating of glucose levels. We also analysed maternal lipid profile and FT4 at baseline in frozen stored maternal blood samples. Apo A1 and Apo B were measured on an Architect ci8200 with use of immunoturbidimetric assay (Abbott Laboratories, Illinois), with an intra-assay coefficient of variation (CV) of respectively 0.8% and 1.4% and an inter-assay CV of 1.6% and 2.4 %. Total cholesterol was measured on a Roche Cobas 502 by use of photometric assay tests (Roche Diagnostics, Mannheim, Germany) with an intra-assay CV of 1.1% and an inter-assay CV of 1.6%. HDL, triglycerides and glucose were measured on a Roche Cobas 702 by use of photometric assay (Roche Diagnostics, Mannheim, Germany) with intra-assay CV of 1.0%, 0.9% and 0.8% and inter-assay CV of 1.3%, 2.0% and 1.3%, respectively. LDL was calculated by subtracting the HDL concentration and 0.45 times the triglycerides concentration of the total cholesterol concentration. C-peptide was measured with luminescence immunometric assay on a Advia Centaur analyser (Siemens Medical Solutions Diagnostics, Malvern, United States of America), with an intra-assay CV of 5% and an inter-assay of 7%. FT4 was measured by a Roche Cobas e602 immuno-analyser (Roche Diagnostics, Mannheim, Germany) with an intra-assay CV of 2.0% and an inter-assay of 2.2 %.

Determinants: measures of HG severity

We assessed HG severity as the severity of vomiting at inclusion (PUQE-24) and in the first three weeks after inclusion (average PUQE-24), the 24-hour energy intake at inclusion, weight change at inclusion compared to pre-pregnancy weight, duration of hospital admissions, readmissions and readmission after the first trimester.

Statistical analysis

Due to the fact that early enteral tube feeding did not affect perinatal and maternal outcomes, we combined the RCT and cohort into one study population to assess associations between HG severity and cord blood outcomes.[18] SPSS Statistics 26.0 for Windows (IBM Corp., Armonk, NY, USA) was used for all analysis.

Descriptive statistics

Normally distributed continuous variables are presented as means with standard deviations (SDs), skewed distributions as medians with interquartile ranges (IQRs) and categorical variables as frequencies with percentages. For FT4 levels we were able to report how many neonates had an FT4 level below (<12.0 pmol/l) or above (>36.0 pmol/l) the national Dutch cord blood reference interval.[28] No corresponding reference intervals are available for any of the other

endocrine or lipid measures in cord blood, therefore we have reported these as continuous measures.

Associations between disease severity and cord blood outcomes

We performed univariable and multivariable linear regression analysis to assess associations between measures of HG severity and endocrine and lipid measures in cord blood. After assessing the effect of HG severity and other maternal and perinatal factors on endocrine and lipid measures in cord blood in univariate analyses, maternal and perinatal factors associated with endocrine and lipid measures in cord blood with a *P*-value below 0.2 on univariable analysis were considered for assessment in multivariable regression models. HG severity was kept as the primary determinant. In the multivariable model for the association of HG measures of severity on cord blood lipid measures, we adjusted for preterm birth, smoking, ethnicity, maternal diabetes (gestational, type I or II) and maternal lipid profile at study inclusion (i.e. maternal Apo-A1, Apo-B, HDL, LDL, total cholesterol or triglycerides, depending on which cord blood lipid profile measure was analysed). In the multivariable model for the association of HG measures of severity on cord blood glucose insulin metabolism (i.e. c-peptide and glucose), we adjusted for preterm birth, SGA, maternal age and pre-pregnancy BMI. In the multivariable model for the association of HG measures of severity on cord blood FT4, we adjusted for smoking, highest finished education and maternal FT4 levels at study inclusion. Associations between HG measures of severity and variables that were normally distributed were reported in differences (β) and 95% confidence intervals (95% CIs). Variables of which residuals were not normally distributed were logarithmically transformed to achieve normality before entry into the multivariable models. Subsequently these variables were back-transformed and reported in proportionate differences (β) and 95% CIs (in percentages).

Effect of nutritional supplementation

The effect of nutritional supplementation was evaluated in the RCT study population alone. We separately performed intention-to-treat, per protocol and 'as treated' analyses, in which the cord blood endocrine and lipid measures of patients in the RCT were reported according to allocation group to which they had been randomized (intention-to-treat), according to allocation and adherence to protocol (per protocol) or according to treatment with enteral tube feeding ('as treated'). For eligibility for the per protocol analysis, participants allocated to the intervention arm had to have received a nasogastric tube within *3 days* after randomization and continued for 7 days or longer, and, participants allocated to standard care had to have received standard care alone in the first *3 days* after randomization. For the 'as treated' analysis, participants allocated to the intervention arm had to have received nasogastric tube within

7 days after randomization and continued for 7 days or longer while participants allocated to standard care had to have received standard care alone in the first *7 days* after randomization. Chi-square test, Mann-Whitney U test and independent Student's *t* test were used for statistical analyses.

Sensitivity analysis

A sensitivity analysis was performed to assess selective participation, by looking for differences in baseline characteristics and measures of HG severity between women with and without cord blood samples available.

Power calculation

A power calculation was performed for the original MOTHER study to determine the sample size of the RCT: 120 participants were randomised, based on finding a mean difference in birth weight of 200 gram between the two intervention arms.[18] A post-hoc power calculation for this study was not performed, since there was already a set number of cord blood samples available together with the fact that post-hoc power calculations for multiple regression analysis are based on the coefficient of determinations (R^2), which differ for each of the performed multivariable regression analyses.

RESULTS

Between 2013 and 2016, 215 women participated in the MOTHER study: 115 participants in the RCT and 100 participants in the observational cohort. From the 210 singleton pregnancies, we had 81 cord blood samples available as shown in **Figure 1**. Baseline characteristics and outcomes of women included in this study are shown in **Table 1.** The median 24-hour energy intake at inclusion was 435 kcal (IQR 152-1021). Sixty two out of 67 women (92.5%) for whom 24-hour energy intake was available, had a 24-hour energy intake that fell short of the recommended daily intake (1870 kcal).[26]

Sensitivity analysis showed that women with cord blood samples available had higher maternal triglyceride levels at inclusion (median 0.83; IQR 0.64-1.18 mmol/L) than women without cord blood samples available (median 0.75; IQR 0.61-1.02 mmol/L, *P*-value 0.045). **(Supplemental Table S1).** Further baseline characteristics and outcome measures did not differ.

Maternal and perinatal factors and their association with endocrine and lipid measures in cord blood

Univariable regression analyses assessing associations between maternal and perinatal factors and each of the endocrine and lipid measures are presented in **Supplemental Table S2**. Higher maternal prepregnancy BMI was associated with higher c-peptide levels in cord blood (β 0.01, 95% CI: 0.00;0.01). Smoking was associated with higher cord blood Apo-B (β 0.13, 95% CI: 0.01;0.25) and higher LDL levels (β 0.46, 95% CI: 0.06;0.87). Maternal gestational diabetes, diabetes type I or II during pregnancy was associated with higher Apo-B (β 0.20, 95% CI: 0.06;0.33), higher LDL (β 0.77, 95% CI: 0.23;1.31) and higher total cholesterol levels in cord blood (β 0.85, 95% CI: 0.11;1.60). No associations were found between maternal age, ethnicity, highest completed educational level and maternal hypo- or hyperthyroidism and any of the endocrine and lipid measures in cord blood.

Regarding perinatal risk factors for endocrine and lipid measures in cord blood, we found that preterm birth was associated with higher Apo-B (β 0.18, 95% CI: 0.05;0.30) and higher total cholesterol levels in cord blood (β 0.75, 95% CI: 0.17;1.34). Preterm birth was also associated with lower cord blood LDL levels (β -0.66, 95% CI: -1.08;-0.24). No associations were found between SGA or fetal sex and any of the endocrine and lipid measures in cord blood.

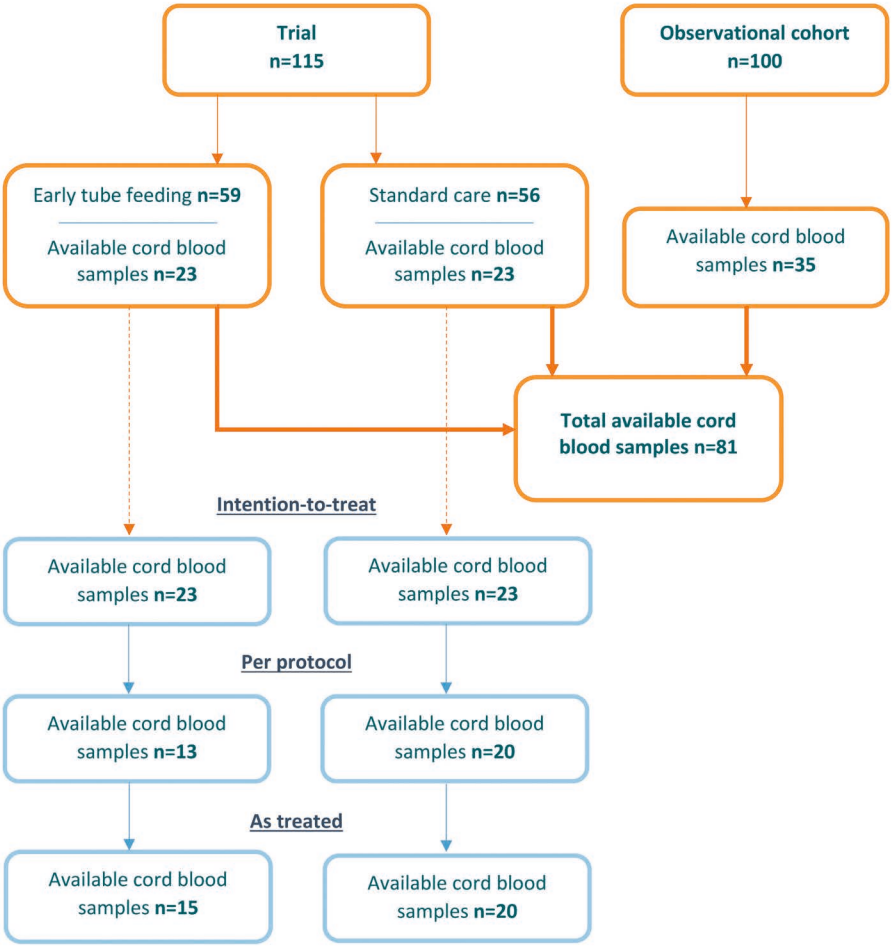


Figure 1. Flowchart available cord blood samples

Table 1. Baseline characteristics and outcome measures of women included in this study

	N=81	Missing, %
Demographics		
Age (years)	28.3 ± 4.2	0.0%
Prepregnancy weight (kg)	73.4 ± 15.2	2.5%
Prepregnancy BMI (kg/m^2)	25.5 ± 4.8	2.5%
Non-western ethnicity	18 (22.2%)	9.9%
Education level		19.8%
- Primary or secondary	38 (46.9%)	
- Higher	27 (33.3%)	
Primigravida	27 (33.3%)	0.0%
Maternal diabetes (gestational, type I or II)	4 (4.9%)	0.0%
Maternal thyroid disease	4 (4.9%)	0.0%
Current smoker	7 (8.6%)	2.5%
Gestational age at onset of HG symptoms (wks)	6.0 (5.0-7.0)	18.5%
Gestational age at inclusion (wks)	9.0 (7.0-11.0)	0.0%
Measures of HG severity		
Weight change (kg)	-3.0 ± 3.8	2.5%
24-hour energy intake at inclusion (kcal)	435 (152-1021)	17.3%
Below recommended daily intake (<1870 kcal)	62 (76.5%)	17.3%
PUQE-24 at inclusion	10.4 ± 3.2	23.5%
Average PUQE-24 in the first three weeks after inclusion	8.7 ± 2.9	27.2%
Total duration of hospital admissions (days)	5.0 (3.0-7.0)	0.0%
Readmission	27 (33.3%)	0.0%
Readmission after the first trimester	21 (25.9%)	0.0%
Perinatal outcomes		
Birth weight (grams)	3315 ± 513	0.0%
SGA (birth weight <10th percentile)	6 (7.5%)	1.2%
Prematurity (< 37 weeks)	5 (6.2%)	0.0%
Apgar score <7 at 5 min	2 (2.5%)	1.2%
Fetal sex (female)	46 (56.8%)	0.0%
Maternal blood measurements		
Apolipoprotein A1 (g/L)	1.43 ± 0.31	19.8%
Apolipoprotein B (g/L)	0.80 ± 0.25	19.8%
HDL (mmol/L)	1.33 ± 0.33	19.8%
LDL (mmol/L)	2.43 ± 0.83	19.8%
Total cholesterol (mmol/L)	4.21 ± 1.07	19.8%

Table 1. *Continued*

	N=81	Missing, %
Triglycerides (mmol/L)	0.83 (0.64-1.18)	19.8%
FT4 (pmol/L)	19.79 ± 5.81	19.8%
Cord blood measurements		
Apolipoprotein A1 (g/L)	0.90 ± 0.18	7.4%
Apolipoprotein B (g/L)	0.28 ± 0.14	7.4%
HDL (mmol/L)	0.85 ± 0.22	8.6%
LDL (mmol/L)	0.95 ± 0.48	13.6%
Total cholesterol (mmol/L)	1.98 ± 0.66	11.1%
Triglycerides (mmol/L)	0.35 (0.21-0.54)	9.9%
Glucose (mmol/L)	4.70 ± 1.26	8.6%
C-peptide (nmol/L)	0.20 ± 0.14	4.9%
FT4 (pmol/L)	17.48 ± 2.07	24.7%

Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). **Abbreviations:** BMI: body mass index. HDL: High Density Lipoprotein. HG: hyperemesis gravidarum. FT4: free thyroxine. LDL: Low Density Lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms. SGA: small for gestational age. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight.

Associations between measures of HG severity and endocrine and lipid measures in cord blood

Lipid profile

Lower maternal weight gain at study entry was associated with higher Apo-B levels (β -0.01 g/L per kilogram maternal weight gain, 95% CI: -0.02;-0.00, *P*=0.005) in multivariable linear regression analysis **(Table 2)**. The 24-hour energy intake at inclusion, nor any of the other measures of HG severity were associated with lipid profile in cord blood.

C-peptide and glucose

Lower maternal weight gain at study entry was associated with higher glucose levels in cord blood in univariable linear regression analysis, (β -0.08 mmol/L per kilogram maternal weight gain, 95% CI: -0.16;-0.00, *P*=0.043), but the association was not sustained in multivariable models, as shown in **Table 3**. None of the measures of severity of HG were associated with c-peptide in cord blood.

FT4

None of the measures of HG disease severity were associated with FT4 in cord blood **(Table 3)**. One neonate had a FT4 cord blood level below the corresponding Dutch national reference

interval (<12.0 pmol/L) of 10.4 pmol/L. In this specific case, the mother had no medical history of hypo- or hyperthyroidism. She herself had normal Thyroid Stimulating Hormone and FT4 concentrations at baseline.[29] Although birth weight was considered normal and the neonate was born at full term, the neonate was admitted to hospital postpartum due to suspicion of infection. No additional information on admission or further thyroid function measures was known. None of the neonates included in this study had a FT4 cord blood level above the corresponding Dutch national reference interval (>36.0 pmol/L).

