

**The case – control method in obstetrics and gynecology:  
etiology of hyperemesis gravidarum**



**Irene Sandven**

**MPH**

Unit of Epidemiology and Biostatistics

Ullevaal Centre for Clinical Research

Ullevål Department of Research Administration

Oslo University Hospital

Norway

2010

© Irene Sandven, 2011

*Series of dissertations submitted to the  
Faculty of Medicine, University of Oslo  
No. 1095*

ISBN 978-82-8264-007-7

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.  
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.  
The thesis is produced by Unipub merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

# Contents

<b>Acknowledgments</b>	<b>5</b>
<b>1. List of papers</b>	<b>6</b>
<b>2. Aim of the thesis</b>	<b>7</b>
<b>3. Introduction</b>	<b>9</b>
3.1. The case-control method and its use in obstetrics and gynecology	9
3.1.1. History of the case-control method	9
3.1.2. General structure of the case-control method	9
3.1.2.1. Conceptual foundations	10
3.1.2.2. Odds ratio approximation in a case-control study	11
3.1.3. The classical case-control design	11
3.1.3.1. Selection bias	12
3.1.3.2. Information bias	13
3.1.3.3. Confounding bias	14
3.1.4. Alternative case-based designs	15
3.1.4.1. Role of the case-based design in obstetrics and gynecology	15
3.1.4.2. The case-cohort design	16
3.1.4.3. The nested case-control design	18
3.1.4.4. The case-crossover design	19
3.1.5. How many controls?	20
3.2. Epidemiology and etiology of hyperemesis gravidarum	21
3.2.1. Definition	21
3.2.2. Diagnosis	21
3.2.3. Prevalence	22
3.2.4. Etiology	22
3.2.4.1. Endocrine factors	22
3.2.4.2. Genetic factors	23
3.2.4.3. Helicobacter pylori infection	23
3.2.4.4. Psychological factors	23
3.2.4.5. Other risk factors	24
3.2.5. Severe consequences of hyperemesis gravidarum	24
<b>4 Material and statistical method</b>	<b>25</b>
4.1 Material	25

4.2	Statistical methods	25
<b>5</b>	<b>Summary of papers</b>	<b>29</b>
<b>6</b>	<b>General discussion</b>	<b>31</b>
6.1	Etiology of hyperemesis gravidarum	31
6.1.1	Internal validity	31
6.1.1.1	Selection bias	31
6.1.1.2	Information bias	32
6.1.1.2.1	Non-response, an issue often forgotten	32
6.1.1.3	Confounding bias	33
6.1.1.3.1	Interaction, a word with two meanings	33
6.1.1.4	Heterogeneity between studies in the meta-analysis	38
6.1.2	Consideration of positive features of causation on the association between H.pylori and hyperemesis gravidarum	39
6.1.2.1	Is there a correct time relationship?	39
6.1.2.2	Is the association strong?	40
6.1.2.3	Is there a dose-response relationship?	40
6.1.2.4	Is there a biological plausible explanation?	41
6.1.3	Web of causation	41
6.1.3.1	Endocrine factors	42
6.1.3.2	Psychological factors	43
6.1.3.3	Genetic factors	43
6.1.4	Need to include gene-environment interaction in the etiology of hyperemesis gravidarum	44
6.2	Critical appraisal of case-control studies on the etiology or risk factors of hyperemesis gravidarum	45
6.2.1	Problems with the tool for critical appraisal	46
<b>7</b>	<b>Conclusions</b>	<b>47</b>
7.1	The case-control and case-based designs in obstetrics and gynecology	47
7.2	Etiology of hyperemesis gravidarum	48
7.2.1	Future perspectives	48
<b>8</b>	<b>References</b>	<b>49</b>
<b>9</b>	<b>Papers 1-3</b>	

## **Acknowledgement**

The present work has been funded by Helse Sør-Øst and Oslo University Hospital, and carried out at the Unit of Epidemiology and Biostatistics, Ullevaal Centre for Clinical Research, Oslo University Hospital in the years 2005-2010.

I am deeply grateful to Michael Abdelnoor, my main supervisor, for his guidance and support. Generously, he has initiated me to the concepts and methods of epidemiology. Without his visions this thesis would never been conceived and completed. Hopefully, we will continue working together on future projects.

I have been fortunate to collaborate with other experienced clinical researchers, Britt-Ingjerd Nesheim (co-supervisor) and Kjetil K Melby, who have stimulated my work. Britt-Ingjerd Nesheim has given me input as far her knowledge in obstetrics and gynecology. Kjetil Melby has introduced me to the problem of H.pylori infection and hyperemesis gravidarum. The assistance of the medical library staff at Oslo University Hospital has been invaluable providing literature, and my special thank to Tiril Bonnevie for assisting with comprehensive database searches. Interesting discussions with my husband, Håkon Wergeland, have been appreciated and I am grateful to him for encouraging me in my work. Finally, I thank all my colleagues for creating a nice and stimulating work environment.

Oslo, June 2010

## **1 List of papers**

### **Paper I**

Sandven I, Abdelnoor M, Wethe M, Nesheim B-I, Vikanes Å, Gjønnnes H, Melby KK. Helicobacter pylori infection and hyperemesis gravidarum. An institution-based case-control study. Eur J Epidemiol 2008; 23: 491-8

### **Paper II**

Sandven I, Abdelnoor M, Nesheim B-I, Melby KK. Helicobacter pylori infection and hyperemesis gravidarum: a systematic review and meta-analysis of case-control studies. Acta Obstet Gynecol Scand 2009; 88: 1190-1200

### **Paper III**

Sandven I, Abdelnoor M. Critical appraisal of case-control studies of risk factors or etiology of hyperemesis gravidarum. Arch gynecol obstet 2010; 282: 1-10

## **2 Aim of the thesis**

In this thesis our aim was to:

- a) Highlight the use of the case-control method and case-based designs in the field of obstetrics and gynecology.
- b) Investigate the association between *Helicobacter pylori* (*H.pylori*) and Hyperemesis gravidarum (HG). The case-control method is a suitable design studying the etiology of HG, as its prevalence is low (0.3 – 2 %).
- c) Perform a survey of the literature as far the published case-control studies of the association of *H.pylori* as the exposure and HG as the outcome. Try to do a pooled estimate of the effect of this association, and check the cause of heterogeneity in this meta-analysis.
- d) Perform critical appraisal of all epidemiological case-control studies of the risk factors of HG.