Table 2. Univariable and multivariable regression analysis assessing the association between measures of HG severity and lipid measures in cord blood

	Apo-A1 (g/L)						Apo-B (g/L)					
	Model 1			Model 2			Model 1			Model 2		
	β	**95% CI**	***n***	***β***	**95% CI**	***n***	***β***	**95% CI**	***n***	***β***	**95% CI**	***n***
- PUQE-24 at inclusion	-0.01	-0.02;0.01	56	-0.01	-0.02;0.01	42	-0.01	-0.01;0.00	56	-0.00	-0.01;0.01	42
- Average PUQE-24 first three weeks after inclusion	0.00	-0.02;0.02	54	-0.01	-0.03;0.01	38	0.01	-0.01;0.02	54	0.00	-0.01;0.01	38
- Weight change (kg)	-0.00	-0.01;0.01	72	-0.01	-0.02;0.01	50	0.00	-0.01;0.01	72	**-0.01**	**-0.02;-0.00**	**50**
- Energy intake at inclusion (kcal)	-0.00	-0.01;0.01	61	0.00	-0.01;0.01	46	-0.00	-0.01;0.00	61	0.00	-0.00;0.00	46
- Duration of hospital admissions (days)	-0.00	-0.01;0.00	74	-0.00	-0.01;0.00	50	0.00	-0.00;0.00	74	-0.00	-0.00;0.00	50
- Readmitted	0.03	-0.06;0.12	74	0.04	-0.05;0.13	50	0.05	-0.02;0.11	74	0.02	-0.04;0.08	50
- Readmission after the first trimester	-0.04	-0.14;0.05	74	-0.03	-0.14;0.08	50	-0.06	-0.13;0.02	74	-0.02	-0.09;0.05	50
	HDL (mmol/L)						**LDL (mmol/L)**					
	Model 1			Model 2			Model 1			Model 2		
	β	**95% CI**	***n***	***β***	**95% CI**	***n***	***β***	**95% CI**	***n***	***β***	**95% CI**	***n***
- PUQE-24 at inclusion	-0.01	-0.03;0.01	55	0.00	-0.02;0.02	41	-0.02	-0.05;0.01	52	-0.02	-0.06;0.02	38
- Average PUQE-24 first three weeks after inclusion	-0.00	-0.03;0.02	53	-0.01	-0.03;0.02	37	0.04	-0.02;0.09	49	0.01	-0.04;0.06	33
- Weight change (kg)	0.00	-0.01;0.02	71	-0.00	-0.02;0.01	49	0.00	-0.03;0.04	67	-0.03	-0.05;-0.00	45
- Energy intake at inclusion (kcal)	0.00	-0.01;0.01	60	0.00	-0.01;0.01	45	-0.01	-0.03;0.01	56	-0.01	-0.02;0.01	41
- Duration of hospital admissions (days)	-0.00	-0.01;0.00	73	-0.00	-0.01;0.00	49	-0.00	-0.02;0.01	69	-0.00	-0.01;0.01	45
- Readmitted	0.55	-0.14;0.08	73	0.01	-0.11;0.14	49	0.12	-0.13;0.36	69	0.09	-0.12;0.29	45
- Readmission after the first trimester	-0.02	-0.13;0.10	73	-0.02	-0.17;0.12	49	-0.10	-0.36;0.16	69	-0.07	-0.17;0.32	45

	Total cholesterol (mmol/L)						Triglycerides ∫ (mmol/L)					
	Model 1			Model 2			Model 1			Model 2		
	β	**95% CI**	***n***	**β**	**95% CI**	***n***	**β**	**95% CI**	***n***	**β**	**95% CI**	***n***
- PUQE-24 at inclusion	-0.03	-0.07;0.02	54	-0.03	-0.09;0.03	40	-2.57	-7.13;2.33	54	-5.16	-11.31;1.41	40
- Average PUQE-24 first three weeks after inclusion	0.04	-0.03;0.11	51	0.01	-0.06;0.08	35	2.53	-2.86;8.11	52	-1.29	-7.87;5.65	36
- Weight change (kg)	0.01	-0.03;0.05	69	-0.03	-0.07;-0.01	47	0.40	-3.34;4.19	70	-1.78	-6.29;3.05	48
- Energy intake at inclusion (kcal)	-0.01	-0.04;0.02	58	-0.00	-0.02;0.02	43	-0.70	-3.05;1.61	59	0.60	-2.27;3.46	44
- Duration of hospital admissions (days)	-0.01	-0.03;0.01	71	-0.01	-0.02;0.01	47	0.10	-1.49;1.82	72	-0.40	-2.18;1.31	48
- Readmitted	0.13	-0.20;0.46	71	0.14	-0.15;0.43	47	30.60	-2.26; 75.07	72	29.18	-9.70; 84.97	48
- Readmission after the first trimester	-0.16	-0.51;0.19	71	-0.02	-0.32;0.36	47	-14.62	-37.56; 16.77	72	2.43	-34.88; 61.12	48

Significant associations with a *P*-value<0.05 are marked in bold. **Abbreviations:** Apo-A1: apolipoprotein A1, Apo-B: apolipoprotein B, HDL: high density lipoprotein, LDL: low density lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score; the PUQE-24 score can range from 3 to 15 with a higher PUQE-24 indicating more severe symptoms. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. Energy intake at inclusion is calculated over 24 hours, measured in kilocalories and divided by 100 for regression analysis. **∫** = log transformed, back transformed and expressed in % of difference. **Model 1:** univariable. **Model 2:** multivariable regression, adjusted for preterm birth, smoking, ethnicity (western or not), maternal diabetes (gestational, type I or II) and maternal lipid profile at inclusion.

Table 3. Univariable and multivariable regression analysis assessing the association between measures of HG severity and endocrine measures in cord blood

				C-peptide (nmol/L)		
		Model 1			**Model 2***	
	β	**95% CI**	***n***	***β***	**95% CI**	***n***
- PUQE-24 at inclusion	-0.00	-0.02;0.01	58	-0.01	-0.02;0.01	56
- Average PUQE-24 first three weeks after inclusion	0.01	-0.01;0.02	56	0.01	-0.01;0.02	54
- Weight change (kg)	-0.01	-0.01;0.00	74	-0.00	-0.01;0.01	73
- Energy intake at inclusion (kcal)	-0.00	-0.01;0.00	63	-0.00	-0.01;0.01	61
- Duration of hospital admissions (days)	-0.00	-0.01;0.00	76	0.00	-0.00;0.00	73
- Readmitted	-0.01	-0.08;0.06	76	-0.00	-0.07;0.07	73
- Readmission after the first trimester	-0.02	-0.09;0.05	76	-0.02	-0.10;0.06	73
				Glucose (mmol/L)		
		Model 1			**Model 2***	
	β	**95% CI**	***n***	***β***	**95% CI**	***n***
- PUQE-24 at inclusion	-0.02	-0.12;0.10	55	-0.04	-0.17;0.08	53
- Average PUQE-24 first three weeks after inclusion	0.06	-0.06;0.17	54	0.06	-0.06;0.18	52
- Weight change (kg)	**-0.08**	**-0.16;-0.00**	71	-0.08	-0.16;0.01	70
- Energy intake at inclusion (kcal)	0.01	-0.04;0.06	60	0.01	-0.04;0.07	58
- Duration of hospital admissions (days)	0.01	-0.02;0.04	73	0.01	-0.03;0.04	70
- Readmitted	0.22	-0.39;0.83	73	0.10	-0.54;0.74	70
- Readmission after the first trimester	0.56	-0.09;1.21	73	0.48	-0.22;1.19	70
				FT4 (pmol/L)		
		Model 1			**Model 2****	
	β	**95% CI**	***n***	***β***	**95% CI**	***n***
- PUQE-24 at inclusion	-0.05	-0.24;0.15	45	-0.19	-0.43;0.06	31
- Average PUQE-24 first three weeks after inclusion	0.17	-0.07;0.41	43	0.20	-0.09;0.50	29
- Weight change (kg)	0.00	-0.14;0.14	58	0.11	-0.08;0.29	37
- Energy intake at inclusion (kcal)	0.08	-0.01;0.17	49	0.06	-0.04;0.17	34
- Duration of hospital admissions (days)	-0.01	-0.07;0.05	60	-0.01	-0.07;0.06	37
- Readmitted	-0.42	-1.55;0.71	60	0.25	-1.16;1.65	37
- Readmission after the first trimester	0.45	-0.76;1.66	60	-0.14	-1.71;1.43	37

Significant associations with a *P*-value<0.05 are marked in bold. **Abbreviations:** FT4: free thyroxine, PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score; the PUQE-24 score can range from 3 to 15 with a higher PUQE-24 indicating more severe symptoms. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight. Energy intake at inclusion is calculated over 24 hours, measured in kilocalories and divided by 100 for regression analysis.

Model 1: univariable regression analysis. **Model 2:** * multivariable regression analysis, adjusted for preterm birth, small for gestational age, maternal age and pre-pregnancy BMI. ** multivariable regression analysis, adjusted for smoking, highest finished education and maternal FT4 levels at study inclusion.

The effect of early enteral tube feeding on endocrine and lipid measures in cord blood

In total, we had 46 cord blood samples available for RCT participants. All included RCT participants with dietary intake at inclusion available (38/46) had a 24-hour energy intake below the recommended daily intake of 1870 kcal. When comparing baseline characteristics, no differences were found in demographics, maternal blood measurements or in perinatal outcomes, but few differences were found in measures of HG severity in intention-to-treat, per protocol and 'as treated' analysis. Participants who received early enteral tube feeding had a lower 24-hour energy intake at inclusion than participants who received standard care, in both intention-to-treat (131 (IQR 43-417) vs 483 (IQR 283-1068) kcal) and per protocol analysis (82 (IQR 15-394) vs 605 (IQR 326-1129) kcal) **(Supplemental Table S3 and S4)**. In the as treated analysis, participants who received early enteral tube feeding had a lower 24-hour energy intake at inclusion (76 (IQR 8-387) vs 605 (IQR 326-1129) kcal), higher vomiting scores at inclusion (PUQE-24: 12.6 ± 2.0 vs 10.3 ± 3.3), longer duration of admission to hospital (6.0 (IQR 3.0-10.0) vs 4.0 (IQR 2.3-5.0) days) and were readmitted more often (53% vs 20%) than participants who received standard care **(Supplemental Table S5)**.

There were no differences in Apo-A1, Apo-B, HDL, LDL, total cholesterol, triglycerides, glucose, c-peptide and FT4 cord blood levels between offspring of participants who received early enteral tube feeding and offspring of participants who received standard care in intention-to-treat, per protocol and 'as treated' analysis **(Supplemental Tables S3-S5)**.

DISCUSSION

Main findings

This small prospective cohort study was the first to find evidence that more severe HG could affect cardiometabolic disease markers in offspring at birth. We found that lower maternal weight gain at inclusion was associated with higher Apo-B and higher glucose levels in offspring cord blood. None of the other cord blood measures of possible cardiometabolic risk were affected by HG severity. We found no evidence that treatment by early enteral tube feeding improved offspring's cardiometabolic markers in cord blood.

Strengths and limitations

One of the strengths of our study is that detailed information about HG severity, including comprehensive nutritional information, was collected prospectively by trained research staff. With this study we are addressing two of the top 10 priority HG research questions that were

recently identified by patients and clinicians, indicating the urgency of the need for guidance in the consequences of nutritional deficiency among women with HG.[30] This study can help meet patients' need for information on the long term consequences of undernutrition associated with HG for their offspring - a source of significant maternal distress.[31] Despite the small sample size, sensitivity analysis did not show any differences between MOTHER participants with and without cord blood samples available. We therefore think the likelihood of selection bias is small, which implies generalizability of our findings to all women with HG. The fact that we collected markers of possible future health by assessing cord blood, enabling estimates of increased cardiometabolic disease risk early in life, is another strength of this study.[7]

Cord blood samples were only available in 81 out of 215 participants (37%). This small sample size limited our statistical power. Furthermore, there was missing data regarding weight change and dietary intake throughout pregnancy, *after* inclusion. For that reason, we were not able to assess gestational weight change and energy intake over time and include those measures as predictors in our regression analysis model. However, since PUQE-24 scores were more often reported in the first three weeks after inclusion, we were able to calculate average PUQE-24 scores and we also collected data of hospital admissions as measures of HG severity and did not find an association with endocrine and lipid measures in cord blood. While we used the validated PUQE-24 questionnaire,[32, 33] information on nutritional intake was based on a non-validated 24-hour food diary. In general however, food diaries have been considered as a reliable method to collect information on dietary intake, since they provide a detailed, prospective method with a low risk of recall bias.[34, 35]

By design, our study did not include a control group of healthy pregnant women. Therefore, we are not able to comment on any possible changes in cardiometabolic markers in cord blood of HG cases compared to those in the offspring of healthy pregnant women. Finally, it is important to notice that, since we performed multiple statistical analyses, there is a considerable probability that some of our results were due to chance. When we applied the Bonferroni correction to our data, none of the associations we found remained significant (data not shown), which may support the notion that our findings were the result of chance.[36]

Interpretation

Our study showed that lower maternal weight gain at study inclusion was associated with higher Apo-B levels in cord blood. Apo-B is a protein that is attached to (very) low and intermediate density lipoproteins and which is used for transporting lipids around the body.[37] Apo-B has an important role in the pathogenesis of atherosclerosis and has previously been

shown to be an independent and better predictor for cardiovascular disease risk than LDL.[37, 38] Birth weight can be a possible mediator for Apo-B levels in cord blood, although literature shows conflicting results.[39-41] We did not find an association between Apo-B cord blood levels and SGA neonates. Our sample size however, prohibited any firm conclusions regarding infrequent outcomes including SGA (6 out of 81 neonates) and preterm birth (5 out of 81 neonates). Our study suggests that per extra kilogram maternal weight loss, Apo-B cord blood levels increase with 0.01 g/L. Considering the reference interval for Apo-B for infants younger than 14 days of 0.10 to 0.67 g/L, the clinical relevance of this difference, even in cases with severe weight loss, might be low.[42]

Currently, there is only a small body of evidence available describing the effects of HG on offspring's cardiovascular disease risk in later life in comparison to that among offspring of uncomplicated pregnancies.[13-16] None of these studies found differences in offspring lipid concentrations in childhood.[13-16] Two studies however, did find that offspring born to women with HG were heavier and more adipose in childhood than offspring born to women without HG,[14, 16] which can in line lead to increased Apo-B levels in later life, and therefore affect offspring cardiovascular disease risk.[43] Taking our findings into account, differences in these studies' findings might be due to differences in HG severity with studies including women with a more severe HG being more likely to find associations than studies including women with a relatively mild HG. Another explanation for differences in results could be the sample size or, more important, the number of HG-exposed offspring included. Although Koot *et al.*[13] included 6462 adolescents in their study, only 42 had mothers with HG. Their study did not find any differences in cardiometabolic risk factors due to HG. Poeran-Bahadoer *et al.*[16] and Grooten *et al.*[14] included a far larger group of women with daily vomiting and severe weight loss (respectively 463 and 533 women) and found significant differences in anthropometric measures and blood pressure. However, neither study has prospectively collected information on HG diagnosis and treatment.

Our results show that a more severe HG, in terms of lower maternal weight gain at inclusion, was associated with higher glucose levels, in the absence of increased c-peptide levels in cord blood. These findings are in line with earlier studies: research from the Dutch famine showed an increased risk of developing diabetes in later life in offspring conceived during the famine and Ayyavoo *et al.*[15] found a lower insulin sensitivity in offspring of women with HG in childhood.[44, 45] Importantly, glucose levels in cord blood are likely highly reflective of maternal blood glucose concentrations.[46] This may point to an increase in maternal hyperglycaemia in more severe HG cases, which corresponds to literature linking maternal starvation to insulin

resistance, which in turn could have independent effects on the chance of offspring adiposity and insulin resistance in later life.[47]

While our study suggests that lower maternal weight gain in HG is associated with some increased markers of cardiometabolic disease risk for offspring, treatment by early enteral tube feeding, a treatment which could potentially restore nutritional intake, did not affect offspring's cardiometabolic markers in cord blood. Given the small sample size, with just 46 women in RCT analyses, our statistical power was limited. Our study would have been able to detect only large mean differences based on treatment (for instance a mean difference of 0.11 g/L for Apo-B cord blood levels, based on using a two group t-test with 80% power and 5% two-sided significance level, assuming that the common standard deviation is 0.13 g/L[39, 41]). Alternatively, it is also possible that women received tube feeding too late or for too short a period to have a beneficial effect on offspring cardiometabolic markers in cord blood. Other treatment aspects of interest, but out of scope for this study, include the impact of adequate antiemetic treatment on symptoms and maternal weight gain, and are of interest for future studies on the efficacy of medications in the treatment of HG. For example, there is some evidence showing that starting antiemetic treatments early in pregnancy can reduce the overall severity of HG, which might have a beneficial effect on neonatal outcomes as well.[48, 49]

Conclusion

Our study showed that lower maternal weight gain was associated with higher Apo-B levels in offspring cord blood, possibly indicating an increased cardiometabolic risk in later life, although this effect is modest and the clinical relevance remains unclear. None of the other measures of HG severity were associated with any of the other endocrine and lipid measures in cord blood, which could also mean that our results are based on coincidence. Additionally, our study provided insufficient evidence to show that early enteral tube feeding amended offspring outcomes. Larger and longer follow-up studies are necessary to further evaluate the possible negative impact of HG on offspring's long term health and to assess the role of HG's disease spectrum or nutritional and other management options.

Acknowledgements

We thank all participating women of the MOTHER study and all staff who made this study possible, including staff from the Amsterdam UMC laboratory who helped us analysing frozen stored blood samples.

Funding

The MOTHER study was conducted with support of a research grant from North West Hospital Group, Alkmaar, the Netherlands (grant number 2013T085). The follow-up study analysing maternal and cord blood samples was conducted with support of a research grant from the Amsterdam Reproduction and Development (AR&D) research institute, Amsterdam UMC, the Netherlands (project number 23346).

Contribution to authorship

IJG, KN, MHK, RCP and TJR: formulated research questions and designed the study; KN and RCP carried out the study; JMJB, CR-S, HAB, DPvdH, WMH, AH, GK, SMK, JOEHvL, JL, FvdM, DP, M-JP, PJP, LvRF, RJR, HCJS, TV, BWM, MHK, IJG and RCP: performed recruitment and data collection of the original MOTHER study; HK: gave perspective on the interpretation of dietary results; KN, supervised by RCP and RvE: analysed data; KN drafted the manuscript. All authors critically reviewed the manuscript and approved the final draft.

REFERENCES

1. Einarson TR, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: a meta analysis. Journal of population therapeutics and clinical pharmacology = Journal de la therapeutique des populations et de la pharamcologie clinique. 2013;20(2):e171-83.
2. van Stuijvenberg ME, Schabort I, Labadarios D, Nel JT. The nutritional status and treatment of patients with hyperemesis gravidarum. American journal of obstetrics and gynecology. 1995;172(5):1585-91.
3. Birkeland E, Stokke G, Tangvik RJ, Torkildsen EA, Boateng J, Wollen AL, et al. Norwegian PUQE (Pregnancy-Unique Quantification of Emesis and nausea) identifies patients with hyperemesis gravidarum and poor nutritional intake: a prospective cohort validation study. PloS one. 2015;10(4):e0119962.
4. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Bleker OP. Plasma lipid profiles in adults after prenatal exposure to the Dutch famine. The American journal of clinical nutrition. 2000;72(5):1101-6.
5. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Schroeder-Tanka JM, et al. Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. Heart (British Cardiac Society). 2000;84(6):595-8.
6. Roseboom TJ, van der Meulen JH, Ravelli AC, Osmond C, Barker DJ, Bleker OP. Plasma fibrinogen and factor VII concentrations in adults after prenatal exposure to famine. British journal of haematology. 2000;111(1):112-7.
7. Juhola J, Magnussen CG, Viikari JS, Kahonen M, Hutri-Kahonen N, Jula A, et al. Tracking of serum lipid levels, blood pressure, and body mass index from childhood to adulthood: the Cardiovascular Risk in Young Finns Study. The Journal of pediatrics. 2011;159(4):584-90.
8. Kaser S, Ebenbichler CF, Wolf HJ, Sandhofer A, Stanzl U, Ritsch A, et al. Lipoprotein profile and cholesteryl ester transfer protein in neonates. Metabolism: clinical and experimental. 2001;50(6):723-8.
9. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: associations with neonatal anthropometrics. Diabetes. 2009;58(2):453-9.
10. Barjaktarovic M, Korevaar TIM, Gaillard R, de Rijke YB, Visser TJ, Jaddoe VWV, et al. Childhood thyroid function, body composition and cardiovascular function. European journal of endocrinology. 2017;177(4):319-27.
11. Maslin K, Shaw V, Brown A, Dean C, Shawe J. What is known about the nutritional intake of women with Hyperemesis Gravidarum?: A scoping review. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2021;257:76-83.
12. Veenendaal MV, van Abeelen AF, Painter RC, van der Post JA, Roseboom TJ. Consequences of hyperemesis gravidarum for offspring: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2011;118(11):1302-13.
13. Koot MH, Grooten IJ, Sebert S, Koiranen M, Jarvelin MR, Kajantie E, et al. Hyperemesis gravidarum and cardiometabolic risk factors in adolescents: a follow-up of the Northern Finland Birth Cohort 1986. BJOG : an international journal of obstetrics and gynaecology. 2017.
14. Grooten IJ, Painter RC, Pontesilli M, van der Post JA, Mol BW, van Eijsden M, et al. Weight loss in pregnancy and cardiometabolic profile in childhood: findings from a longitudinal birth cohort. BJOG : an international journal of obstetrics and gynaecology. 2015;122(12):1664-73.
15. Ayyavoo A, Derraik JG, Hofman PL, Biggs J, Bloomfield FH, Cormack BE, et al. Severe hyperemesis gravidarum is associated with reduced insulin sensitivity in the offspring in childhood. The Journal of clinical endocrinology and metabolism. 2013;98(8):3263-8.

16. Poeran-Bahadoer S, Jaddoe VWV, Gishti O, Grooten IJ, Franco OH, Hofman A, et al. Maternal vomiting during early pregnancy and cardiovascular risk factors at school age: the Generation R Study. J Dev Orig Health Dis. 2020;11(2):118-26.

17. Royal College of Obstetricians and Gynaecologists. The Management of Nausea and Vomiting of Pregnancy and Hyperemesis Gravidarum: RCOG Green-Top Guideline No. 69 2016 [updated June 2016]. Available from: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg69/.

18. Grooten IJ, Koot MH, van der Post JA, Bais JM, Ris-Stalpers C, Naaktgeboren C, et al. Early enteral tube feeding in optimizing treatment of hyperemesis gravidarum: the Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding (MOTHER) randomized controlled trial. The American journal of clinical nutrition. 2017;106(3):812-20.

19. Grooten IJ, Mol BW, van der Post JA, Ris-Stalpers C, Kok M, Bais JM, et al. Early nasogastric tube feeding in optimising treatment for hyperemesis gravidarum: the MOTHER randomised controlled trial (Maternal and Offspring outcomes after Treatment of HyperEmesis by Refeeding). BMC pregnancy and childbirth. 2016;16:22.

20. Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM. New Dutch reference curves for birthweight by gestational age. Early Hum Dev. 2009;85(12):737-44.

21. Organization for Economic Cooperation and Development (OECD). 2021 [Available from: https://www.oecd.org/about/members-and-partners/.

22. Koren G, Piwko C, Ahn E, Boskovic R, Maltepe C, Einarson A, et al. Validation studies of the Pregnancy Unique-Quantification of Emesis (PUQE) scores. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2005;25(3):241-4.

23. National Institute for Public Health and the Environment. NEVO online version 2019/6.0, RIVM, Bilthoven [13 October 2020]. Available from: https://nevo-online.rivm.nl/.

24. National Institute for Public Health and the Environment. Portie-online version 2020/1.4, RIVM, Bilthoven. [13 October 2020]. Available from: https://portie-online.rivm.nl/.

25. National Institute for Public Health and the Environment. Dutch National Food Consumption Survey 2007-2010. [Internet]. Available from: https://www.rivm.nl/documenten/vcp-2007-2010-deel-2-voedingsmiddelen-nevo-codes.

26. Brink L, Postma-Smeets A, Stafleu A, Wolvers D. Richtlijnen Schijf van Vijf. Den Haag: Stichting Voedingscentrum Nederland; 2020. 182 p.

27. Geurts M, Toxopeus I, Van Rossum C, Vennemann F, Buurma-Rethans E, Ocké M. MEMO: Achtergrondgegevens van referentievoedingen voor de Richtlijnen Schijf van Vijf 2016. Bilthoven: National Institute for Public Health and the Environment; 2016.

28. Noordam C, Rotteveel J, Schroor EJ. Werkboek Kinderendocrinologie (workbook pediatric endocrinology). Amsterdam: VU boekhandel; 2010. Available from: https://www.nvk.nl/themas/kwaliteit/werkboeken.

29. Federation of Medical Specialists. Schildklierfunctiestoornissen (thyroid disfunction) 2012 [Available from: https://richtlijnendatabase.nl/richtlijn/schildklierfunctiestoornissen/schildklierfunctiestoornissen_-_korte_beschrijving.html.

30. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient–clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ Open. 2021;11(1):e041254.

31. van Vliet R, Bink M, Polman J, Suntharan A, Grooten I, Zwolsman SE, et al. Patient Preferences and Experiences in Hyperemesis Gravidarum Treatment: A Qualitative Study. J Pregnancy. 2018;2018:5378502-.

32. Ebrahimi N, Maltepe C, Bournissen F, Koren G. Nausea and vomiting of pregnancy: using the 24-hour Pregnancy-Unique Quantification of Emesis (PUQE-24) scale. J obstet Gynaecol Can. 2009;31(9):803-7.

33. Lacasse A, Rey E, Ferreira E, Morin C, Berard A. Validity of a modified Pregnancy-Unique Quantification of Emesis and Nausea (PUQE) scoring index to assess severity of nausea and vomiting of pregnancy. American journal of obstetrics and gynecology. 2008;198(1).

34. Crozier SR, Inskip HM, Godfrey KM, Robinson SM. Dietary patterns in pregnant women: a comparison of food-frequency questionnaires and 4 d prospective diaries. British Journal of Nutrition. 2008;99(4):869-75.

35. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014;36:e2014009.

36. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. Bmj. 1995;310(6973):170.

37. Wilkins JT, Li RC, Sniderman A, Chan C, Lloyd-Jones DM. Discordance Between Apolipoprotein B and LDL-Cholesterol in Young Adults Predicts Coronary Artery Calcification. Journal of the American College of Cardiology. 2016;67(2):193-201.

38. Walldius G, Jungner I, Holme I, Aastveit AH, Kolar W, Steiner E. High apolipoprotein B, low apolipoprotein A-I, and improvement in the prediction of fatal myocardial infarction (AMORIS study): a prospective study. Lancet (London, England). 2001;358(9298):2026-33.

39. Agrawal A, Shrivastava J, Dwivedi R, Siddiqui M. Assessment of serum apolipoprotein B and apolipoprotein A-1 and their ratio in healthy full term small for gestational age newborns. Journal of Neonatal-Perinatal Medicine. 2017;10:49-53.

40. Katragadda T, Mahabala RS, Shetty S, Baliga S. Comparison of Cord Blood Lipid Profile in Preterm Small for Gestational Age and Appropriate for Gestational Age Newborns. J Clin Diagn Res. 2017;11(1):SC05-SC7.

41. Kharb S, Kaur R, Singh V, Sangwan K. Birth weight, cord blood lipoprotein and apolipoprotein levels in Indian newborns. Int J Prev Med. 2010;1(1):29-33.

42. Bakker A, Weel J, Wolthuis A. Laboratoriumdiagnostiek bij kinderen: een praktische handleiding. (Laboratory diagnostics in children: a practical manual). Houten: Prelum Uitgevers; 2015.

43. Bays HE, Toth PP, Kris-Etherton PM, Abate N, Aronne LJ, Brown WV, et al. Obesity, adiposity, and dyslipidemia: a consensus statement from the National Lipid Association. J Clin Lipidol. 2013;7(4):304-83.

44. de Rooij SR, Painter RC, Roseboom TJ, Phillips DI, Osmond C, Barker DJ, et al. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. Diabetologia. 2006;49(4):637-43.

45. Ravelli AC, van der Meulen JH, Michels RP, Osmond C, Barker DJ, Hales CN, et al. Glucose tolerance in adults after prenatal exposure to famine. Lancet (London, England). 1998;351(9097):173-7.

46. Koh D, Hume R, Eisenhofer G, Ogston S, Watson J, Williams F. Maternal and fetal factors which influence cord blood glucose levels in term infants delivered by cesarean section. Journal of perinatal medicine. 2015;43(3):339-46.

47. Soeters MR, Soeters PB. The evolutionary benefit of insulin resistance. Clin Nutr. 2012;31(6):1002-7.

48. Dean C. Helping women prepare for hyperemesis gravidarum. British Journal of Midwifery. 2014;22(12):847-52.

49. Koren G, Maltepe C. Pre-emptive therapy for severe nausea and vomiting of pregnancy and hyperemesis gravidarum. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2004;24(5):530-3.

APPENDIX A

Detailed methods dietary intake and corresponding 24-hour energy intake calculations

During the MOTHER study, participants kept a weekly self-reported 24-hour food diary. The food diary consisted of a list of pre-specified food products, as shown below in **Figure A1**. Participants were able to fill in their food intake in pieces or millilitres (ml) at six different moments: breakfast, lunch, dinner and three in-between mealtime snacks. Besides the pre-specified food item list, participants could also specify food products, which were not included in the standard checklist below 'other'.

In collaboration with a dietician, databases of the Dutch National Institute for Public Health and the Environment (RIVM) were used to determine portion sizes and energy content of reported food products. First, we determined standard portion sizes of each of the listed food products by using the database containing portion sizes.[1] Secondly, the Dutch Food Composition Database 2019 was used to extract the corresponding energy content in kilocalories (kcal) per 100 gram and to calculate the energy content per portion size.[2]

Few of the listed food products were not specified in details, which made it difficult to find corresponding NEVO-codes in the Dutch Food Composition Database. For some of these products, an average NEVO-code was available, such as for meat products and sweet spreads. For many products however, average NEVO-codes were missing. In these cases, we used the most recent available Dutch National Food Consumption Survey (DNFCS) of 2012-2016 to calculate weighted averages of the five most consumed food products in the corresponding food category among adult women between 18 and 50 years.[3]

Furthermore, a 'hot meal' was also included in the pre-specified food diary, but was not further described in detail. The Dutch National food Consumption Survey 2012-2016 measured that 34.1% of the daily energy intake was consumed during diner and that, on average, the energy intake of adult women between 18 and 50 years was 1898 calories.[3] Therefore, we assumed that one portion size of a consumed hot meal contained 647 calories. **Supplement Table A1** shows each of the listed food products with corresponding portion sizes and energy content per 100 gram as well as per portion size.

Lastly, a 24-hour energy intake was calculated, adding up all caloric quantities of each of the reported food products per woman. For women reporting that they had no intake at all, the

24-hour energy intake was set at zero. For this study, the 24-hour energy intake at inclusion was used as a measure of HG severity and as predictor in regression analysis.

REFERENCES APPENDIX

1. National Institute for Public Health and the Environment. Portie-online version 2020/1.4, RIVM, Bilthoven. [13 October 2020]. Available from: https://portie-online.rivm.nl/.
2. National Institute for Public Health and the Environment. NEVO online version 2019/6.0, RIVM, Bilthoven [13 October 2020]. Available from: https://nevo-online.rivm.nl/.
3. National Institute for Public Health the Environment, van Rossum C, Nelis K, Wilson C, Ocké M. National dietary survey in 2012-2016 on the general population aged 1-79 years in the Netherlands. EFSA Supporting Publications. 2018;15(9):1488E

Product	Breakfast		In-between moment #1		Lunch		In-between moment #2		Dinner		In-between moment #3	
	piece	ml	piece	ml	piece	ml	piece	ml	piece	ml	piece	ml
Biscuit/cracker												
Bread												
Currant bun												
Ginger bread												
Margarine/ butter												
Cheese												
Meat products												
Sweet spreads												
Porridge												
Muesli + yoghurt												
Tea/coffee												
Sugar												
Semi-skimmed milk												
Whole milk												
Buttermilk												
Chocolate milk												
Fruit juice												
Lemonade												
Soft drink (with sugar)												
Water/broth												
Cup a soup												
Nutridrink												
Soup (thick)												
Hot meal 1 portion												
Hot meal ½ portion												
Gravy												
Fruit compote												
Yoghurt												
Custard, pudding												
Fruit												
Other:												
.....												
.....												

Figure A1. 24-hour food diary

Supplemental Table A1. List of food products reported in 24-hour food diary with corresponding portion sizes and energy content

Product	NEVO code	General portion sizes	Energy content *(kcal/100 gram)*	% of top 5 of DNFCS [1]	Kcal per food product portion
Biscuit/cracker		11.8 g			**46**
Crispbread whole meal	1779		363	47.0%	
Crispbread sesame	975		421	19.8%	
Crispbread gold-brown	596		410	11.8%	
Crisp bakes Dutch	227		408	10.8%	
Cracker mini unflavored	5095		442	10.7%	
Bread		35 g			**88**
Bread brown wheat	236		236	24.8%	
Bread whole meal fine	2811		233	24.0%	
Bread multigrain average with seeds	2350		261	23.3%	
Roll white soft	230		262	14.5%	
Roll white hard	2795		277	13.4%	
Currant bun	2803	50 g	268	NA	**134**
Ginger bread		30 g			**92**
Dutch spiced cake	240		308	65.5%	
Dutch spiced cake whole meal	925		308	13.9%	
Dutch spiced cake low sugar	2329		287	10.5%	
Dutch spiced cake with nuts	2397		337	6.0%	
Dutch spiced cake with sugar candy	2398		312	4.1%	
Margarine/butter		6 g			**27**
Low fat margarine 40% fat (<17 g), salted	2059		356	49.6%	
Butter unsalted	310		737	16.2%	
Low fat margarine (Blue Band Goede Start)	1961		349	12.3%	
Low fat margarine (tub Becel Light)	1839		285	11.8%	
Margarine 80% fat (>24 g), salted	2063		719	9.2%	
Cheese		20 g			**74**
Cheese Gouda 48+, age 8 weeks-4 months	2757		370	41.0%	
Cheese Gouda 48+, age 4-8 weeks	2756		364	24,8%	
Cheese Gouda 48+, age 4-7 months	2758		377	18.0%	
Cheese Edam 40+	511		324	8.6%	
Cheese Gouda 48+, age 10-12 months	2759		414	7.6%	

Supplemental Table A1. List of food products reported in 24-hour food diary with corresponding portion sizes and energy *content*