## 3 Introduction

### 3.1 The case-control method and its use in obstetrics and gynecology

#### 3.1.1 History of the case-control method

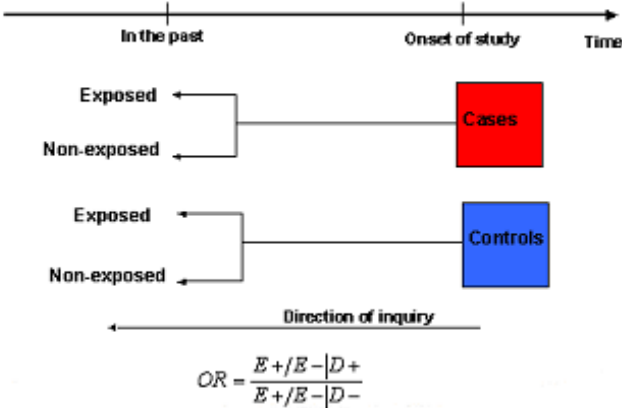
The case-control study design has a long history beginning in the 19<sup>th</sup> century, with interesting examples of studies containing elements of the design (Paneth et al)<sup>1</sup>. A modern form of the case-control study appeared early in the 20<sup>th</sup> century when Lane-Clayton investigated the risk of female breast cancer, and the method developed further in the years following World War II (Paneth et al)<sup>2</sup>. We can still learn from many fine case-control studies done during the 1950<sup>ies</sup>, for example the classical study by Doll and Hill (Doll et al)<sup>3</sup> with good suggestions on how to minimize selection and information bias. During the same decade a number of landmarks in the development of the case-control method occurred, including the calculation of the odds ratio by Cornfield (Cornfield et al)<sup>4</sup> and the Mantel-Haenzel method (Mantel et al)<sup>5</sup> controlling for confounding. In a way, the story of cigarette smoking and lung cancer is the story of the development of the case-control method.

The popularity of the case-control method grew. Over a period of 20 years, from 1956 to 1977, the number of case-control studies published in selected medical journals increased four to seven fold (Cole et al)<sup>6</sup>. In gynecology, significant findings from case-control studies have been many. Classical examples of causal exposure-disease relationships are Diethylstilbestrol and vaginal adenocarcinoma (Herbst et al)<sup>7</sup>, tampon use and toxic-shock syndrome (Reingold et al)<sup>8</sup>, and estrogens treatment and endometrial cancer (Ziel et al)<sup>9</sup> (Smith et al)<sup>10</sup>. In the years since 1950, the case-control method has been further refined and has become an important tool in the armamentarium of today's investigators of health problems.

#### 3.1.2 General structure of the case-control method

Essentially the case-control method is a comparison of exposure history for a group of persons with a certain disease or condition (cases) with another group of persons who do not have that disease or condition (controls). The following résumé is inspired by Kleinbaum et al (Kleinbaum et al)<sup>11</sup>. As illustrated in Figure 1, a case-control study works backwards from disease status to prior exposure status, while the timing can be either prospective or

retrospective. Ideally, incident cases should be selected from a clearly defined source-population during a specified period of time, and the ideal control group should be representative of that source-population. The case-control method is able to calculate a ratio of odds of exposure in cases and controls as a measure of association. Major advantages include suitability to study rare diseases or those with long latency, and allowing several exposures to be evaluated. It is relatively quick to conduct and reasonably inexpensive compared to cohort studies. Further, it is often the only study design ethically applicable to a clinical problem because the disease has already occurred. Disadvantages include susceptibility to selection and information bias, not allowing estimation of risk, not considering more than one disease and not feasible for rare exposures.



**Figure 1.**  
The case-control study design

*3.1.2.1 Conceptual foundations*

Basically, two kinds of observational analytic study designs in epidemiology are the cohort study and the case-control study. There are clear links between these two study designs Rothman argues (Rothman)<sup>12</sup>. All case-control studies are generated from an existing or hypothetical mother cohort. Sampling is done on the bases of outcome from such an ongoing real or imagined cohort study. Therefore it can be helpful in the planning stage of a case-control study to first plan the ideal cohort study that would have been conducted investigating the same hypothesis if unlimited recourses were available.

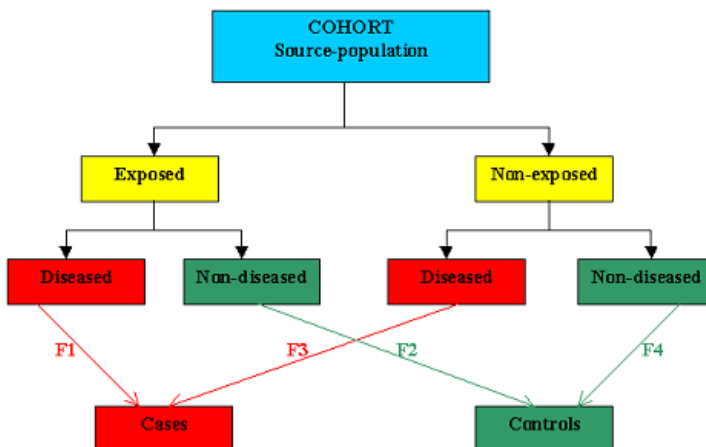
### 3.1.2.2 Odds ratio approximation in a case-control study

In a cohort study two fundamental measures of effect can be estimated, the odds ratio (OR) and the risk ratio (RR). While in a case-control study we can estimate the OR only. When the disease of interest is rare or the prevalence of the disease is low over the study period (a cumulative incidence of less than 10%), the OR will be approximately equal to the RR. When we use only incident cases and the controls are a random sample of non-diseased coming from the same source-population, it can be showed mathematically that the OR estimated from case-control design is the same as the one from the true or fictitious mother cohort (Kleinbaum et al)<sup>11</sup>.

### 3.1.3 The classical case-control design

In this option we select the cases and controls from the true or fictitious mother cohort (source-population) during the follow up period (Figure 2), for example a random sample of the survivors at the end of follow up. During a given period of time, enrolled cases are a fraction  $F_1$  of the exposed who developed the disease plus a fraction  $F_3$  of the unexposed who developed the disease. Similarly, enrolled controls are a fraction  $F_2$  of the exposed who did not developed the disease plus a fraction  $F_4$  of the unexposed who did not developed the disease. The different fraction of all the cases and controls from the cohort study that is included in the case-control study will represent the major selection bias (Kleinbaum et al)<sup>11</sup>.

There will be a selection bias if the ratio of  $\frac{F_1 * F_4}{F_2 * F_3}$  is different to one.