Product	NEVO code	General portion sizes	Energy content *(kcal/100 gram)*	% of top 5 of DNFCS [1]	Kcal per food product portion
Meat products	344	15 g	236	NA	**35**
Sweet spreads	464	20 g	393	NA	**79**
Porridge		150 g			**115**
Porridge rice	298		86	19.1%	
Porridge semolina	1722		93	9.8%	
Porridge oatmeal prepared with semi-skimmed milk	3050		69	9.8%	
Porridge oatmeal prepared with whole milk	288		85	9.4%	
Porridge barley with raisins (Bessola)	605		73	6.5%	
Muesli + yoghurt	2278	125 g	110	NA	**138**
Tea/coffee		175 ml			**2**
Tea prepared	645		0	68.2%	
Coffee prepared	644		1	28.2%	
Cappuccino freshly made	2476		31	1.6%	
Coffee with milk from vending machine	2648		12	1.2%	
Cappuccino instant prepared	2477		37	0.8%	
Sugar		3.9 g			**16**
Sugar granulated	377		400	96.3%	
Castor sugar white	375		396	2.0%	
Castor sugar brown	374		396	1.7%	
Semi-skimmed milk	286	200 ml	45	NA	**90**
Whole milk	279	200 ml	61	NA	**122**
Buttermilk	289	200 ml	30	NA	**60**
Chocolate milk		200 ml			**154**
Whole chocolate milk	272		89	34.9%	
Semi-skimmed chocolate milk	1464		77	26.9%	
Hot chocolate from vending machine	2760		66	17.2%	
Semi-skimmed chocolate milk with sweetener	1970		56	10.5%	
Semi-skimmed chocolate milk with sweetened cacao powder	2495		78	10.5%	
Fruit juice		200 ml			**90**
Freshly squeezed orange juice	2755		44	32.8%	
Pasteurized orange juice	410		45	30.6%	
Multi-fruit juice	2507		47	17.0%	

Supplemental Table A1. List of food products reported in 24-hour food diary with corresponding portion sizes and energy *content*

Product	NEVO code	General portion sizes	Energy content *(kcal/100 gram)*	% of top 5 of DNFCS [1]	Kcal per food product portion
Apple juice	383		46	14.7%	
Apple nectar juice	3218		37	4.9%	
Lemonade		40 g [2]			**60**
Fruit drink concentrated	463		233	26.1%	
Fruit drink concentrated light	2289		8	24.5%	
Fruit drink concentrated (Karvan Cevitam)	1810		238	24.0%	
Fruit drink concentrated with sugar & sweeteners 40-45g	2287		170	15.7%	
Fruit drink concentrated with sugar & sweeteners 10-15g	2831		47	9.7%	
Soft drink (with sugar)		200 ml			**73**
Cola with caffeine	395		41	54.8%	
Ice tea	2086		31	15.1%	
Soft drink without caffeine	400		38	19.6%	
Ice tea with sugar & sweetener	2088		19	8.5%	
Soft drink with sugar, sweetener (5-8 g) & caffeine	2665		27	1.7%	
Water/broth	3192	175 ml	6	NA	**11**
Cup a soup	2932	175 ml	40	NA	**70**
Nutridrink [3]	-	200 ml	150	NA	**300**
Soup (thick) [4]		250 ml			**117**
Soup thickened with vegetables	763		36	33.3%	
Soup thickened with meat (beef/ chicken)	764		64	33.3%	
Soup thickened, no filling	2561		41	33.3%	
Hot meal 1 portion [5]	NA	NA	NA	NA	**647**
Hot meal ½ portion [5]	NA	NA	NA	NA	**324**
Gravy		25 g			**77**
Gravy 50% fat , prepared without instant gravy powder	2459		439	34.3%	
Instant gravy 25% fat, thickened	2451		233	20.0%	
Instant gravy 50% fat, thickened	2461		429	18.1%	
Instant gravy 0% fat	2650		28	14.3%	
Gravy 25% fat, prepared without instant gravy powder	2588		219	13.3%	
Fruit compote	179	100 g	76	NA	**76**
Yoghurt		120 g			**60**
Yoghurt low fat	301		37	46.6%	

Supplemental Table A1. List of food products reported in 24-hour food diary with corresponding portion sizes and energy *content*

Product	**NEVO code**	**General portion sizes**	**Energy content *(kcal/100 gram)***	**% of top 5 of DNFCS [1]**	**Kcal per food product portion**
Yoghurt full fat	278		58	19.5%	
Yoghurt reduced fat	1.502		50	15.2%	
Yoghurt low fat with fruit	284		73	11.0%	
Yoghurt vanilla reduced fat	1.721		78	7.7%	
Custard, pudding		150 ml			**119**
Quark low fat	305		58	29.4%	
Custard several flavors full fat	1720		95	22.4%	
Custard vanilla full fat	282		93	18.2%	
Quark low fat with fruit/vanilla with sweetener	2246		43	17.1%	
Quark reduced fat with fruit/ vanilla	917		130	12.9%	
Fruit		140 g			**71**
Fresh fruit average, including citrus	172		48	82.2%	
Fresh fruit average, excluding citrus	173		62	17.8%	

[1] In case no corresponding NEVO code was available for a specific food product, we calculated a weighted average calories of the 5 most consumed food products in the DNFCS 2012-2016 among adult women between 18 and 50 years. [2] Portion size of undiluted lemonade. [3] Nutritional drink. [4] Soup was not present in the DNFCS, so we calculated an average of the three available thick soup NEVO codes. [5] The energy intake for a hot meal was calculated using the DNFCS 2012-2016, which measured that 34.1% of the daily energy intake was consumed during diner. On average, the energy intake of adult women between 18 and 50 years was 1898 calories, leading to the assumption that a full size hot meal contained 647 calories. **Abbreviations**: DNFCS: Dutch National Food Consumption Survey, g: gram, kcal: kilocalories, ml: milliliters.

Supplemental Table S1. Sensitivity analysis to assess differences between in- and excluded women in this study

	Included women (cord blood available)	**Excluded women** (no cord blood available)	**Missing**	***P***
Demographics	**n=81**	**n=134**		
Age (years)	28.3 ± 4.2	29.2 ± 5.2	0.0%	0.21
Prepregnancy weight (kg)	73.4 ± 15.2	69.8 ± 14.8	2.3%	0.09
Prepregnancy BMI (kg/m^2)	25.5 ± 4.8	24.9 ± 4.9	3.7%	0.37
Non-western ethnicity	18 (22.2%)	35 (26.1%)	18.1%	0.18
Education level			33.0%	0.65
- Primary or secondary	38 (46.9%)	49 (36.5%)		
- Higher	27 (33.3%)	30 (22.4%)		
Primigravida	27 (33.3%)	37 (27.6%)	0.0%	0.37
Maternal diabetes (gestational, type I or II)	4 (4.9%)	10 (7.5%)	0.0%	0.47
Maternal thyroid disease	4 (4.9%)	5 (3.7%)	0.0%	0.73
Current smoker	7 (8.6%)	4 (3.0%)	5.6%	0.11
Gestational age at onset of HG symptoms (weeks)	6.0 (5.0-7.0)	6.0 (5.5-7.0)	23.3%	0.88
Gestational age at inclusion (weeks)	9.0 (7.0-11.0)	8.5 (7.0-11.0)	0.0%	0.36
Measures of HG severity				
Weight change (kg)	-3.0 ± 3.8	-2.8 ± 4.3	2.8%	0.73
24-hour energy intake at inclusion (kcal)	435 (152-1021)	547 (227-1087)	35.3%	0.36
Below recommended daily intake (<1870 kcal)	62 (76.5%)	68 (50.7%)	35.3%	0.74
PUQE-24 at inclusion	10.4 ± 3.2	9.6 ± 3.4	37.2%	0.17
Average PUQE-24 in the first three weeks after inclusion	8.7 ± 2.9	8.2 ± 2.8	36.3%	0.30
Total duration of hospital admissions (days)	5.0 (3.0-7.0)	5.0 (3.0-8.0)	0.0%	0.82
Readmitted	27 (33.3%)	44 (32.8%)	0.0%	0.94
Readmission after the first trimester	21 (25.9%)	22 (16.4%)	0.0%	0.09
Perinatal outcomes				
Birth weight (grams)	3315 ± 513	3233 ± 760	1.9%	0.35
SGA (birth weight <10th percentile)	6 (7.5%)	14 (10.4%)	3.3%	0.41
Prematurity (< 37 weeks)	5 (6.2%)	15 (11.2%)	1.9%	0.20
Apgar score <7 at 5 min	2 (2.5%)	4 (3.0%)	2.3%	1.00
Fetal sex (female)	46 (56.8%)	74 (55.2%)	1.9%	0.99
Maternal blood measurements				
Apolipoprotein A1 (g/L)	1.43 ± 0.31	1.44 ± 0.29	28.8%	0.87
Apolipoprotein B (g/L)	0.80 ± 0.25	0.75 ± 0.21	28.8%	0.21

Supplemental Table S1. Sensitivity analysis to assess differences between in- and excluded women in this study. *Continued*

	Included women (cord blood available)	**Excluded women** (no cord blood available)	**Missing**	***P***
HDL (mmol/L)	1.33 ± 0.33	1.39 ± 0.31	28.8%	0.26
LDL (mmol/L)	2.43 ± 0.83	2.31 ± 0.70	28.8%	0.32
Total cholesterol (mmol/L)	4.21 ± 1.07	4.07 ± 0.88	28.8%	0.36
Triglycerides (mmol/L)	0.83 (0.64-1.18)	0.75 (0.61-1.02)	28.8%	**0.045**
FT4 (pmol/L)	19.79 ± 5.81	19.35 ± 4.96	29.3%	0.62

Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). *P*-values <0.05 were considered statistically significant. Abbreviations: BMI: body mass index, HDL: High Density Lipoprotein, HG: hyperemesis gravidarum, FT4: free thyroxine, LDL: Low Density Lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms. SGA: small for gestational age. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight.

Supplemental Table S2. Associations between maternal and perinatal factors and endocrine and lipid measures in cord blood in univariable regression analysis

	Apo-A1			Apo-B			HDL			LDL			Total cholesterol		
	β	95% CI	*P*	*β*	95% CI	*P*	*β*	95% CI	*P*	*β*	95% CI	*P*	*β*	95% CI	*P*
- Maternal age (years)	-0.00	-0.01;0.01	0.72	0.00	-0.00;0.01	0.36	-0.00	-0.02;0.01	0.72	0.00	-0.03;0.03	0.86	-0.00	-0.04;0.04	0.93
- Pre-pregnancy BMI (kg)	0.00	-0.01;0.01	0.97	0.00	-0.01;0.01	0.62	-0.00	-0.01;0.01	0.78	0.01	-0.02;0.04	0.41	0.01	-0.03;0.05	0.59
- Ethnicity (western or not)	-0.05	-0.15;0.06	0.37	-0.04	-0.13;0.04	0.30	-0.09	-0.21;0.04	0.16†	-0.20	-0.50;0.10	0.18†	-0.32	-0.71;0.08	0.11†
- Highest finished education	0.02	-0.08;0.12	0.62	-0.02	-0.10;0.06	0.66	0.04	-0.09;0.16	0.55	-0.09	-0.37;0.21	0.56	-0.09	-0.48;0.29	0.62
- Current smoker	0.02	-0.14;0.17	0.82	0.13	0.01;0.25	0.03*	-0.03	-0.22;0.17	0.77	0.46	0.06;0.87	0.03*	0.48	-0.08;1.04	0.09†
- Maternal diabetes (gestational, type I or II)	0.06	-0.13;0.24	0.55	0.20	0.06;0.33	0.01*	0.06	-0.17;0.29	0.62	0.77	0.23;1.31	0.01*	0.85	0.11;1.60	0.03*
- Preterm birth	-0.00	-0.17;0.16	0.98	0.18	0.05;0.30	0.01*	0.09	-0.12;0.30	0.40	-0.66	-1.08;-0.24	<0.01*	0.75	0.17;1.34	0.01*
- Small for gestational age	0.04	-0.12;0.20	0.64	-0.01	-0.14;0.12	0.88	-0.03	-0.24;0.18	0.78	0.00	-0.45;0.45	1.00	0.01	-0.60;0.63	0.97
- Fetal sex	-0.02	-0.10;0.06	0.65	-0.00	-0.07;0.06	0.90	-0.03	0.13;0.08	0.59	0.02	-0.21;0.26	0.84	-0.04	-0.36;0.27	0.78

	Triglycerides ∫			C-peptide			Glucose			FT4		
	β	95% CI	*P*	*β*	95% CI	*P*	*β*	95% CI	*P*	*β*	95% CI	*P*
- Maternal age (years)	0.80	-2.57;4.39	0.62	0.00	-0.01;0.01	0.49	-0.05	-0.12;0.02	0.14†	-0.07	-0.19;0.06	0.30
- Pre-pregnancy BMI (kg)	-0.40	-3.34;2.63	0.79	0.01	0.00;0.01	0.04*	0.00	-0.06;0.07	0.90	-0.05	-0.16;0.06	0.37
- Ethnicity (western or not)	-4.02	-31.48;34.58	0.81	-0.02	-0.11;0.06	0.57	-0.01	-0.75;0.72	0.97	-0.81	-2.08;0.46	0.21
- Highest finished education	-5.64	-31.34;29.69	0.72	-0.01	-0.08;0.07	0.91	0.13	-0.58;0.85	0.71	0.95	-0.28;2.17	0.13†
- Current smoker	13.54	-32.50;90.98	0.63	-0.05	-0.17;0.07	0.39	-0.30	-1.32;0.71	0.55	-1.31	-3.06;0.45	0.14†
- Maternal diabetes (gestational, type I or II)	10.41	-40.49;104.83	0.75	0.00	-0.15;0.15	0.96	0.51	-0.77;1.79	0.43	-0.96	-3.42;1.49	0.44
- Maternal hypothyroidism	-	-	-	-	-	-	-	-	-	-1.30	-5.49;2.89	0.54
- Maternal hyperthyroidism	-	-	-	-	-	-	-	-	-	-1.01	-4.00;1.97	0.50
- Preterm birth	-11.57	-49.29;54.34	0.66	0.02	-0.11;0.15	0.75	-1.05	-2.19;0.10	0.07†	-1.24	-3.37;0.90	0.25
- Small for gestational age	37.30	-21.34;139.41	0.26	-0.13	-0.26;0.01	0.06†	0.17	-1.01;1.34	0.78	-0.48	-2.24;1.46	0.62
- Fetal sex	-3.44	-27.39;28.27	0.81	0.03	-0.04;0.09	0.44	-0.12	-0.71;0.47	0.68	0.01	-1.07;1.10	0.98

Significant associations with *P*-values <0.05 are marked with an asterisk (*). Associations with *a P*-value <0.20 are marked with a dagger (†) and are included as confounders in multivariable regression analysis assessing the association between measures of HG severity and endocrine and lipid measures in cord blood. β is the unstandardized regression coefficient, 95% CI is the 95% confidence interval. ∫ = log transformed, back transformed and expressed in % of difference. **Abbreviations:** Apo-A1: apolipoprotein A1, Apo-B: apolipoprotein B, FT4: free thyroxine, HDL: high density lipoprotein, LDL: low density lipoprotein. Small for gestational age is defined as birth weight below the 10th percentile.

Supplemental Table S3. RCT analysis according to intention-to-treat

	Early tube feeding	Standard care	Missing	*P*
Demographics	**n=23**	**n=23**		
Age (years)	27.8 ± 4.3	28.4 ± 3.8	0.0%	0.64
Prepregnancy weight (kg)	75.7 ± 15.4	76.3 ± 16.6	2.2%	0.90
Prepregnancy BMI (kg/m^2)	27.2 ± 5.3	25.4 ± 4.7	2.2%	0.24
Non-western ethnicity	3 (13.0%)	7 (30.4%)	10.9%	0.16
Education level			17.4%	0.32
- Primary or secondary	10 (43.5%)	14 (60.9%)		
- Higher	8 (34.8%)	6 (26.1%)		
Primigravida	9 (39.1%)	8 (34.8%)	0.0%	0.76
Maternal diabetes (gestational, type I or II)	3 (13.0%)	1 (4.3%)	0.0%	0.61
Maternal thyroid disease	0 (0.0%)	2 (8.7%)	0.0%	0.49
Current smoker	3 (13.0%)	3 (13.0%)	4.3%	1.00
Gestational age at onset of HG symptoms (wks)	6.0 (5.0-7.5)	6.0 (5.0-6.1)	19.6%	0.52
Gestational age at inclusion (wks)	9.0 (8.0-11.0)	9.0 (7.0-12.0)	0.0%	0.52
Measures of HG severity				
Weight change (kg)	-2.6 ± 4.1	-5.0 ± 4.0	2.2%	0.06
24-hour energy intake at inclusion (kcal)	131 (43-417)	483 (283-1068)	17.4%	**0.04**
Below recommended daily intake (<1870 kcal)	18 (78.3%)	20 (87.0%)	17.4%	1.00
PUQE-24 at inclusion	11.4 ±2.9	10.7 ±3.2	21.7%	0.48
Average PUQE-24 in the first three weeks after inclusion	9.0 ± 3.6	8.7 ± 2.9	17.4%	0.74
Total duration of hospital admissions (days)	6.0 (3.0-10.0)	4.0 (3.0-5.0)	0.0%	0.18
Readmitted	9 (39.1%)	7 (30.4%)	0.0%	0.54
Readmission after the first trimester	7 (30.4%)	7 (30.4%)	0.0%	1.00
Perinatal outcomes				
Birth weight (grams)	3235 ± 562	3498 ± 485	0.0%	0.10
SGA (birth weight <10th percentile)	1 (4.3%)	1 (4.3%)	0.0%	1.00
Prematurity (< 37 weeks)	2 (8.7%)	2 (8.7%)	0.0%	1.00
Apgar score <7 at 5 min	1 (4.3%)	1 (4.3%)	0.0%	1.00
Fetal sex (female)	13 (56.5%)	13 (56.5%)	0.0%	1.00
Maternal blood measurements				
Apolipoprotein A1 (g/L)	1.52 ± 0.25	1.45 ± 0.30	13.0%	0.43
Apolipoprotein B (g/L)	0.82 ± 0.28	0.84 ± 0.27	13.0%	0.81

Supplemental Table S3. RCT analysis according to intention-to-treat. *Continued*

	Early tube feeding	**Standard care**	**Missing**	***P***
HDL (mmol/L)	1.38 ± 0.26	1.33 ± 0.29	13.0%	0.56
LDL (mmol/L)	2.46 ± 0.89	2.58 ± 0.90	13.0%	0.67
Total cholesterol (mmol/L)	4.31 ± 1.03	4.37 ± 1.17	13.0%	0.85
Triglycerides (mmol/L)	0.86 (0.70-1.59)	0.86 (0.70-1.18)	13.0%	1.00
FT4 (pmol/L)	20.38 ± 8.67	19.44 ± 3.88	13.0%	0.66
Cord blood measurements				
Apolipoprotein A1 (g/L)	0.89 ± 0.16	0.91 ± 0.22	6.5%	0.70
Apolipoprotein B (g/L)	0.30 ± 0.22	0.28 ± 0.12	6.5%	0.68
HDL (mmol/L)	0.84 ± 0.19	0.86 ± 0.26	6.5%	0.72
LDL (mmol/L)	1.05 ± 0.78	0.94 ± 0.38	13.0%	0.56
Total cholesterol (mmol/L)	2.09 ± 0.93	1.96 ± 0.62	10.9%	0.59
Triglycerides (mmol/L)	0.40 (0.18-0.73)	0.30 (0.17-0.44)	8.7%	0.15
Glucose (mmol/L)	4.26 ±1.35	5.06 ± 1.28	10.9%	0.06
C-peptide (nmol/L)	0.20 ± 0.10	0.25 ± 0.18	4.3%	0.26
FT4 (pmol/L)	16.52 ±2.40	17.59 ± 1.75	21.7%	0.13

P-values <0.05 are considered statistically significant and marked in bold. Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). Abbreviations: BMI: body mass index, HDL: High Density Lipoprotein, HG: Hyperemesis Gravidarum, FT4: free thyroxine, LDL: Low Density Lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms. SGA: small for gestational age. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight.

Supplemental Table S4. RCT analysis according to per protocol

	Early tube feeding	Standard care	Missing	*P*
Demographics	**n=13**	**n=20**		
Age (years)	27.6 ± 4.6	28.7 ±3.6	0.0%	0.46
Prepregnancy weight (kg)	76.4 ± 17.1	74.7 ± 16.6	3.0%	0.79
Prepregnancy BMI (kg/m^2)	26.8 ± 4.6	24.9 ± 4.5	3.0%	0.26
Non-western ethnicity	1 (7.7%)	7 (35.0%)	9.1%	0.10
Education level			12.1%	0.32
- Primary or secondary	6 (46.2%)	12 (60.0%)		
- Higher	5 (38.5%)	6 (30.0%)		
Primigravida	7 (53.8%)	8 (40.0%)	0.0%	0.44
Maternal diabetes (gestational, type I or II)	2 (15.4%)	1 (5.0%)	0.0%	0.55
Maternal thyroid disease	0 (0.0%)	1 (5.0%)	0.0%	1.00
Current smoker	2 (15.4%)	3 (15.0%)	6.1%	1.00
Gestational age at onset of HG symptoms (wks)	5.8 (5.0-7.0)	6.0 (5.5-6.3)	12.1%	0.81
Gestational age at inclusion (wks)	8.0 (7.5-10.5)	9.0 (7.3-12.8)	0.0%	0.48
Measures of HG severity				
Weight change (kg)	-2.7 ± 4.6	-4.0 ± 1.9	3.0%	0.34
24-hour energy intake at inclusion (kcal)	82 (15-394)	605 (326-1129)	12.1%	**<0.01**
Below recommended daily intake (<1870 kcal)	11 (84.6%)	18 (90.0%)	12.1%	1.00
PUQE-24 at inclusion	12.5 ± 2.2	10.3 ± 3.3	21.2%	0.07
Average PUQE-24 in the first three weeks after inclusion	9.0 ± 3.0	8.3 ± 3.1	12.1%	0.55
Total duration of hospital admissions (days)	6.0 (3.0-10.5)	4.0 (2.3-5.0)	0.0%	0.10
Readmitted	6 (46.2%)	4 (20.0%)	0.0%	0.14
Readmission after the first trimester	3 (23.1%)	6 (30.0%)	0.0%	1.00
Perinatal outcomes				
Birth weight (grams)	3356 ± 696	3420 ± 461	0.0%	0.75
SGA (birth weight <10th percentile)	1 (7.7%)	1 (5.0%)	0.0%	1.00
Prematurity (< 37 weeks)	2 (15.4%)	2 (10.0%)	0.0%	1.00
Apgar score <7 at 5 min	1 (7.7%)	1 (5.0%)	0.0%	1.00
Fetal sex (female)	7 (53.8%)	12 (60.0%)	0.0%	0.73
Maternal blood measurements				
Apolipoprotein A1 (g/L)	1.48 ± 0.26	1.48 ± 0.32	15.2%	0.95
Apolipoprotein B (g/L)	0.78 ± 0.24	0.82 ± 0.28	15.2%	0.71

Supplemental Table S4. RCT analysis according to per protocol. *Continued*

	Early tube feeding	Standard care	Missing	*P*
HDL (mmol/L)	1.37 ± 0.29	1.36 ± 0.30	15.2%	0.91
LDL (mmol/L)	2.36 ± 0.79	2.44 ± 0.88	15.2%	0.80
Total cholesterol (mmol/L)	4.18 ± 0.93	4.28 ± 1.22	15.2%	0.80
Triglycerides (mmol/L)	0.78 (0.59-1.52)	0.88 (0.71-1.18)	15.2%	0.71
FT4 (pmol/L)	18.12 ± 2.05	19.14 ± 4.10	15.2%	0.39
Cord blood measurements				
Apolipoprotein A1 (g/L)	0.93 ± 0.18	0.92 ± 0.22	3.0%	0.87
Apolipoprotein B (g/L)	0.33 ± 0.28	0.27 ± 0.11	3.0%	0.44
HDL (mmol/L)	0.91 ± 0.19	0.89 ± 0.27	3.0%	0.80
LDL (mmol/L)	1.17 ± 0.99	0.95 ± 0.33	9.1%	0.51
Total cholesterol (mmol/L)	2.26 ±1.13	1.99 ± 0.57	6.1%	0.37
Triglycerides (mmol/L)	0.42 (0.21-0.72)	0.30 (0.17-0.42)	6.1%	0.24
Glucose (mmol/L)	4.52 ± 1.46	4.82 ± 0.89	9.1%	0.54
C-peptide (nmol/L)	0.23 ± 0.12	0.25 ± 0.19	3.0%	0.79
FT4 (pmol/L)	16.29 ± 2.86	17.58 ± 1.83	12.1%	0.15

P-values <0.05 are considered statistically significant and marked in bold. Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). Abbreviations: BMI: body mass index. HDL: High Density Lipoprotein. HG: hyperemesis gravidarum. FT4: free thyroxine. LDL: Low Density Lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms. SGA: small for gestational age. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight.

Supplemental Table S5. RCT analysis according to as treated

	Early tube feeding	Standard care	Missing	*P*
Demographics	**n=15**	**n=20**		
Age (years)	27.2 ± 4.7	28.7 ± 3.6	0.0%	0.29
Prepregnancy weight (kg)	78.9 ± 17.2	74.7 ± 16.6	4.0%	0.48
Prepregnancy BMI (kg/m^2)	27.5 ± 4.6	24.9 ± 4.5	4.0%	0.11
Non-western ethnicity	1 (6.7%)	7 (35.0%)	16.0%	0.10
Education level			20.0%	0.54
- Primary or secondary	7 (46.7%)	12 (60.0%)		
- Higher	5 (33.3%)	6 (30.0%)		
Primigravida	7 (46.7%)	8 (40.0%)	0.0%	0.69
Maternal diabetes (gestational, type I or II)	2 (13.3%)	1 (5.0%)	0.0%	0.57
Maternal thyroid disease	1 (6.7%)	1 (5.0%)	0.0%	1.00
Current smoker	2 (13.3%)	3 (15.0%)	5.7%	1.00
Gestational age at onset of HG symptoms (wks)	5.8 (5.0-7.0)	6.0 (5.5-6.3)	24.0%	0.81
Gestational age at inclusion (wks)	8.0 (7.0-10.0)	9.0 (7.3-12.8)	0.0%	0.27
Measures of HG severity				
Weight change (kg)	-4.5 ± 6.4	-4.0 ± 1.9	4.0%	0.80
24-hour energy intake at inclusion (kcal)	76 (8-387)	605 (326-1129)	14.3%	**<0.01**
Below recommended daily intake (<1870 kcal)	12 (80.0%)	18 (90.0%)	14.3%	1.00
PUQE-24 at inclusion	12.6 ± 2.0	10.3 ± 3.3	28.0%	**0.04**
Average PUQE-24 in the first three weeks after inclusion	9.3 ± 2.9	8.3 ± 3.1	11.4%	0.40
Total duration of hospital admissions (days)	6.0 (3.0-10.0)	4.0 (2.3-5.0)	0.0%	**0.04**
Readmitted	8 (53.3%)	4 (20.0%)	0.0%	**0.04**
Readmission after the first trimester	3 (20.0%)	6 (30.0%)	0.0%	0.70
Perinatal outcomes				
Birth weight (grams)	3438 ± 692	3420 ± 461	0.0%	0.93
SGA (birth weight <10th percentile)	1 (6.7%)	1 (5.0%)	0.0%	1.00
Prematurity (< 37 weeks)	2 (13.3%)	2 (10.0%)	0.0%	1.00
Apgar score <7 at 5 min	1 (6.7%)	1 (5.0%)	0.0%	1.00
Fetal sex (female)	8 (53.3%)	12 (60.0%)	0.0%	0.69
Maternal blood measurements				
Apolipoprotein A1 (g/L)	1.45 ± 0.26	1.48 ± 0.32	14.3%	0.80
Apolipoprotein B (g/L)	0.82 ± 0.25	0.82 ± 0.28	14.3%	0.98

Supplemental Table S5. RCT analysis according to as treated. *Continued*

	Early tube feeding	Standard care	Missing	*P*
HDL (mmol/L)	1.33 ± 0.29	1.36 ± 0.30	14.3%	0.76
LDL (mmol/L)	2.55 ± 0.87	2.44 ± 0.88	14.3%	0.76
Total cholesterol (mmol/L)	4.32 ± 0.96	4.28 ± 1.22	14.3%	0.93
Triglycerides (mmol/L)	0.83 (0.67-1.39)	0.88 (0.71-1.18)	14.3%	0.90
FT4 (pmol/L)	18.61 ± 2.34	19.14 ± 4.10	14.3%	0.66
Cord blood measurements				
Apolipoprotein A1 (g/L)	0.93 ± 0.19	0.92 ± 0.22	4.0%	0.90
Apolipoprotein B (g/L)	0.33 ± 0.27	0.27 ± 0.11	4.0%	0.36
HDL (mmol/L)	0.89 ± 0.19	0.89 ± 0.27	4.0%	0.98
LDL (mmol/L)	1.17 ± 0.93	0.95 ± 0.33	12.0%	0.45
Total cholesterol (mmol/L)	2.24 ±1.09	1.99 ± 0.57	8.0%	0.40
Triglycerides (mmol/L)	0.42 (0.17-0.71)	0.30 (0.17-0.42)	8.0%	0.32
Glucose (mmol/L)	4.86 ± 1.87	4.82 ± 0.89	12.0%	0.94
C-peptide (nmol/L)	0.24 ± 0.12	0.25 ± 0.19	4.0%	0.90
FT4 (pmol/L)	16.53 ± 2.69	17.58 ± 1.83	16.0%	0.21

P-values <0.05 are considered statistically significant and marked in bold. Data represented with mean±SD and median (IQR), unless stated otherwise (frequency (%)). Abbreviations: BMI: body mass index, HDL: High Density Lipoprotein, HG: Hyperemesis Gravidarum, FT4: free thyroxine, LDL: Low Density Lipoprotein. PUQE-24: 24-hour Pregnancy Unique Quantification of Emesis and nausea score: a higher PUQE-24 indicates more severe symptoms. SGA: small for gestational age. Weight change is weight at baseline minus prepregnancy weight and can be < 0 if women lost weight and can be > 0 if women gained weight.

CHAPTER 10

General discussion

In this thesis, we have described that there are substantial knowledge gaps in the top 10 priority research questions in HG. We found that the HG recurrence rate for people that previously suffered from HG is high and that maternal mental health sequelae among these patients, both during and after pregnancy, are common. We also illustrated in two systematic reviews that HG can lead to both adverse perinatal and long-term effects in offspring, and that a more severe HG disease course possibly leads to an increased cardiovascular disease risk based on analysed cord blood samples of neonates born to mothers with HG in a cohort study. Patient and public involvement (PPI) was an important attribute to this thesis: patient representatives had an initiating, leading and interpreting role in prioritising the top 10 most urgent research questions in HG. Patient representatives also were an integral part of the design of the follow-up study evaluating recurrence rates and mental health symptoms. We experienced that PPI in the research described in this thesis contributed to higher quality evidence and to insights and outcomes of greater relevance to patients. The general discussion will take a closer look at PPI in general and in current HG research: How has PPI become incorporated in current healthcare and medical research? What are the benefits of PPI? To what extent are patients and the public currently involved in HG research, other medical fields and research agendas? Lastly, suggestions for PPI in future research will also be given.

History and current involvement of patients in healthcare and research

Patient centred care (PCC) has become a concept that is widely applied in current clinical practice. Although there is no worldwide uniform definition for PCC, all definitions in literature revolve around patients being involved in healthcare to increase individualisation of care and a more holistic approach of healthcare, that includes focussing on the whole patient instead of focussing solely on patients' disease or symtoms.[1] Hippocrates already said "*I would rather know what sort of person has a disease than what sort of disease a person has*". The term PCC was introduced in the literature in the early nineteen eighties and PCC only became common practice in healthcare in the past forty years.[2] Before the late 20th century, the doctor had a dominant position in the patient-doctor relation, which over time has evolved into a more equal relationship.[3] The introduction of Freud's psychology theories led to the emergence of psychology in medical healthcare. It is thought that this has led to a change in doctors' insight in patient perspectives: patients became more individual persons to doctors instead of being an object, which emphasized the importance of doctors listening to patients.[3] An increase in accessibility to healthcare, an upward trend in education level and the onset of a more equal society because of the end of the class society in the 20th century have likely contributed in the evolvement of PCC. Additionally, the introduction of informed consent and the right to self-determination has led to an increase in patients' autonomy and the patients' right to make

their own choices and decisions in medical care. Taken together with the fact that medical information was becoming available to a wider audience through for example the internet, an increase of patients being involved in current healthcare practices has come about.[3] In medical care, the benefits of PCC include higher patient satisfaction and increased perceived quality of care.[4] In the Netherlands, a cross-sectional survey study showed that PCC led to an increase in patient satisfaction and patients' physical and social well-being.[5]

Interestingly enough, when considering that PCC was already introduced in the eighties, its equivalent in research is of much more recent date. PPI in research was launched about 25 years ago: INVOLVE was the first program that was set-up to support active public involvement in research as well as in healthcare and public health in 1996 in the United Kingdom.[6] INVOLVE defined PPI as follows: research carried out *with* or *by* patients and their relatives as well as members of the general public rather than *to*, *about* or *for* them. PPI was first described in HIV and cancer research and later expanded into other fields of medicine.[7] The United Kingdom (UK) was the first country to actively incorporate PPI in research, and to this day has maintained a leading position in PPI. For example, a recent systematic review that evaluated available PPI frameworks concluded that a majority of the available frameworks originated from the UK (34 out of 65 frameworks), whereas only three frameworks originated from the Netherlands.[8] Additionally, PPI is one of the requirements to receive funding from the National Institute of Health Research, the leading funder in the UK.[9] When comparing this with other European countries, for example the Nordic countries, it is only mandatory to include PPI in research proposals to receive funding from governmental institutes in Norway, but not in Denmark, Sweden or Finland.[10] In for example Germany, two of the largest governmental funding institutes include a section regarding whether patients are, or will be involved in their research proposal application forms, but do not state that PPI is mandatory.[11] Throughout Europe, the European Patients' Academy on Therapeutic Innovation (EUPATI) was launched in 2017, which is an European education program for patients to increase their knowledge on healthcare and research to improve PPI.[12] Currently 21 European countries are involved and in 2019 also a Dutch version was launched.[12, 13]