**Figure 2.**  
Sampling from a true or fictitious mother cohort

### *3.1.3.1 Selection bias*

Selection bias concerns a distortion of study results that occurs because of the way subjects are selected into the study. Of major concern is that selection of cases, controls, or both might be influenced by the exposure of interest. To minimize bias it is preferable to use incident cases rather than prevalent cases and population-based controls rather than hospital-based controls (Kleinbaum et al)<sup>11</sup>. The following example is commonly used to illustrate the problem with selection bias and how difficult it sometimes can be to determine whether study results are biased. In the 1970's there was a debate among researchers whether selection bias could explain a strong harmful effect of estrogens replacement in menopausal women on the risk of endometrial cancer, reported in early case-control studies (Ziel et al)<sup>9</sup> (Smith et al)<sup>10</sup>. The controls typically used were women with gynaecologic cancers other than endometrial cancer (ovarian, cervical, vulva). Critics claimed that because estrogens often cause vaginal bleeding irrespective of cancer, estrogens users with endometrial cancer would receive closer medical surveillance compared to nonusers with endometrial cancer. Consequently, one would expect cases being more likely estrogens users than controls, and the OR biased away from the null. An alternative sampling method was later suggested by Horwitz and Feinstein (Horwitz et al)<sup>13</sup>. In their study they tried to compensate for a possible unequal chance for detection of endometrial cancer among users and nonusers of estrogens, by selecting cases and controls from a source population of women who underwent curettage or hysterectomy because of uterine bleeding. The investigators showed a large decrease in the magnitude of the association between estrogens and endometrial cancer. However, the debate of selection bias continued (Hutchison et al)<sup>14</sup>, arguing that estrogens are potentially harmful developing endometrial cancer.

Unfortunately, we are unable to quantify selection bias unless we have reliable estimates of the underlying selection ratio parameters ( $F_1, F_2, F_3, F_4$ ), such as in studies nested within a cohort of known size. The best we usually can do at the design stage of our study is to make decisions minimizing selection bias, and at the publication stage address potential bias qualitatively in the discussion section of our article.

### 3.1.3.2 Information bias

A difficulty in case-control studies involves information on exposure status. The consequences of having inaccurate information about exposure variables is misclassification of exposure that leads to bias in the resulting OR (Kleinbaum et al)<sup>11</sup>. The underlying parameters considered when assessing information bias are sensitivity and specificity. Non-differential misclassification occurs whenever the sensitivities and specificities for misclassifying exposure do not vary with disease status, and will always dilute the estimated OR. With differential misclassification, the sensitivities and specificities for misclassifying exposure differ by disease status, and the estimated OR will be biased either toward or away from the null.

Misclassification of exposure happens if the data collecting process is not independent from the case-control status. The most important biases that can lead to misclassification are recall bias, non-response bias and interviewer bias. **Recall bias** can arise when information on past exposure relies on memory. Furthermore, cases may remember exposures to putative risk factors differently than controls. For example, Lindfors-Harris et al (Lindfors-Harris et al)<sup>15</sup> suggested that the tendency toward increased risk of breast cancer with induced abortion could be explained by recall bias. In this study, information on exposure (previous abortion) was gathered by interviews and by objectively documented information from a nationwide register. When information from interview was compared to information from vital statistics, a systematic underreporting of abortions was observed among controls compared to cases. The differential recall between cases and controls indicated a 50 percent increase in estimated risk of breast cancer associated with induced abortion. **Non-response bias** occurs if cases and controls have differential response rates to exposure information. In recent years, non-response has been recognized as an increasing problem in case-control studies (Feychting et al)<sup>16</sup>. In a comparison of the characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women, Madigan et al (Madigan et al)<sup>17</sup> reported a number of differences between the two study groups, some of which was related to the study hypothesis. **Interviewer bias** may occur when the interviewer favours a certain response differentially and unconsciously over another between cases and controls, leading to misclassification. Such bias can also be generated when interviewers are aware of case-control status and may try to obtain information about exposure with more intensity from cases than from controls.

Ideally we should use tested and validated instruments with known sensitivity and specificity when measuring exposure, which enables us to correct for misclassification (Kleinbaum et al)<sup>11</sup>. Furthermore, data collectors should be blinded to the research hypothesis and to whether the patient is a case or control. When exposure is measured in this optimal way, misclassification will be non-differential and the diluted effect on OR depend on the sensitivity and specificity of the instrument used. In other words, the actual association will turn out to be bigger. Rothman have been arguing that epidemiological studies with weak effect should be published because any real effect may be stronger than the reported one (Taubes et al)<sup>18</sup>.

### *3.1.3.3 Confounding bias*

Confounding is an important problem in observational clinical research. Potential confounders of an exposure (E) – disease (D) association are risk factor of D, associated with E, but not in the causal pathway between E and D. If we fail to consider other variables such as age and gender when assessing an E – D relationship, the estimated OR will be biased. In contrast to selection bias and information bias, the effect of confounding can easily be controlled, provided that confounding is anticipated and the requisite information gathered.

Confounding can be dealt with in the design stage by restriction or matching, but generally it is preferable to handle it in the analysis stage with analytical techniques such as the Mantel-Haenzel (M-H) method based on stratification and mathematical modelling (Kleinbaum et al)<sup>11</sup>. With restriction, the control for confounding is achieved by selecting into the study only individuals with certain homogenous levels of the potential confounder (Kleinbaum et al)<sup>11</sup>. For example, if cigarette smoking is suspected to be a confounder, only non-smokers are included. Restriction might increase the internal validity of a study, but at the cost of poorer external validity. With matching, controls are selected to each case in such a way that the potential confounders are identically distributed among cases and controls (Kleinbaum et al)<sup>11</sup>. For example, cases and controls can be matched by smoking status when smoking is a potential confounder. However, advantages of matching to control for confounding are limited and well described by Kleinbaum et al (Kleinbaum et al)<sup>11</sup>. Essentially, matching does not control for confounding in the case-control design and since it easily can be controlled at the analysis stage, whether to match or not is an issue of efficiency. Two problems with matching in case-control studies have also been highlighted and discussed by Sørensen et al.

(Sørensen et al)<sup>19</sup>, confounding introduced by the matching variable itself and loss of study efficiency due to overmatching.

#### 3.1.4 Alternative case-based designs

In later years we have seen the emergence of strategies for selecting and analyzing only a sample of subjects within a cohort. The two most common strategies are known as the nested case-control and case-cohort designs. Usually, all cases are selected but only a fraction of the cohort is selected as controls. The advantage of having all cases and controls derive from the same well-defined source population is maintained. These strategies are useful for large cohorts with few cases, where analyzing the entire cohort adds little efficiency but increases the cost. Another strategy developed more recently is the case-crossover design, which is a variant of the matched case-control study that is intended to be less prone to bias than the standard case-control design because of the way controls are selected.

##### *3.1.4.1 Role of the case-based design in obstetrics and gynaecology*

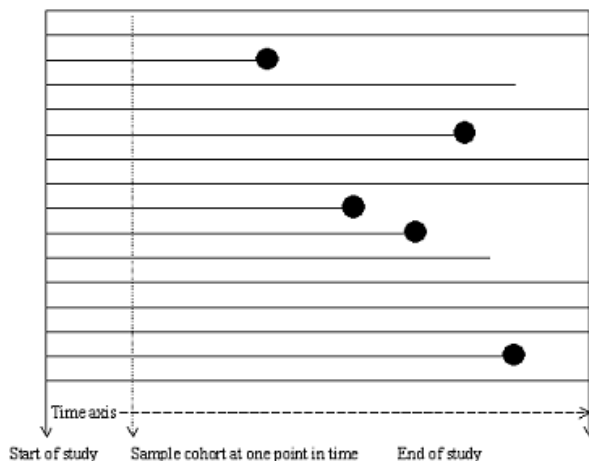
Alternative case-based designs seem to be used infrequently in obstetrics and gynecology, with most investigators choosing nested case-control designs when compared to the classical case-control design (Table 1). A PubMed search January 20<sup>th</sup> 2010, shows 397 occurrences of the keywords “nested case-control” versus 41 occurrences of “case-cohort” and 13 occurrences of “case-crossover”. These figures may have some limitations since mislabelling of case-control studies is a recognized problem in obstetrics and gynaecology journals (Grimes)<sup>20</sup>. However, our search shows a trend towards infrequent use of alternative case-based designs in obstetrics and gynecology. Its application may have been discouraged in some investigators because of perceived analytic complexity and lack of software. Others may simply be unaware of these design alternatives and its potential advantages.

**Table 1.** Popularity of the case-based design in obstetrics and gynaecology per January 20<sup>th</sup> 2010

<i>The classical case-control design (reference-group) versus alternative case-based designs</i>				
	<b>Case-Control</b>	<b>Nested case-control</b>	<b>Case-Cohort</b>	<b>Case-crossover</b>
<b>PubMed count</b>	<b>11834</b>	<b>397</b>	<b>41</b>	<b>13</b>
<i>Ratio</i>	<i>1</i>	<i>0.03</i>	<i>0.003</i>	<i>0.001</i>
<b>Google Scholar count</b>	<b>42600</b>	<b>2490</b>	<b>363</b>	<b>106</b>
<i>Ratio</i>	<i>1</i>	<i>0.06</i>	<i>0.008</i>	<i>0.002</i>

### 3.1.4.2 The case-cohort design

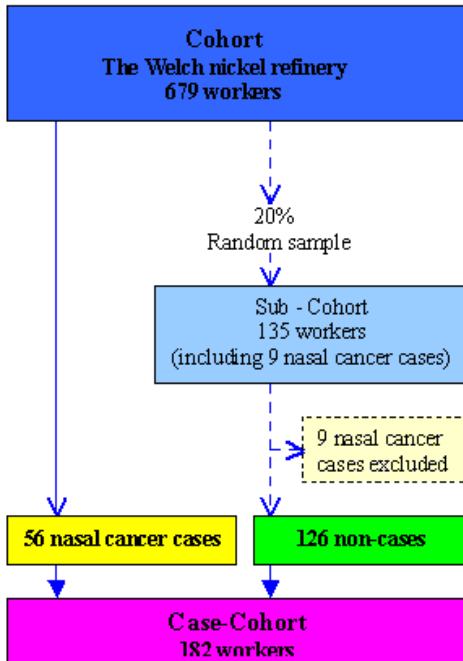
Performing a case-cohort study we begin by defining the cohort time axes. Following the total cohort (Figure 3), all individuals who develop the disease or outcome of interest (all cases) are included in the analysis. Controls are selected by taking a simple random sample from the entire cohort at one point in time (for instance at the start point of study). Depending on the sampling time point, control subjects in this sub-cohort possibly can become a case later during the follow up time. In Figure 4, the sampling method is illustrated considering the Welch nickel refinery workers and subsequent development of nasal cancer (Barlow et al)<sup>21</sup>.



**Figure 3.**

Case-cohort design. A horizontal line represents each individual at an entry time and ending at an exit time corresponding to censoring or failure. Failure times are indicated by black dots.





**Figure 4.**

Example of the Case-Cohort sampling method (Barlow et al)<sup>21</sup>. On a time axis of 29 years, there are 56 nasal cancer cases among 679 employees at the Welch nickel refinery. A 20% random sample of the total cohort, or “sub-cohort”, is designated as the source of comparisons for the observed cases. All cases are included whether they occur in the random sample or not, but non-cases are included only if in the sub-cohort.

Analyses of case-cohort data use a standard Cox regression approach with some modifications. According to the Barlow method (Barlow et al)<sup>21</sup>, the case weight is always one and the sub-cohort controls are weighted by the inverse of the sampling fraction. In the Welch nickel refinery example (Figure 4), the sampling fraction is 20%, and each sub-cohort control has a weight equal five. The analysis is done using a robust estimate of the standard error, which provides some protection against model misspecification. Usually those robust standard errors are larger than the incorrect naive standard error estimated by the usual model. In this example the variance of the case-cohort (robust SE = 0.259) is bigger than in the full cohort (robust SE = 0.170), which means that this case-cohort analysis is 43 % as efficient relative to the full cohort analysis:

$$Efficiency = \left( \frac{SE_{cohort}}{SE_{case-cohort}} \right)^2 = \left( \frac{0.170}{0.259} \right)^2 = 0.43$$

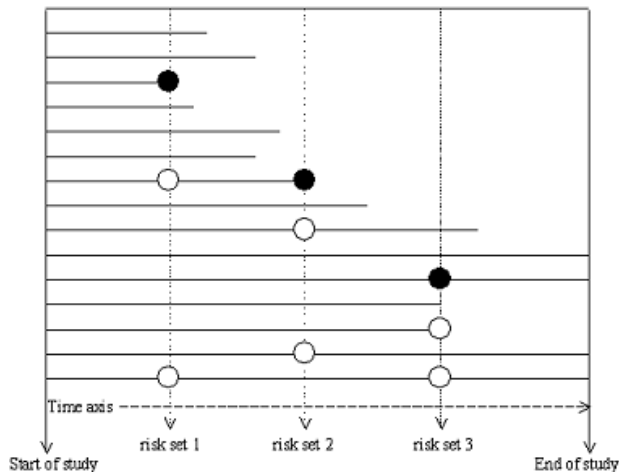
In general, the case-cohort design provides insurance that the controls derive from the source population from which the cases developed. Also, since cases are determined from follow up, several diseases can be studied. Furthermore, the case-cohort design can be an economic

alternative to the standard cohort design, especially when the incidence of the disease or event of interest is low. Studying rare diseases or uncommon events, as mortality, the cohort design requires collection of covariate information on all individuals in the cohort even if only a small fraction of these actually get diseased or die. A case-cohort sampling technique will reduce the costs, because it requires covariate information only from a random sample of non-diseased individuals or those who stay alive (Borgan et al)<sup>22</sup>. Good examples of the case-cohort design used in obstetrics and gynecology are Bell et al (Bell et al)<sup>23</sup> investigating the potential association between agricultural pesticide applications near maternal residence and fetal death, and Wernli et al (Wernli et al)<sup>24</sup> investigating the associations between occupational exposures in silk industry and endometrial cancer.