Benefits of patient and public involvement in research

Several advantages of involving patients in research have been previously demonstrated. Identifying research topics that are important to patients and involving patients in developing study designs are thought to lead to an increase of patient recruitment and participation in studies, although the evidence to support this claim is scarce, and the size of possible benefits is not well known.[14] Survey questionnaires and instruments can be improved by involving

patients. This is for example shown in a previously published study that involved patients and family members in exploring older patients' needs in receiving pre-operative information and deciding on whether to undergo high-risk surgery.[15] Available evidence from qualitative studies and experiences from patients and their relatives were discussed during monthly meetings, which resulted in a questionnaire that was used at the outpatient clinic to improve shared decision making in whether older patients should undergo surgery. Moreover, PPI can lead to improvement of communication of study results. For example when patient organisations have been involved in research, they can contribute to the dissemination of study results, in which they are also likely to reach a wider audience of patients.[14] Trust, commitment and a good interaction between researchers and involved patients however, have been highlighted to be key elements to most optimally involve patients and public in research.[16-18] In a recently published systematic review that included 36 studies, most studies reported a positive attitude towards PPI among professionals, although there were also studies that concluded that the role of PPI remained marginal or tokenistic and that there was reluctance towards PPI implementation among professionals.[19]

Current patient and public involvement in HG research

In HG research, patients have been increasingly involved, or even taking the lead, in the last decade. Patient organisations from different countries, including the HER foundation, Pregnancy Sickness Support (PSS) and Zwangerschapsmisselijkheid en hyperemesis gravidarum (ZEHG), have been involved in HG research and contributed to emancipation of HG sufferers in both research and clinical healthcare. From 2015 on, the International Collaboration for Hyperemesis Gravidarum (ICHG) research, a working group that consists of patient organisations, researchers, and practising clinicians, has co-organised scientific conferences and has the ambition to collaboratively push the research agenda forward. Patients have been involved in the process of developing a core outcome set and definition for HG as well as in the identification of the top 10 most priority research questions in HG.[20-22] Even though this is an important step in PPI in HG research, there is still plenty of room for improvement: *Chapter 2* of this thesis includes a systematic review that collected evidence on the top 10 most priority research questions in HG and also reported on PPI of the included studies. In about 8% (n=31 studies) of the over 400 included studies patients were involved in conducting HG research. Although some studies included PPI, this percentage of 8% is still very low. When looking more specifically at the type of study design and PPI, we noticed that 35% (11/31) of the studies that reported PPI were survey studies. In studies with a different design, for example randomised controlled trials (RCTs), none of the completed trials described PPI. Only one included RCT protocol of an ongoing trial reported having included PPI in the protocol development.[23] When looking at registered

HG clinical trials in Europe, the only other registered, recently completed RCT in the United Kingdom (the EMPOWER study: EMesis in Pregnancy - Ondansetron With mEtoclopRamide, results not yet published),[24] also report PPI in the development and execution of the study. This could mean that there is an upward trend in PPI in high quality research, which is an important step in current research practices. Another explanation could be that previously PPI was simply not reported and that nowadays, to receive funding from governmental institutes, PPI is mandatory, just like reporting of PPI in study protocols and published studies, whereas earlier this was not the case. This might lead to PPI not always being reported in previously published literature, despite the fact that they have been involved.

Current PPI in HG guidelines and thus research implementation is sparse. HG guidelines have been published by the American College Of Obstetricians and Gynaecologists, the UK's Royal College of Obstetricians and Gynaecologists and the Nordic Federation Of Societies of Obstetrics and Gynaecology, in which only the guideline of the Royal College of Obstetricians and Gynaecologists reports involvement of a patient representative.[25-27] The first Dutch guideline for HG is currently being developed by the Dutch Association Of Obstetrics and Gynaecology (NVOG) and will be available mid-2023. The Dutch patient organisation ZEHG is involved in its development. The guideline, and the active patient involvement in its development, will address a long-held wish of the HG patient community in the Netherlands, which is to have a more consistent national approach with joined primary and secondary care.[28]

Patient and public involvement in setting up research agendas within obstetrics and gynaecology

Research agendas have been developed to prioritise research topics and to reduce research waste in the future.[29] Research agendas aim for a better match in priorities of patients and healthcare professionals and research questions that are being addressed by researchers.[30] Additionally, research agendas increase awareness among funders which topics are most important for patients and healthcare professionals and that need funding. In the obstetrics and gynaecology sector in the Netherlands, the NVOG started to develop a national research agenda in 2016 (published in 2017 and updated in 2020) in order to identify current knowledge gaps in the field.[31, 32] In these research agendas the top 10 most urgent research questions in obstetrics and gynaecology together as well as the top 10 most urgent research questions in obstetrics alone were identified. This framework has its limitations. For example, since HG is just a small part of the obstetrics and gynaecology field, none of the research questions included in the 2017 agenda, or in any of the subsequent versions, concerned HG. Additionally, PPI is

limited: current PPI consists of including the patient federation in the process of suggesting research questions and priority voting, and of involving only the largest patient organisations.

A commonly used methodology to identify top 10 priority research questions in a research topic is the James Lind Alliance (JLA) methodology, which originates from the UK.[33] This method was also used to identify the top 10 most urgent research questions in HG, as described in *chapter 2*. The JLA methodology consists of setting up a Priority Setting Partnership, including multiple stakeholders such as patients and their carers, healthcare professionals and researchers, that identify so far unanswered, priority research questions with use of surveys and online meetings.[33] The JLA method consists of five steps: Step 1) setting up a steering group; Step 2) gathering uncertainties by use of for example an online survey, although in some cases also interviews and/or discussion groups take place; Step 3) data processing and verifying uncertainties by reviewing and interpreting survey responses; Step 4) interim priority setting which consists of reducing the number of research questions collected; Step 5) final priority setting where all questions are ranked in priority order.

The JLA has previously identified top 10 research questions in other subspecialties of Gynaecology and Obstetrics, as for example endometriosis, lichen sclerosus, miscarriages and preterm birth.[34-37] When taking a closer look at for instance the top 10 priority research questions of endometriosis, it is interesting to see the differences in ranking between patients and their carers and researchers and healthcare professionals. Healthcare professionals and researches ranked research topics such as the "the cause, risk and pathology", "diagnosis and screening", "treatments" and "fertility" higher than patients and their carers.[38] On the other hand, patients and their carers designated "education and awareness", "emotional impact" and "comorbid conditions" as more important topics.[38] Similar differences between patients and professionals perspectives were found in the emergence of priority research questions in preterm birth.[37] Both groups listed "Which treatments are most effective to predict or prevent preterm birth?" as the most important research question. The second most important question to patients however ("The optimum milk feeding strategy and guidance for the best long-term outcomes of premature babies?"), eventually ended up as question number six in the final top 10, while the third most important research question to patients ("How do stress, trauma and physical workload contribute to the risk of preterm birth?") did not make the top 10, and ended up as question number 18.

Patient and public involvement in setting up research agendas in other fields of medicine

Previously we described the JLA method that includes professionals, researchers and patients to prioritise research questions with use of surveys and online meetings. Another similar method to prioritise research questions is the Dialogue Model, a method that has been developed and validated in the Netherlands.[39] The Dialogue Model comprises six phases and, similar to the JLA, includes multiple stakeholders.[39] The first phase consists of a exploration phase in which participants are gathered and first information is collected by a literature search. In the second consultation phase, multiple different stakeholders are consulted separately in order to set up a list of research questions. In the third prioritisation phase, research questions identified in phase two are prioritised, and in the fourth phase these research questions are combined into one research agenda including the top 10 research priorities after discussing them between the different stakeholders. Additionally, there are phase 5 and 6 that include programming and implementation of the research agenda to receive funding and make sure that the research agenda is used among researchers. The dialogue and JLA method have many similarities. However, in the JLA method information from different participants is gathered mostly through surveys, with the additional possibility to include interviews and discussions between patients and participants, while the main focus of the dialogue method lies on the dialogue meetings in phase 4, where patients get into conversations with professionals.

In 2006 the Dutch Burns Foundation developed a research agenda according to the Dialogue Model.[40] In the first phase patients ranked "itchiness" as the top priority research topic, while researchers and healthcare professionals did not consider it as an important research topic. However, eventually in the fourth phase, when conducting dialogue meetings between burn survivors, researchers and clinicians, "itchiness" was ranked as the second most important research topic among the theme "tissue regeneration".[40] Another example can be found within the Netherlands Asthma Foundation, that also created a revised research agenda in 2009 with use of the Dialogue Model. Involvement of patients in setting up this research agenda led to inclusion of co-morbidity, fatigue and psychological problems such as stress and depression among Asthma and COPD patients in the top 10 list, topics that were not previously recognised as important matters.[41]

While PPI was included in prioritising research topics in burn care as well as in the Asthma and COPD field, phase 5 and 6 of the Dialogue Model are eventually most important. As described earlier, phase 5 evaluates if research about one of the priority topics has been carried out or whether funding has been sought or granted, while phase 6 evaluates whether patients

were involved in research projects on a regular basis after the research agenda was set up.[39] A qualitative study looked into whether phase 5 and 6 were applied in eight different medical health categories in the Netherlands that previously developed a research agenda with use of the Dialogue Model.[18] The follow-up of nine different research agendas were critically assessed, among which the earlier discussed research agendas of the Dutch Burns Foundation and Asthma and COPD. The development of three out of nine research agendas was funded by the local government. These three research agendas did not lead to a formal funding program, neither were the large part of the included research topics funded in a later stage: in two out of three research agendas only two research topics eventually received funding, while none of the research topics of the third research agenda received funding. Only some patient involvement was described in implementation in these three studies. The other six out of the nine research agendas resulted in a formal funding program and active implementation of the research agendas, although patient involvement was described as limited. All six of these research agendas were funded by charity foundations of the disease in question, which could be a reflection of a higher level of commitment or higher resource levels among the patient representatives involved, compared to PPI in governmental funded research agendas. The research agenda of the Dutch Lung Foundation was the only research agenda that followed formal programming and implementation including PPI, with for example currently all grant applications being reviewed by a patient pool, which was shown to encourage implementation of the research agenda.[42]

Future implications

PPI played an important role in the design, the conduction and interpretation of data from research that was included in this thesis, which was invaluable for the reseach and its translation to clinical practice. PPI in research is in an upward trend and advantages of PPI in research have been described in the current chapter. However, we also described that there is only a relatively small amount of evidence available and that the results of the effects of including PPI in research are not consistent.[14] Some studies stated that there is a likely possibility that PPI is simply being reported since PPI is required for funders instead of them actively being involved and that PPI is only included in the final stages of research projects instead of from the beginning.[43] Previously, a systematic review also showed that there is insufficient evidence available on how to optimally involve patient representatives in research.[44] For those reasons, some professionals are still reluctant to include PPI in research.[19] An important role in future research therefore lays in further exploring the benefits of PPI in research and how PPI can most optimally be incorporated.

When looking specifically at HG research, patients and public have been involved in some important research developments over the past years: the development of a core outcome set, creating a uniform worldwide definition for HG and in prioritising the top 10 most urgent research questions.[20, 22, 45] An important next step in PPI in HG research is to also include PPI in the following programming and implementation phases, since previous research in other fields of medicine showed that PPI in these phases often still need improvement.[18] This includes further gathering evidence (or funding to support research) on the top 10 most priority research questions in HG, with the questions ranked highest in the top 10 being most urgent, and to incorporate PPI in early stages of research projects and in updating guidelines. This will lead to research that suits the needs of both healthcare professionals and patients and to an improvement of HG care.

REFERENCES

1. Robinson JH, Callister LC, Berry JA, Dearing KA. Patient-centered care and adherence: Definitions and applications to improve outcomes. Journal of the American Academy of Nurse Practitioners. 2008;20(12):600-7.
2. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med. 2012;27(10):1361-7.
3. Kaba R, Sooriakumaran P. The evolution of the doctor-patient relationship. International Journal of Surgery. 2007;5(1):57-65.
4. McMillan SS, Kendall E, Sav A, King MA, Whitty JA, Kelly F, et al. Patient-centered approaches to health care: a systematic review of randomized controlled trials. Med Care Res Rev. 2013;70(6):567-96.
5. Kuipers SJ, Cramm JM, Nieboer AP. The importance of patient-centered care and co-creation of care for satisfaction with care and physical and social well-being of patients with multi-morbidity in the primary care setting. BMC Health Services Research. 2019;19(1):13.
6. NIHR - INVOLVE. [Available from: https://www.invo.org.uk/about-involve/.
7. Stewart D, Wilson R, Selby P, Darbyshire J. Patient and public involvement. Annals of Oncology. 2011;22:vii54-vii6.
8. Greenhalgh T, Hinton L, Finlay T, Macfarlane A, Fahy N, Clyde B, et al. Frameworks for supporting patient and public involvement in research: Systematic review and co-design pilot. Health Expect. 2019;22(4):785-801.
9. Denegri S. Going the extra mile: improving the nation's health and wellbeing through public involvement in research. London: NIHR; 2015.
10. Sand A-S, Grimsgaard S, Pettersen I. Patient and public involvement in health research: A Nordic perspective. Scandinavian Journal of Public Health. 2020;48(1):119-21.
11. Schilling I, Bleidorn J, Ehrmann U, Müller-Fries E, Rathjen KI, Saedler K. Chancen und Herausforderungen der aktiven Beteiligung von Patient*innen an klinischer Forschung in Deutschland – Eine Betrachtung aus den Perspektiven eines Patient*innenvertreters, einer klinischen Forscherin und zweier Mitarbeiterinnen des Forschungsfördermanagements. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2021;163:66-75.
12. Spindler P, Lima BS. Editorial: The European Patients Academy on Therapeutic Innovation (EUPATI) Guidelines on Patient Involvement in Research and Development. Front Med (Lausanne). 2018;5:310-.
13. van Rensen A, Voogdt-Pruis HR, Vroonland E. The Launch of the European Patients' Academy on Therapeutic Innovation in the Netherlands: A Qualitative Multi-Stakeholder Analysis. Front Med (Lausanne). 2020;7:558-.
14. Lee DJ, Avulova S, Conwill R, Barocas DA. Patient engagement in the design and execution of urologic oncology research. Urol Oncol. 2017;35(9):552-8.
15. Steffens NM, Tucholka JL, Nabozny MJ, Schmick AE, Brasel KJ, Schwarze ML. Engaging Patients, Health Care Professionals, and Community Members to Improve Preoperative Decision Making for Older Adults Facing High-Risk Surgery. JAMA Surgery. 2016;151(10):938-45.
16. Barber R, Beresford P, Boote J, Cooper C, Faulkner A. Evaluating the impact of service user involvement on research: a prospective case study. International Journal of Consumer Studies. 2011;35(6):609-15.
17. Lindenmeyer A, Hearnshaw H, Sturt J, Ormerod R, Aitchison G. Assessment of the benefits of user involvement in health research from the Warwick Diabetes Care Research User Group: a qualitative case study. Health Expectations. 2007;10(3):268-77.

18. Abma TA, Pittens CACM, Visse M, Elberse JE, Broerse JEW. Patient involvement in research programming and implementation. Health Expectations. 2015;18(6):2449-64.

19. Biddle MSY, Gibson A, Evans D. Attitudes and approaches to patient and public involvement across Europe: A systematic review. Health & Social Care in the Community. 2021;29(1):18-27.

20. Jansen L, Koot M, van't Hooft J, Dean C, Duffy J, Ganzevoort W, et al. A core outcome set for hyperemesis gravidarum research: an international consensus study. BJOG: An International Journal of Obstetrics & Gynaecology. 2020;127(8):983-92.

21. Jansen LAW, Koot MH, Van 't Hooft J, Dean CR, Bossuyt PMM, Ganzevoort W, et al. The Windsor Definition for Hyperemesis Gravidarum: a multistakeholder International Consensus Definition. American Journal of Obstetrics & Gynecology. 2021. Submitted.

22. Dean CR, Bierma H, Clarke R, Cleary B, Ellis P, Gadsby R, et al. A patient-clinician James Lind Alliance partnership to identify research priorities for hyperemesis gravidarum. BMJ open. 2021;11(1):e041254-e.

23. Ostenfeld A, Petersen TS, Futtrup TB, Andersen JT, Jensen AK, Westergaard HB, et al. Validating the effect of Ondansetron and Mirtazapine In Treating hyperemesis gravidarum (VOMIT): protocol for a randomised placebo-controlled trial. BMJ Open. 2020;10(3):e034712.

24. National Institute for Health Research. EMPOWER: EMesis in Pregnancy - Ondansetron With mEtoclopRamide [Available from: https://fundingawards.nihr.ac.uk/award/16/15/03.

25. ACOG Practice Bulletin No. 189: Nausea And Vomiting Of Pregnancy. Obstet Gynecol. 2018;131(1):e15-e30.

26. Royal College of Obstetricians and Gynaecologists. The Management of Nausea and Vomiting of Pregnancy and Hyperemesis Gravidarum: RCOG Green-Top Guideline No. 69 2016 [updated June 2016]. Available from: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg69/.

27. Vikanes A, Trovik J, Tellum T, Lomsdal S, Stensløkken A, Nesheim B. Emesis & hyperemesis gravidarum 2014 [Available from: http://www.nfog.org/files/guidelines/7%20NGF%20Obst%20hyperemisis%20Vikanes.pdf.

28. van Vliet R, Bink M, Polman J, Suntharan A, Grooten I, Zwolsman SE, et al. Patient Preferences and Experiences in Hyperemesis Gravidarum Treatment: A Qualitative Study. J Pregnancy. 2018;2018:5378502.

29. Dechartres A, Ravaud P. Better prioritization to increase research value and decrease waste. BMC Medicine. 2015;13(1):244.

30. Crowe S, Fenton M, Hall M, Cowan K, Chalmers I. Patients', clinicians' and the research communities' priorities for treatment research: there is an important mismatch. Research Involvement and Engagement. 2015;1(1):2.

31. Nederlandse Vereniging Obstetrie en Gynaecologie. NVOG-Kennisagenda 2017-2020 [Available from: https://www.nvog.nl/wp-content/uploads/2017/12/NVOG-kennisagenda-2017-2020.pdf.

32. Nederlandse Vereniging Obstetrie en Gynaecologie. NVOG-Kennisagenda 2020-2023 [Available from: https://www.nvog.nl/wp-content/uploads/2019/07/NVOG-kennisagenda-2020-2023_def.pdf.

33. James Lind Alliance. The James Lind Alliance Guidebook 2021. Available from: https://www.jla.nihr.ac.uk/jla-guidebook/downloads/JLA-Guidebook-Version-10-March-2021.pdf.

34. Horne AW, Saunders PTK, Abokhrais IM, Hogg L. Top ten endometriosis research priorities in the UK and Ireland. Lancet. 2017;389(10085):2191-2.

35. Simpson RC, Cooper SM, Kirtschig G, Larsen S, Lawton S, McPhee M, et al. Future research priorities for lichen sclerosus – results of a James Lind Alliance Priority Setting Partnership. British Journal of Dermatology. 2019;180(5):1236-7.

36. Prior M, Bagness C, Brewin J, Coomarasamy A, Easthope L, Hepworth-Jones B, et al. Priorities for research in miscarriage: a priority setting partnership between people affected by miscarriage and professionals following the James Lind Alliance methodology. BMJ Open. 2017;7(8):e016571.

37. Oliver S, Uhm S, Duley L, Crowe S, David AL, James CP, et al. Top research priorities for preterm birth: results of a prioritisation partnership between people affected by preterm birth and healthcare professionals. BMC Pregnancy and Childbirth. 2019;19(1):528.

38. Brady PC, Horne AW, Saunders PTK, Thomas AM, Missmer SA, Farland LV. Research priorities for endometriosis differ among patients, clinicians, and researchers. American journal of obstetrics and gynecology. 2020;222(6):630-2.

39. Abma TA, Broerse JE. Patient participation as dialogue: setting research agendas. Health Expect. 2010;13(2):160-73.

40. Broerse JE, Zweekhorst MB, van Rensen AJ, de Haan MJ. Involving burn survivors in agenda setting on burn research: an added value? Burns. 2010;36(2):217-31.

41. Elberse J, Laan D, de Cock Buning T, Teunissen T, Broerse J, de Boer W. Patient involvement in agenda setting for respiratory research in the Netherlands. European Respiratory Journal. 2012;40(2):508-10.

42. Pittens CACM, Elberse JE, Visse M, Abma TA, Broerse JEW. Research agendas involving patients: Factors that facilitate or impede translation of patients' perspectives in programming and implementation. Science and Public Policy. 2014;41(6):809-20.

43. Martin GP. Representativeness, legitimacy and power in public involvement in health-service management. Soc Sci Med. 2008;67(11):1757-65.

44. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res. 2014;14:89.

45. Jansen LAW, Koot MH, Van't Hooft J, Dean CR, Bossuyt PMM, Ganzevoort W, et al. The windsor definition for hyperemesis gravidarum: A multistakeholder international consensus definition. Eur J Obstet Gynecol Reprod Biol. 2021;266:15-22.

10

SUMMARY

In *chapter 1*, the general introduction of this thesis, we give a summary about HG in general. We describe that it is likely that HG has a multifactorial aetiology and we give an overview of the current available anti-emetic medications and lack of any curative treatment. We also illustrate that HG has an enormous impact on people suffering from HG. HG can lead to anxiety, depression and post-traumatic stress disorder symptoms that can maintain *after* suffering from HG. Lastly, we describe that there is a need for a more up-to-date review about the current available evidence on perinatal and long-term effects of HG in offspring.

Part I, including *chapter 2*, presents the top 10 priority research questions in HG that were recently identified. Current knowledge gaps and available evidence for these top 10 research questions were systematically assessed by including 406 articles that were displayed in an evidence map. The evidence map highlighted that the least evidence was available on finding a cure for HG (priority question 1), preventative treatment (priority question 4) and on how to achieve nutritional requirements in pregnancy (priority question 10). There was also lack of evidence for most other top 10 questions.

Part II addresses the effects of HG on maternal future health. In *chapter 3* we investigated whether thyroid function could be used as a marker for monitoring en predicting the disease course of HG and concluded that this was not the case, due to inconsistent findings. *Chapter 4* presents a prospective follow-up cohort study that showed that the chance of people suffering from HG again is high (89%). We also demonstrated that many people postponed their pregnancy due to suffering from HG in the past (40%) and that people even considered terminating their otherwise wanted pregnancy because of recurrent HG (23%). *Chapter 5* demonstrates maternal mental health outcomes in a follow-up study of a prospective cohort of people that previously suffered from HG. Anxiety and depression symptoms were frequently reported at six weeks postpartum (24 and 20%), as well as at follow-up about 4.5 years later (40 and 27%), which is much higher than the incidence among pregnant or postpartum people without HG. One out of five participants also had symptoms of a post-traumatic stress disorder at follow-up. In *chapter 6* we present the results of a systematic review on HG-related vitamin K deficiency and corresponding maternal and neonatal complications. One cohort study and 14 case-reports were included that reported on maternal complications, which consisted of coagulopathy-related haemorrhage, and neonatal complications, which consisted of intracranial haemorrhage and embryopathy. We were unable to assess the incidence of vitamin K deficiency among HG patients and their offspring.

Part III describes the effects of HG on offspring's future health. In *chapter 7* we present the results of a large systematic review and meta-analysis on perinatal outcomes of offspring of people suffering from HG. Sixty-one studies were identified reporting on 20,532,671 infants in total. Meta-analyses showed that maternal HG was more often associated with preterm birth <34weeks, birth weight <1500 grams, neonatal resuscitation, neonatal intensive care unit admission and placental abruption, and less often associated with birthweight >4000 grams and stillbirths. No associations were found between HG and Apgar scores <7 at 1 and 5 minutes, fetal loss, perinatal deaths and neonatal deaths in meta-analyses. In *chapter 8* results of a systematic review and meta-analysis on long-term outcomes of offspring born to people with HG were displayed. Nineteen studies were included reporting on 1,814,785 offspring. Meta-analyses showed that HG exposure in utero was associated with anxiety and sleep disorders in offspring and possibly testicular cancer in male offspring. Narrative synthesis showed that all six included studies reporting on attention deficit (hyperactivity) disorders and autism spectrum disorders found that these disorders were more common among HG-exposed offspring compared to non-exposed offspring. The majority of included studies reporting on cognitive and motor problems found an increase among HG-exposed children compared to non-exposed offspring. Findings of anthropometry and cardiometabolic disease markers were inconsistent among the five included studies. Lastly, in *chapter 9* we analysed cardiometabolic disease markers in cord blood of offspring of people with HG and addressed whether there were differences among the disease spectrum of HG and between offspring of people who were allocated to receive tube feeding and to standard care. Multivariable analysis showed that lower maternal weight gain was associated with higher levels of apolipoprotein-B in cord blood. No associations were found between other HG severity measures or allocation to enteral tube feeding and cord blood cardiometabolic disease markers.

Chapter 10 includes the general discussion and future implications of this thesis. Here, we describe patient and public involvement (PPI) in research included in this thesis and PPI in general. We found that the possible benefits of PPI and how to optimally incorporate PPI in research have not been well investigated. We describe how PPI is currently applied in other fields of medicine and give examples of different research agenda's that were developed with PPI and which methods were used to develop them. Future implications for further implementation and improvement of PPI in HG research consisted of future research focussing on the top 10 priority research questions and the involvement of patients representatives in updating/developing HG guidelines and early stages of research.

APPENDICES

A

Nederlandse samenvatting

In *hoofdstuk 1*, de introductie van dit proefschrift, geven we een samenvatting over HG in het algemeen. We beschrijven dat HG waarschijnlijk een multifactoriële etiologie heeft en we geven daarnaast een overzicht van de huidige beschikbare anti-emetica en het gebrek aan een curatieve behandeling. Tevens laten we zien dat HG een enorme impact heeft op mensen die lijden aan HG. HG kan leiden tot angst, depressie en post-traumatische stressstoornis klachten, die kunnen aanhouden tot na de bevalling. Als laatste beschrijven we dat er behoefte is aan een meer up-to-date overzicht van de huidige beschikbare literatuur over perinatale en lange termijn uitkomsten van nakomelingen van moeders met HG.

In **deel I** en *hoofdstuk 2* presenteren we de recent opgestelde 10 meest belangrijke onderzoeksvragen in HG. Met behulp van een systematische review evalueerden we zowel de beschikbare literatuur als de huidige kennislacunes voor deze top 10 onderzoeksvragen. Vierhonderd zes artikelen werden geïncludeerd en weergegeven in een zogeheten "evidence map": een digitale, visuele weergave van de huidige beschikbare literatuur. Deze *evidence map* liet zien dat de minste literatuur beschikbaar is over het vinden van een curatieve behandeling voor HG (vraag #1), preventieve behandelingsmogelijkheden (vraag #4) en over de benodigde voedingsstoffen tijdens de zwangerschap (vraag #10). Er was ook onvoldoende literatuur aanwezig om de meeste andere top 10 onderzoeksvragen te beantwoorden.

In **deel II** beschrijven we de effecten van HG op de gezondheid van moeders, zowel tijdens als na afloop van de zwangerschap. In *hoofdstuk 3* onderzochten we of de schildklierfunctie gebruikt kon worden als voorspeller voor het ziekteverloop van HG en concludeerden we dat dit niet het geval was vanwege tegenstrijdige resultaten. In *hoofdstuk 4* laat een Nederlandse prospectieve cohortstudie van patiënten die werden opgenomen voor HG zien dat er een hoog herhalingsrisico (89%) is op het opnieuw ontstaan van HG in een volgende zwangerschap. Daarnaast zagen we dat HG gepaard ging met het uitstellen van een kinderwens in verband met de kans op een recidief (40%), en dat zwangeren zelfs overwogen om hun gewenste zwangerschap te beëindigen vanwege het opnieuw doormaken van HG (23%). In *hoofdstuk 5* laten we de gevolgen van HG zien voor maternale psychiatrische klachten in een prospectief cohort van patiënten die waren opgenomen voor HG in een eerdere zwangerschap. Angst- en depressieklachten waren veelvoorkomend zes weken postpartum (24 en 20%), en ten tijde van de vervolgstudie ongeveer 4,5 jaar later (40 en 27%), wat beduidend hoger is dan de incidentie onder zwangeren/kraamvrouwen zonder HG. Een op de vijf deelnemers had ongeveer 4,5 jaar na de zwangerschap met HG ook symptomen van een posttraumatische

stressstoornis. In *hoofdstuk 6* presenteren we de resultaten van een systematische review over HG gerelateerde vitamine K deficiëntie en de bijbehorende maternale en neonatale complicaties. Eén cohortstudie en 14 case reports werden geïncludeerd die rapporteerden over maternale complicaties, bestaande uit stolling gerelateerde bloedingen, en neonatale complicaties, bestaande uit intracraniële bloedingen en embryopathie. De incidentie van vitamine K deficiëntie onder HG patiënten en hun nakomelingen kon niet worden bepaald.

Deel III geeft de effecten van HG op de gezondheid van nakomelingen weer. In *hoofdstuk 7* presenteren we de resultaten van een grote systematische review en meta-analyse over perinatale uitkomsten van nakomelingen van mensen met HG. Er werden eenenzestig studies met in totaal 20,532,671 neonaten geïncludeerd. Meta-analyses toonde aan dat HG geassocieerd was met vroeggeboorte <34 weken, geboortegewicht <1500 gram, reanimatie van de pasgeborene, neonatale intensive care unit opnames en placenta loslating. HG was tevens geassocieerd met een afname in een geboortegewicht >4000 gram en een afname van intra uteriene vruchtdood. Er werd geen associatie gevonden tussen HG en Apgar scores<7 bij 1 en 5 minuten, miskramen, perinatale sterfte en neonatale sterfte in meta-analyses. In *hoofdstuk 8* worden de resultaten van een systematische review en meta-analyse over de lange termijn gevolgen van nakomelingen van mensen met HG weergegeven. Negentien studies werden geïncludeerd met in totaal 1.814.785 nakomelingen. Meta-analyses lieten zien dat HG geassocieerd was met angststoornissen, slaapproblemen en mogelijk teelbalkanker bij nakomelingen. Alle zes studies over aandachtsproblemen (met of zonder hyperactiviteit) en problemen in het autismespectrum rapporteerden een toename hiervan bij kinderen van moeders met HG in vergelijking met kinderen van niet-aangedane moeders. In de meerderheid van de geïncludeerde studies werden bij kinderen van moeders met HG meer cognitieve en motorische problemen waargenomen dan bij kinderen van niet-aangedane moeders. Studies naar cardiale en metabole ziekten toonden wisselende resultaten. Ten slotte hebben we in *hoofdstuk 9* cardiale en metabole ziektemarkers in navelstrengbloed van nakomelingen van vrouwen met HG bepaald. We onderzochten of er verschillen waren tussen patiënten met een milde en ernstige HG, en tussen nakomelingen van HG patiënten die gerandomiseerd werden voor sondevoeding of voor standaardzorg. Multivariabele regressie toonde aan dat een lagere maternale gewichtstoename in de zwangerschap geassocieerd was met hogere apolipoproteine B waardes in navelstrengbloed van nakomelingen. Er werd geen associatie gevonden tussen cardiale en metabole ziektemarkers en andere maten van de ernst van HG. Tevens werden geen verschillen gevonden in cardiale en metabole ziektemarkers in navelstrengbloed van nakomelingen van HG patiënten die gerandomiseerd werden voor sondevoeding of voor standaardzorg.

Hoofdstuk 10 bevat de discussie en aanbevelingen voor de toekomst van dit proefschrift. We beschrijven hier de patiënt participatie (PPI) in het onderzoek dat is opgenomen in dit proefschrift en in onderzoek in het algemeen. Wij zagen dat de mogelijke voordelen van PPI en het zo optimaal mogelijk implementeren van PPI niet goed onderzocht is. We beschrijven tevens hoe PPI momenteel wordt toegepast in andere sectoren in de gezondheidszorg en geven voorbeelden van verschillende onderzoek agenda's die met behulp van PPI zijn opgesteld en welke methoden hiervoor zijn gebruikt. Aanbevelingen voor de toekomst voor verdere implementatie en het verbeteren van patiënt betrokkenheid bij HG onderzoek bestaan uit het focussen op de top 10 meest belangrijke onderzoeksvragen in HG en het betrekken van patiëntvertegenwoordigers in het updaten/opstellen van HG richtlijnen en in vroege stadia van onderzoeksprojecten.