#### *3.1.4.3 The nested case-control design*

This approach selects cases and controls longitudinally throughout the cohort time axis, described as risk set sampling or density sampling (Kleinbaum et al)<sup>11</sup>. A person becomes a case at the time when a particular event or outcome of interest occurs. The process of control selection begins with the definition of a risk set for each case, from which the controls will be selected. Each risk set consists of all non-cases present in the cohort at the time each case becomes a case. Once the risk set is defined, controls for each case are randomly selected from the risk set corresponding to the case (Figure 5). Important to notice is that a future case may be selected as a control for a prior case, and that a given subject may be selected as a control for two different cases. Exposure and covariate information for each control selected from each case's risk set must reflect values at the time of selection (Lubin)<sup>25</sup>. This is particularly important for exposure and covariates that vary with time (White et al)<sup>26</sup>. Time matching is an essential feature of the nested case-control design (Wacholder)<sup>27</sup> (Ernststein)<sup>28</sup>, and matched designs require matched analysis (Kleinbaum et al)<sup>11</sup>. The nested case-control design can be used if the time at which subjects become cases is known, and controls are matched to the cases at the time of case diagnosis.

As with a case-cohort study, a nested case-control design provides insurance that the controls derive from the source population from which the cases developed. Also the nested case-control study is more cost and time-efficient than a cohort study, since a much smaller number of non-cases are observed.



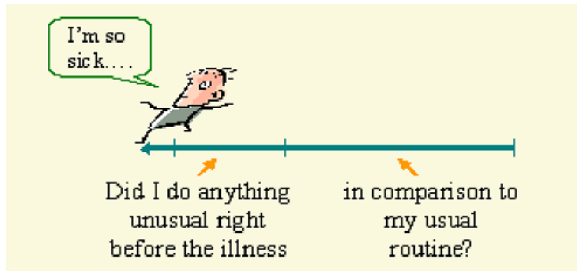
**Figure 5:** The nested case-control design. There are two controls (white circle) for every case (black circle). Follow up for all members of the cohort (horizontal black lines) begins at zero time at the time axis. Controls are randomly selected from the risk set (vertical dashed lines) for each case.

In obstetrics and gynecology the nested case-control design was used investigating whether smoking is an independent risk factor for invasive cervical cancer (Kapeu et al.)<sup>29</sup>. The study was nested in population-based serum banks and cancer registers of Nordic countries. For each case of invasive cervical cancer, five female controls that were alive and free of cancer at the time of the case's diagnosis were randomly selected. Analysis was done using conditional logistic regression for matched case-control sets. In the literature it was hard to find more good examples of the nested case-control design applied in obstetrics and gynecology research. Unfortunately, it seems to be some misunderstanding about the nested case-control design. When reviewing a sample of the 397 references labelled as "nested case-control" studies in PubMed, they rather appeared as classical case-control studies nested within a cohort without matching on the risk-set time variable using unconditional analysis.

#### 3.1.4.4 The case-crossover design

This method involves cases only and each individual serves as his/her own control. The same persons are studied at two points in time, once as cases and a second time as controls, and exposures antecedent to the two states are compared. The case-crossover design is a powerful method used to assess the effect of transient exposures and behaviours on the risk of acute events with a short latency (Maclure)<sup>30</sup>. Since exposure history is compared within the same individual at two different time windows, all time-independent potential confounders are

automatically controlled for in the design. As each case period is matched to a control period, analysis are done using matched methods (Mittelman et al)<sup>31</sup>.



**Figure 6.** The case-crossover design is a scientific way to ask and answer the question clinicians often ask patients: “Were you doing anything unusual just before the event”?

In obstetrics and gynecology the case-crossover design was used examining the relation between maternal smoking during pregnancy and birth of an infant with oral cleft (Meyer et al)<sup>32</sup>. In this study, data from the Swedish Medical Birth Registry was used to identify women who gave birth to a child with oral cleft. The control period was the births preceding the case birth. The OR was estimated using conditional logistic regression, indicating an association between smoking and oral cleft. Another study investigated whether condom use is associated with bacterial vaginosis (Hutchinson et al)<sup>33</sup>. Women at high risk for sexually transmitted diseases were screened for bacterial vaginosis at baseline and thereafter with 6 months intervals. Condom use (exposure history) was registered between each of the 6-month follow-up visits. Case periods were those in which bacterial vaginosis was diagnosed at the end of the time period, and control periods those in which women were diagnosed with normal vaginal microflora at the end of the time period. Conditional logistic regression analysis was used, and consistent condom use associated with a decreased risk for bacterial vaginosis.

### 3.1.5 How many controls?

For matched classical case-control studies, the widely accepted answer to this question is that there is little gain in precision by using more than four controls per case (Kleinbaum et al)<sup>11</sup>.

The Pitman efficiency formula compares the precision of pair matching relative to R to 1

matching:  $\frac{Var(adjustedOR_{1to1})}{Var(adjustedOR_{Rto1})} = \frac{2R}{(R+1)}$  which has been calculated for several values of R

(the table below), showing that efficiency increase 33.3% as R goes from 1 to 2, but only

4.2% as R goes from 4 to 5. The percent increase in efficiency is clearly smaller when R gets beyond 4. Moreover, the maximum possible efficiency is 2 and at R = 4 the efficiency is 1.6.

R	1	2	3	4	5	6	$\infty$
$\frac{2R}{(R+1)}$	1.000	1.333	1.500	1.600	1.667	1.714	2
% increase	-	33.3 %	12.5 %	6.7 %	4.2 %	2.8 %	-

For nested case-control studies, Pang (Pang)<sup>34</sup> showed that selecting a sufficient number of controls per case is important minimizing power loss, particularly for small numbers of risk sets. For example, when the relative risk was 1.5 and the number of risk sets was 20, relative powers for 1, 2, 4, 8, 16 and 32 controls per case were 19%, 39%, 60%, 76%, 87% and 94% respectively.

## 3.2 Epidemiology and etiology of hyperemesis gravidarum

### 3.2.1 Definition

Hyperemesis gravidarum (HG) is inconsistently defined. The Fairwater criteria define HG as vomiting more than three times a day, weight loss, ketonuria, electrolyte imbalance and volume depletion, with typical onset at 4-8 weeks of pregnancy continuing throughout week 14-16 (Fairwater)<sup>35</sup>. The International Statistical Classification of Disease and Related Health Problems ICD-9 Code 643 defines HG as persistent and excessive vomiting starting before the end of the 22<sup>nd</sup> week of gestation, and subdivides the condition as mild or unspecified (ICD-9 code 643.0) and severe with metabolic disturbance (ICD-9 code 643.1).