List of co-authors

J.M.J. Bais	Department of Obstetrics and Gynaecology, Noordwest Ziekenhuisgroep, Alkmaar, the Netherlands
H.A. Bremer	Department of Obstetrics and Gynaecology, Reinier de Graaf Hospital, Delft, the Netherlands
C.R. Dean	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands Pregnancy Sickness Support, Bodmin, UK
R. van Eekelen	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands Department of Epidemiology & Data Science, Amsterdam UMC, Vrije Universiteit Medical Centre, Amsterdam, the Netherlands
M.J.J. Finken	Department of Paediatric Endocrinology, Emma Children's Hospital, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
I.J. Grooten	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
D.P. van der Ham	Department of Obstetrics and Gynaecology, Martini Hospital, Groningen, the Netherlands
W.M. Heidema	Department of Obstetrics and Gynaecology, Radboud University Medical Center, Nijmegen, the Netherlands
A. Huisjes	Department of Obstetrics and Gynaecology, Gelre Hospital, Apeldoorn, the Netherlands
L.A.W. Jansen	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands Department of Obstetrics and Gynaecology, Amphia Hospital, Breda, the Netherlands Department of Epidemiology & Data Science, Amsterdam UMC, Vrije Universiteit Medical Centre, Amsterdam, the Netherlands
G. Kleiverda	Department of Obstetrics and Gynaecology, Flevo Hospital, Almere, the Netherlands
M.H. Koot	Department of Obstetrics and Gynaecology, Gelre Hospital, Apeldoorn, the Netherlands Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
H. Kruizenga	Department of Nutrition and Dietetics, Internal Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
S.M. Kuppens	Department of Obstetrics and Gynaecology, Catharina Hospital, Eindhoven, the Netherlands

J.O.E.H. van Laar	Department of Obstetrics and Gynaecology, Máxima Medical Center, Veldhoven, the Netherlands
J. Langeveld	Department of Obstetrics and Gynaecology, Zuyderland Hospital, Heerlen, the Netherlands
J.L. Limpens	Medical Library, Research Support, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
F. van der Made	Department of Obstetrics and Gynaecology, Franciscus Gasthuis, Rotterdam, the Netherlands
S. Middeldorp	Department of Vascular Medicine, Amsterdam UMC, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam, the Netherlands
L. van der Minnen	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
B.W. Mol	Department of Obstetrics and Gynaecology, Monash University, Clayton, Victoria, Australia Aberdeen Centre for Women's Health Research, Institute of Applied Health Sciences, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK
C. Naaktgeboren	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
M.E. O'Hara	Pregnancy Sickness Support, Bodmin, UK
M. Olff	Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam the Netherlands ARQ National Psychotrauma Centre, Diemen, the Netherlands
R.C. Painter	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands Department of Obstetrics and Gynecology, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands
D. Papatsonis	Department of Obstetrics and Gynaecology, Amphia Hospital, Breda, the Netherlands
M.J. Pelinck	Department of Obstetrics and Gynaecology, Scheper Hospital, Emmen, the Netherlands
P.J. Pernet	Department of Obstetrics and Gynaecology, Spaarne Gasthuis, Haarlem, the Netherlands
J.A.M. van der Post	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
L. van Rheenen-Flach	Department of Obstetrics and Gynaecology, OLVG, Amsterdam, the Netherlands
R.J. Rijnders	Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, 's-Hertogenbosch, the Netherlands

C. Ris-Stalpers	Laboratory of Reproductive Biology, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
T.J. Roseboom	Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands Department of Epidemiology & Data Science, Amsterdam UMC, Vrije Universiteit Medical Centre, Amsterdam, the Netherlands
H.C.J. Scheepers	Department of Obstetrics and Gynaecology, Maastricht University Medical Center, Maastricht, the Netherlands
S.E. Siegelaar	Department of Internal Medicine, Endocrinology and Metabolism, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands
R. Spijker	Department of Epidemiology & Data Science, Amsterdam UMC, Vrije Universiteit Medical Centre, Amsterdam, the Netherlands
T. Vogelvang	Department of Obstetrics and Gynaecology, Diakonessenhuis, Utrecht, the Netherlands
H.M.G. Wiegers	Department of Vascular Medicine, Amsterdam UMC, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam, the Netherlands

List of publications

Nijsten K, van der Minnen LM, Dean C, Bais JMJ, Ris-Stalpers C, van Eekelen R, et al. Depression, anxiety and post-traumatic stress disorder symptoms after hyperemesis gravidarum: a prospective cohort study. J Matern Fetal Neonatal Med. 2022:1-9.
doi: 10.1080/14767058.2022.2089550

Nijsten K, Jansen LAW, Limpens J, Finken MJJ, Koot MH, Grooten IJ, et al. Long-term health outcomes of children born to mothers with hyperemesis gravidarum: a systematic review and meta-analysis. Am J Obstet Gynecol. 2022.
doi: 10.1016/j.ajog.2022.03.052

Nijsten K, Koot MH, Bais JM, Ris-Stalpers C, van Eekelen R, Bremer HA, et al. Hyperemesis gravidarum severity, enteral tube feeding and cardiometabolic markers in offspring cord blood. Br J Nutr. 2022:1-30.
doi: 10.1017/S0007114522000587

Nijsten K, Dean C, van der Minnen LM, Bais JMJ, Ris-Stalpers C, van Eekelen R, et al. Recurrence, postponing pregnancy, and termination rates after hyperemesis gravidarum: Follow up of the MOTHER study. Acta Obstet Gynecol Scand. 2021;100(9):1636-43.
doi: 10.1111/aogs.14197

Nijsten K, Koot MH, van der Post JAM, Bais JMJ, Ris-Stalpers C, Naaktgeboren C, et al. Thyroid-stimulating hormone and free thyroxine fail to predict the severity and clinical course of hyperemesis gravidarum: A prospective cohort study. Acta Obstet Gynecol Scand. 2021;100(8):1419-29.
doi: 10.1111/aogs.14131

Nijsten K,* van der Minnen L,* Wiegers HMG, Koot MH, Middeldorp S, Roseboom TJ, et al. Hyperemesis gravidarum and vitamin K deficiency: a systematic review. Br J Nutr. 2021:1-13.
doi: 10.1017/S0007114521002865
**shared first authorship*

Dean C, **Nijsten K**, Spijker R, O'Hara ME, Roseboom TJ, Painter RC. A systematic evidence map of evidence addressing the top ten priority research questions for hyperemesis gravidarum. Accepted for publication in The British Medical Journal Open.

Jansen LAW, **Nijsten K**, Limpens J, van Eekelen R, Koot MH, Grooten IJ, et al. Perinatal outcomes of infants born to mothers with hyperemesis gravidarum: a systematic review and meta-analysis. Submitted.

PhD portfolio

Name:	Kelly Nijsten
PhD period:	October 2019 – June 2022
Promotor:	Prof. Dr. Tessa J. Roseboom
Co-promotor:	Prof. Dr. Rebecca C. Painter

1. PhD training	Year	ECTS
Courses		
GCP light plus course: training modules 'basics, start en uitvoer'	2019	0.2
Basic Course in Legislation and Organization for Clinical Researchers (BROK)	2020	1.5
Clinical epidemiology: systematic reviews	2020	0.7
Seminars		
Weekly department seminars, Obstetrics & Gynaecology department AMC	2018-2021	4.0
Monthly journal club, Obstetrics & Gynaecology department AMC	2018-2021	2.0
Two-weekly department seminars, Life course Epidemiology Club	2021-2022	2.0
Monthly journal club, Obstetrics & Gynaecology department Flevoziekenhuis	2021-2022	2.0
Oral presentations		
The effects of in utero undernutrition due to hyperemesis gravidarum on offspring's cardiometabolic disease risk *AR&D symposium, Amsterdam, the Netherlands*	2019	0.5
Hyperemesis gravidarum: maternal and neonatal future health *KEBB seminar, AMC, Amsterdam, the Netherlands*	2020	0.5
Poster presentations		
TSH fails to predict the severity and clinical course of Hyperemesis gravidarum: a prospective cohort study *ICHG 2019, Amsterdam, the Netherlands*	2019	0.5
Recurrence rate of hyperemesis gravidarum and maternal future health: follow up of the MOTHER RCT and associated cohort. A study protocol *ICHG 2019, Amsterdam, the Netherlands*	2019	0.5
The effects of in utero undernutrition due to hyperemesis gravidarum on offspring's cardiometabolic disease risk and the possible beneficial effect of maternal early nutritional intervention *ICHG 2019, Amsterdam, the Netherlands*	2019	0.5

(Inter)national conferences		
AR&D symposium, Amsterdam, the Netherlands	2019	0.25
3rd International Colloquium on Hyperemesis Gravidarum 2019 (2 day event)	2019	0.5
AR&D symposium, Amsterdam, the Netherlands	2020	0.25
Gynaecongres 2021	2021	0.25

Other scientific activities		
NTOG publication: *'Implementatie van een nieuw mictie protocol postpartum: over op intermitterend katheteriseren bij symptomatische urineretentie'*	2019	3

2. Teaching	**Year**	**ECTS**
Loïs van der Minnen, bachelor thesis Medicine *(daily supervisor, 8 weeks)*	2019	1.0
Brita Roell, bachelor thesis Medicine *(daily supervisor, 8 weeks)*	2019	1.0
Martina Astorga Alsina, research internship Amsterdam university college *(daily supervisor, 8 weeks)*	2020	1.0
Nianne Schuitema, bachelor thesis of Nutrition & Dietetics, Amsterdam University of Applied Sciences *(daily supervisor, 20 weeks)*	2020	1.5
Judith van den Brink, bachelor thesis of Nutrition & Dietetics, Amsterdam University of Applied Sciences *(daily supervisor, 20 weeks)*	2020	1.5
Roos Anna de Jong, bachelor thesis of Nutrition & Dietetics, Amsterdam University of Applied Sciences *(daily supervisor, 20 weeks)*	2020	1.5

3. Parameters of Esteem	**Year**	**ECTS**
Grants		
AR&D grant 'Start small, think big'	2019	2.0

Dankwoord

Ik wil iedereen bedanken die heeft bijgedragen aan dit proefschrift. Zonder jullie was dit nooit gelukt! Graag wil ik enkele personen in het bijzonder bedanken.

Beste prof. T.J. Roseboom, beste Tessa, dank voor al je hulp en steun de afgelopen jaren tijdens mijn promotie. Dank voor al jouw vertrouwen in mijn soms wat ambitieuze plannen om te verrichten naast mijn ANIOS baan en dat jij er voor zorgde dat ik ook een periode fulltime aan de slag kon. Jij zorgde er voor dat ik altijd met een fijn gevoel vol vertrouwen de deur uit liep na onze gesprekken.

Beste dr. R.C. Painter, beste Rebecca. Vijf jaar geleden, tijdens mijn wetenschappelijke stage, leerden wij elkaar voor het eerst kennen en was ik vooral onder de indruk van je enorme kennis. Inmiddels heb ik je leren kennen als iemand met een enorme toewijding voor je patiënten en PhD studenten en als iemand die zich met passie met de wetenschap bezighoudt. Dank voor al je hulp de afgelopen jaren, van het begeleiden van mijn wetenschappelijk stage, tot aan het schrijven van mijn AR&D beurs aanvraag en mijn thesis, maar vooral ook voor je luisterende oor en het sparren over van alles en nog wat.

Lieve Marjette, tijdens mijn wetenschappelijke stage in hyperemesis is mijn enthousiasme voor zowel de wetenschap als voor dit onderwerp ontstaan, mede dankzij jou! Jij hebt mij kennis laten maken met de verschillende aspecten van onderzoek doen. Daarnaast heb jij mij geholpen met het schrijven van de AR&D beurs aanvraag, wat uiteindelijk tot deze thesis heeft geleid. Dankjewel hiervoor!

Dank aan iedereen die betrokken was bij het uitvoeren van de follow-up studies: beste Mirjam Dijkstra en Femke Schrauwen, trialcoördinatoren van het lab: dankjewel voor jullie hulp rondom het analyseren van de ingevroren bloedafnames. Natuurlijk ook veel dank aan alle deelnemende vrouwen van de MOTHER studie.

Geachte leden van de promotiecommissie, prof. dr. M. Goddijn, prof. dr. K.F.M. Joosten, dr. M.R. Soeters, prof. dr. M.J.M. Serlie, prof. dr. J. de Jonge en prof. dr. K.W.M. Bloemenkamp, veel dank dat jullie bereid waren om mijn proefschrift kritisch te beoordelen en plaats te nemen in mijn promotiecommissie.

Lieve (inmiddels oude) AMC collega's: AIOS, ANIOS en alle arts-onderzoekers van H4. Dankjewel voor al jullie steun tijdens mij PhD traject. Dank voor alle lekkere koffietjes, etentjes, vrijdagmiddagborrels en wandelingen in corona tijd. Vooral aan mijn mede parttime arts-onderzoekers en ANIOS: samen hebben we heel wat uurtjes achter de computers in de arts-assistentenkamer doorgebracht. Gelachen, gezeurd, maar vooral ook elkaar er samen doorheen gesleept.

In het bijzonder lieve Claartje: mijn buddy in mijn PhD en ANIOS tijd en samen begonnen aan de opleiding. Dit betekent nog meer sparren over werk, maar vooral ook nog meer gezellige koffietjes, lunches, borrels en wandelingetjes.

Lieve Joost, mijn voormalige mentor! Dank voor al je oppeppende woorden, je hulp bij het voorbereiden van mijn sollicitatie, je vertrouwen in mij, maar vooral ook je gezelligheid! Ik hoop dat onze paden elkaar nog meerdere keren kruisen in de toekomst.

Lieve vrienden en vriendinnen, de maidon, de huismiepen, vriendinnen van thuis, Ems, Johanna en nog veel meer. Dank jullie wel voor al jullie steun de afgelopen jaren. Met jullie kon ik heerlijk ontspannen na werk tijdens etentjes, feestjes, filmbezoekjes, sportlesjes en noem het maar op. Zonder jullie support was dit nooit gelukt!

Dear Caitlin, I really enjoyed working with you this final year. Your ambition and dedication in HG research is admiring and I hope we continue working on HG projects together in the future! And finally have some drinks in real-life again after many, many zoom meetings.

Lieve Larissa, mijn andere hyperemesis-buddy. Wij werken niet alleen heel fijn samen, maar hebben het ook nog eens mega gezellig samen, een top combinatie! Dat dit boekje er eindelijk ligt, betekent zeker niet dat er geen etentjes meer gaan volgen.

Lieve, lieve Anna. Hoe zou mijn tijd in het Flevoziekenhuis zijn geweest zonder jou? Een stuk minder gezellig denk ik;) Dankjewel dat ik altijd bij je mag sparren, zeuren, maar vooral dat we erg kunnen lachen samen. Ik zou niet weten wat ik zonder jou zou moeten! Op nog veel mooie etentjes in het bruisende Almere, waarvan ik zeker weet dat we die ook (ergens anders) gaan voortzetten na onze tijd in het Flevoziekenhuis.

Mijn paranimfen. Lieve Jazzy, al sinds kind af aan zijn wij twee handen op één buik en daarom voor mij ook meer dan logisch om jou als mijn paranimf te vragen. Ik ben heel blij dat je hier

vandaag aan mijn zijde staat. Lieve Liselotte, lieve lies. Al sinds dat wij elkaar ontmoeten op onze eerste dag van onze studententijd in Amsterdam bij de introductieweek van biomedische wetenschappen zijn wij onafscheidelijk en hebben we zoveel raakvlakken in zowel ons privéleven als op werk gebied. Dankjewel voor al je hulp en steun de afgelopen 13 jaar.

Lieve pap en mam, dank jullie wel voor al jullie hulp door de jaren heen. Jullie steun heb ik altijd gehad, in wat ik ook ging doen. Heerlijk om jullie nuchterheid om mij heen te hebben en om te weten dat jullie altijd trots zijn. Zonder jullie was dit niet gelukt!

Lieve Joyce. Als oudere zus, die ook geneeskunde heeft gestudeerd en inmiddels kinderarts is, heb jij mij in zoveel mijlpalen van mijn carrière geholpen en gesteund: van de decentrale selectie van geneeskunde, tot het solliciteren voor mijn eerste ANIOS baan en het solliciteren voor de opleiding gynaecologie. Zonder jou was dit allemaal, inclusief deze promotie, niet gelukt. Dankjewel hiervoor!

Lieve Quirijn, jij kreeg met al mijn kanten te maken afgelopen jaren: van blij, naar moe, naar soms een tikkeltje overwerkt. Hoe dan ook, je was er altijd voor mij en hielp me altijd mijn gedachten op een rijtje te zetten en te ontspannen wanneer ik thuis kwam. Ik zou niet weten hoe ik dit zonder jou had moeten doen. Op een hele mooie toekomst samen!

About the author

Kelly Nijsten was born on the 5th of January 1991 in Venlo, the Netherlands. She graduated from the Valuascollege in Venlo and moved to Amsterdam in 2009 to start studying Biomedical Science at the University of Amsterdam. A year later, she was accepted to medical school, which she completed 7 years later. While waiting to start on her clinical internships, she travelled to Southeast Asia. In her final year of her medical school, she travelled to Cape Town, South-Africa for an international internship at the Gynaecology and Obstetrics department of the Groote Schuur Hospital and the Mowbray Maternity Hospital. For her Master thesis at the University of Amsterdam, she became involved in research on hyperemesis gravidarum while writing her master thesis under supervision of Marjette Koot and Rebecca Painter, entitled: '*Thyroid-stimulating hormone and free thyroxine fail to predict the severity and clinical course of hyperemesis gravidarum: a prospective cohort study*'. She continued doing research in hyperemesis gravidarum alongside to her work as a resident in the Obstetrics and Gynaecology department of the St. Antonius Hospital. Subsequently, she started working as a resident in the Obstetrics and Gynaecology department of the Amsterdam UMC, location AMC. She acquired an AR&D research grant for analysing cardiometabolic markers in cord blood of offspring of people with hyperemesis gravidarum to investigate their possible long term health risk, which formed the basis for this thesis. Initially, she carried out research for her PhD part time alongside to her clinical work as a resident, but from April 2020 onwards she was able to continue her PhD full-time. In April 2021, she started as a junior registrar in Obstetrics and Gynaecology at the Flevoziekenhuis in Almere.