### 3.2.2 Diagnosis

Clinical symptoms of HG usually occur 4-7 weeks after the last menstrual period (Goodwin)<sup>36</sup>. Gradually, weight loss and symptoms of dehydration occur. The condition is assessed by physical examination, anamnesis, urine testing for ketones and serum testing for electrolytes, liver enzymes, amylase and thyroid function. Laboratory findings assist in assessing the severity of the condition and ruling out other causes of persistent vomiting.

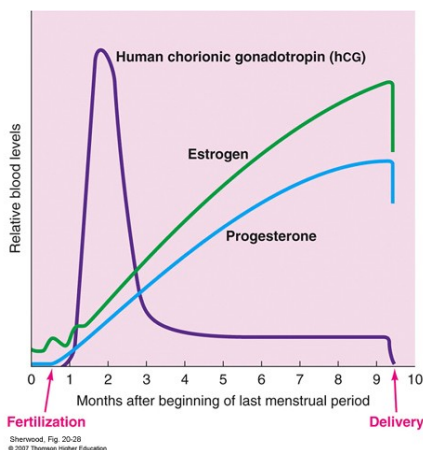
### 3.2.3 Prevalence

The prevalence of HG is low, ranging between 0.3 and 2 % of all deliveries (Ismail et al)<sup>37</sup>. In Norway the prevalence of HG has been estimated 0.8 – 3.2 % dependent on ethnicity (Vikanes et al)<sup>38</sup>. In Canada the incidence of hospitalization for HG was 473 per 100,000 live births (0.5 %) estimated from a large prospective cohort study during 1999 (Bailit)<sup>39</sup>. In an Asian population, the frequency of HG was 3.6 %, estimated from a retrospective cohort study between January 1,2002 and December 31, 2005 in Kaizuka City (Matsuo et al)<sup>40</sup>.

### 3.2.4 Etiology

#### 3.2.4.1 Endocrine factors

Human Chorionic gonadotropin (HCG) has been proposed as the most likely endocrine contributor to the etiology of HG because HCG levels rise rapidly in the first trimester and peak between 10<sup>th</sup> and 12<sup>th</sup> weeks' gestation (Figure 6) corresponding with the onset and peak of women's symptoms of HG (Furmeaux et al)<sup>41</sup>. Further, high HCG levels may cause transient hyperthyroidism because HCG stimulates thyroid activity (Hershman)<sup>42</sup>, (Tan et al)<sup>43</sup>. Also elevated estrogen levels found in HG patients might be induced by HCG and linked to the vomiting (Goodwin et al)<sup>44</sup>. These hypotheses have been supported because obstetric conditions associated with elevated HCG levels were frequently observed in HG patients, such as twin pregnancies (Basso et al)<sup>45</sup> and female sex of the offspring (Askling et al)<sup>46</sup>, (Del Mar Melero-Montes et al)<sup>47</sup>, (Schiff et al)<sup>48</sup>.



**Figure 6.**

HCG levels rise rapidly in the first trimester and peak between 10<sup>th</sup> and 12<sup>th</sup> week of gestation.

#### 3.2.4.2 Genetic factors

Maternal genetic susceptibility developing HG during pregnancy has been hypothesized. As far the suggestion that gestational hyperthyroidism is associated with HG, a mutant thyrotropin receptor hypersensitive to HCG was pinpointed in a case report describing a woman and her mother having recurrent gestational hyperthyroidism and normal HCG concentrations (Rodien et al)<sup>49</sup>. A familial component to HG was demonstrated in a survey (Fejzo et al)<sup>50</sup>. A Norwegian twin study showed a stronger correlation with pregnancy related nausea and vomiting in monozygotic female twins compared to dizygotic female twins (Corey et al)<sup>51</sup>. The role of maternal and paternal genome in HG has been pinpointed and discussed in a Norwegian population-based cohort study (Trogstad et al)<sup>52</sup>, arguing that a maternal genetic factor or permanent maternal environmental factor might be involved in the etiology of HG. Also observed differences in the prevalence of HG dependent on ethnicity have raised questions of possible genetic determinants of HG (Vikanes et al)<sup>38</sup>. Consanguinity was used as a marker for homozygosity in HG associated alleles in the fetus (Grjibovski et al)<sup>53</sup>, but no association between biologically related partners and HG was identified.

#### 3.2.4.3 *Helicobacter pylori* infection

*Helicobacter pylori* (*H.pylori*) have been associated with an increased risk of HG. A meta-analysis based on fourteen published studies showed a four-fold increased risk of HG associated with *H.pylori* (Goldberg et al)<sup>54</sup>.

#### 3.2.4.4 Psychological factors

For many years it was presumed that women suffering from HG were psychologically predisposed to the condition through an infantile or immature personality associated with hysteria (Fairwater)<sup>35</sup>, concluding that psychiatric factors accounted for around 70% of HG cases. Later, it was argued that psychological symptoms might be a result of stress arising from the physical burden of HG rather than being the cause of the condition (Simpson et al.)<sup>55</sup>, (Swallow et al)<sup>56</sup>. Further, psychological explanations of HG have been discussed and criticised from a social science point of view (Munch)<sup>57</sup>. On the other hand, a recent study supports the hypothesis that pre-pregnancy psychiatric conditions might be causally related to HG (Seng et al)<sup>58</sup>.

#### 3.2.4.5 *Other risk factors*

Risk factors of HG were recently evaluated from a population-based cohort of all deliveries in Nova Scotia, Canada between 1988 and 2002 (Fell et al)<sup>59</sup>. The findings from this study showed that maternal conditions as pre-existing diabetes, psychiatric illness, hyperthyroid disorders, gastrointestinal disorders and asthma, and pregnancy related conditions as multiple gestation pregnancy, the presence of a female fetus and history of molar pregnancy were independent risk factors of HG. Increasing maternal age and smoking during pregnancy were associated with decreased risk of HG, a tendency also reported by others (Källén et al)<sup>60</sup>. Ethnicity was pinpointed as a strong risk factor of HG beside female sex of the fetus (Vilming et al)<sup>61</sup>.

#### 3.2.5 Severe consequences of HG

Some severe physical consequences of HG described in the literature are Wernicke's encephalopathy, hepatic dysfunction and renal failure (Blackburn)<sup>62</sup> (De Swiet )<sup>63</sup>. Also the psychosocial burden of HG can profoundly affect multiple aspects of women's lives. Some women terminate their pregnancy due to HG (Poursharif et al)<sup>64</sup>, and socioeconomic changes with respect to career or educational problems, difficulties at work and even job loss have been reported (Poursharif et al)<sup>65</sup>.



## 4 Material and statistical methods

### 4.1 Material

The thesis is based on three different materials:

**Paper I:** Sandven I, Abdelnoor M, Wethe M, Nesheim B-I, Vikanes Å, Gjønnnes H, Melby KK. *Helicobacter pylori* infection and hyperemesis gravidarum. An institution - based case - control study. *European Journal of Epidemiology* 2008; 23: 491-8

The material concerns 244 cases of hyperemesis gravidarum (HG) and 244 controls at Aker University Hospital, Oslo, Norway in the years 1994 - 1999.

**Paper II:** Sandven I, Abdelnoor M, Nesheim B-I, Melby KK. *Helicobacter pylori* infection and hyperemesis gravidarum: a systematic review and meta-analysis of case-control studies. *Acta Obstetrica et Gynecologica Scandinavica* 2009; 88: 1190-1200

The material concerns 25 case-control studies examining the association between *Helicobacter pylori* (H.pylori) and HG.

**Paper III:** Sandven I, Abdelnoor M. *Critical appraisal of case-control studies of risk factors or etiology of hyperemesis gravidarum. Archives of Gynecology and Obstetrics* 2010; 282: 1-10

The material concerns 48 case-control studies examining the etiology or risk factors of HG from the international literature.

### 4.2 Statistical methods

**Paper I:** Sandven I, Abdelnoor M, Wethe M, Nesheim B-I, Vikanes Å, Gjønnnes H, Melby KK. *Helicobacter pylori* infection and hyperemesis gravidarum. An institution - based case - control study. *European Journal of Epidemiology* 2008; 23: 491-8

The odds ratio (OR) and its 95% confidence interval (CI) were used to quantify association between an exposure and disease. A Breslow and Day test of heterogeneity was done to pinpoint effect modification. Confounders were identified and controlled univariately using the Mantel-Haenszel method. The confounding effect was quantified by comparing the

adjusted Mantel-Haenszel OR to the crude OR. Assessment of interaction or effect modification was done also at the additive scale, as it is more important in public health and adequate with biological interaction (Rothman)<sup>12</sup>. Unconditional logistic regression model was used to control for multiconfounders. A Chi-square test of interaction using the log-likelihood ratio was used (Kleinbaum et al)<sup>66</sup>. Information bias, non-differential misclassification of exposure was corrected for using the formula of Kleinbaum et al (Kleinbaum et al)<sup>11</sup>. Dose-response effect was investigated by the Mantel-Haenszel test of linear trend for different exposure levels (Mantel)<sup>67</sup>. The etiologic fraction was estimated in the total study population. All statistical analyses were performed with EpiInfo version 3.3.2.

**Paper II:** *Sandven I, Abdelnoor M, Nesheim B-I, Melby KK. Helicobacter pylori infection and hyperemesis gravidarum: a systematic review and meta-analysis of case-control studies. Acta Obstetrica et Gynecologica Scandinavica 2009; 88: 1190-1200*

#### *Statistical pooling*

In order to calculate the pooled odds ratio (OR) estimate and its confidence interval (CI) we first used a fixed effect model (Mantel and Haenszel method)(Mantel et al)<sup>5</sup>. As the homogeneity hypothesis appeared to be irrelevant, we worked with a random effect model. The estimation was done according to the DerSimonian and Laird method (DerSimonian et al)<sup>68</sup> and fitted for a pooled analysis. For each model a test for overall effect was done. Because of the heterogeneity of our studies the random effect model is more appropriate. Using this model, the estimate of the pooled effect measure and its confidence interval incorporate the additional variability due to inter-study variance ( $\tau^2$ ).

#### *Evaluation of heterogeneity*

Statistical heterogeneity among studies was assessed with the Cochran's Q test (Cochran)<sup>69</sup>. The magnitude of heterogeneity was evaluated by the  $I^2$  statistic that describes the proportion of total variation due to heterogeneity rather than chance (Higgins et al)<sup>70</sup>. In order to investigate possible sources of heterogeneity two different methods were used: subgroup analysis and meta-regression (Thomson et al)<sup>71</sup>. After the sub-group analysis, we run a random effect meta-regression analysis, where the outcome variable was the observed log OR from every study indicating exposition effect and study characteristics (covariates). We used an iterative method with restricted maximum likelihood (REML). A multivariate backward

elimination procedure was performed. Source of heterogeneity was considered as important if the covariate decreased the between study variance. The estimate of  $\tau^2$  in the presence of a covariate in comparison to that when the covariate is omitted, allows the proportion of the heterogeneity variance explained by the covariate to be calculated (Berkey et al)<sup>72</sup> (Thompson et al)<sup>73</sup>.

#### *Publication bias*

In order to assess potential publication bias we used the funnel plot, which is a good visual evaluation of sampling bias. A test of asymmetry (Eggers test of the intercept) applied on the funnel plot was performed (Sterne et al)<sup>74</sup> (Egger et al)<sup>75</sup>. In addition the Orwin faile-safe N test was performed (Orwin)<sup>76</sup>. This test talk about the possibilities that studies are missing from our analysis and that these studies, if included in the analysis, would shift the effect toward the null.

All statistical analyses were performed with the Meta-analysis in Medical Research program developed by the British Medical Journal (Gioacchino)<sup>77</sup>, Comprehensive meta-analysis ([www.meta-analysis.com](http://www.meta-analysis.com)) and STATA 8.0 ([www.stata.com](http://www.stata.com)).

**Paper III:** *Sandven I, Abdelnoor M. Critical appraisal of case-control studies of risk factors or etiology of hyperemesis gravidarum. Archives of Gynecology and Obstetrics 2010; 282: 1-10*

All cases-control studies examining the etiology or risk factors of hyperemesis gravidarum (HG) were considered. Inclusion criteria for studies were according to a general definition of a case-control study.

#### *Search strategy*

A database search was conducted in Medline/PubMed from 1966. Initially the MeSH-terms “Case-Control Studies” and “Hyperemesis gravidarum” were combined. Additionally these terms were searched as text words. Searches were limited to English, French and Scandinavian languages. Closing date of literature search was February 12<sup>th</sup>, 2009. One reviewer (IS) assessed the title and abstract of all identified studies, and made the first decision whether they met the inclusion and exclusion criteria. Full text of eligible papers was

retrieved and read by two reviewers (IS and MA). The final decision about which studies to be included in the review was made in a consensus meeting.

#### *Assessment of studies*

Data regarding methodology issues were abstracted by two reviewers (IS and MA) and entered into a devised data extraction form developed on basis of guidelines for case-control studies (von Elm et al)<sup>78</sup> and checklists for critical appraisal of case-control studies developed by other experienced investigators (Lichtenstein et al)<sup>79</sup> (Elwood)<sup>80</sup>. The checklist included core items assessing a) methods for selecting study participants, b) methods for measuring main exposure variables and c) appropriate use of statistical methods and control for confounding. To assess Type-II error, we recorded whether a power analysis was done. Finally we assessed whether the authors answered causality and discussed internal and external validity of their study. As recommended by Sanderson et al (Sanderson et al.)<sup>81</sup>, we preferred a checklist rather than a scale when assessing study quality.

## 5 Summary of papers

**Paper I:** Sandven I, Abdelnoor M, Wethe M, Nesheim B-I, Vikanes Å, Gjønnnes H, Melby KK. *Helicobacter pylori* infection and hyperemesis gravidarum. An institution - based case - control study. *European Journal of Epidemiology* 2008; 23: 491-8

In this case-control study we used an etiological strategy testing the association between *Helicobacter pylori* (H.pylori) as the exposure and hyperemesis gravidarum (HG) as disease. The source-population was defined as pregnant women admitted to Aker University Hospital (AUH), which served a population mainly from seven out of fifteen city districts in Oslo during the calendar years 1994 throughout 1999. A case of HG was defined according to ICD-9 Code 643.1. Controls were pregnant women, not suffering from HG, consulting the obstetric outpatient clinic. One control on each case was selected consecutively. Total sample size was 488. Exposure to H.pylori was determined by serology testing. Variables of interest as potential confounders or effect modifiers of the association between H.pylori and HG were obtained from medical records, including maternal age, parity, ethnicity, employment and smoking status. We observed that H.pylori was associated with an increased risk of HG, quantified by the odds ratio (OR) and its 95% confidence interval (CI) adjusted for maternal age. African ethnicity was pinpointed as an effect modifier on this association, investigated on the additive scale. A gradient effect in the risk of HG with increasing levels of H.pylori IgG antibodies was found more pronounced in Africans versus non-Africans, using the Mantel-Haenszel test of linear trend. In H.pylori seropositive Africans the etiological fraction was twice as high as compared to non-Africans. One of four HG cases in the total population can be attributable to H.pylori infection on the condition that the association was causal.

**Paper II:** Sandven I, Abdelnoor M, Nesheim B-I, Melby KK. *Helicobacter pylori* infection and hyperemesis gravidarum: a systematic review and meta-analysis of case-control studies. *Acta Obstetrica et Gynecologica Scandinavica* 2009; 88: 1190-1200

In the second paper, we summarized the evidence from 25 case-control studies examining the association between *Helicobacter pylori* (H.pylori) infection and hyperemesis gravidarum (HG). No cohort studies were identified. Included studies, published between 1998 and 2008, provided a pooled sample size of 1455 HG cases and 1970 controls. No publication bias was

indicated, using the funnel plot and a test of asymmetry applied on the funnel plot. Exposure to H.pylori was associated with an increased risk of HG, expressed by the pooled OR and its 95% CI using a random effect model. High heterogeneity between the studies was pinpointed, assessed by the Cochran's Q test and quantified by the index of heterogeneity statistic. Possible sources of heterogeneity between studies was investigated by subgroup analysis and meta-regression considering six study characteristics that could influence the association between H.pylori and HG: method of diagnosing HG, year of publication, method of assessing H.pylori, matching process, geographical location and the north latitude where the study took place. Two of these characteristics were important. A weaker association was found in studies with a clear definition of HG compared to studies without this condition, and weaker association in recent studies compared to earlier studies. Year of publication and a clear definition of HG explained 35% and 20% of between study heterogeneity respectively. Using the random effect meta-regression analysis in the multivariate mode, those two study characteristics explained 40% of between study heterogeneity.

**Paper III:** *Sandven I, Abdelnoor M. Critical appraisal of case-control studies of risk factors or etiology of hyperemesis gravidarum. Archives of Gynecology and Obstetrics 2010; 282: 1-10*

In the third paper, critical appraisal based on 48 case-control studies examining the etiology or risk factors of hyperemesis gravidarum (HG) was performed, using a structured assessment devised for the purpose. All studies were identified by a database search in Medline/PubMed from 1966 until February 12<sup>th</sup> 2009, and included according to predetermined inclusion and exclusion criteria. The study design was rarely indicated with the commonly used term "case-control study" in the title or abstract. The majority investigated hormonal or gastrointestinal risk factors associated with HG (42 studies). Selection bias and non-response received insufficient attention in many articles. Sample sizes varied among the studies, some apparently small without giving any power calculation to justify their size. Association uncommonly was expressed by OR's and CI's (9 studies), confounding rarely controlled for (10 studies) and few examined effect modification (2 studies). Few authors discussed internal validity of their study and the possible impact of bias on study findings. Causality criteria were rarely judged, while authors frequently tried to pinpoint biological plausibility by discussing biological mechanisms explaining their findings. Our study raised concern regarding inadequacies in the design, analysis and reporting of case-control studies of HG

pathology. There is a crucial need to follow correct clinical research methodology and the STROBE guidelines in designing, executing and reporting such studies.

## **6 General discussion**

### **6.1 Etiology of hyperemesis gravidarum**

The last decade, *Helicobacter pylori* (*H.pylori*) has been a candidate etiologic factor for hyperemesis gravidarum (HG). The association between *H.pylori* and HG was investigated in our case-control study (Paper I), and evidence from 25 case-control studies on the same association summarized in our meta-analysis (Paper II). In our third article 48 case-control studies examining the etiology or risk factors of HG were critically appraised (Paper III).

#### 6.1.1 Internal validity

##### *6.1.1.1 Selection bias*

The possibility of selection bias cannot be excluded. Ideally, HG cases should be series of all incident cases from a defined eligible and source population during a specified period of time (Kleinbaum et al)<sup>11</sup>. In our case-control study (Paper I) these recommendations were met, but not in all studies included in the meta-analysis (Paper II) and the 48 studies critical appraised (Paper III). The ideal control group should be representative of non-diseased pregnant women from the same source population during the same time period as the cases are identified (Kleinbaum et al)<sup>11</sup>, and random sampling is an ideal method to select controls (Kupper et al)<sup>82</sup>. In our case-control study (Paper I) control subjects were selected consecutively from the same source population during the same time period as the cases were identified, but random sampling would have been the optimal choice. Random sampling was used in none studies included in our meta-analysis (Paper II) and in five of 48 studies reviewed for critical appraisal (Paper III).

### *6.1.1.2 Information bias*

The potential of recall bias is small. Ideally, past exposure should be determined objectively in cases and controls (Kleinbaum et al)<sup>11</sup>. In our case-control study (Paper I) exposure to *H.pylori* was determined by a serology test, and in 23 of 25 studies included in the meta-analysis (Paper II). Serology tests were also used determining main exposure status in 39 of 48 studies reviewed for critical appraisal (Paper III). When serological tests are used, analyzes is done by laboratory technicians unaware of the status of the patient being a case or control. In those situations there is possibility of non-differential misclassification of exposure, which dilutes the association between exposure and disease (Kleinbaum et al)<sup>66</sup>. The diluted effect can be quantified and corrected for when the sensitivity and specificity of the tests are known (Kleinbaum et al)<sup>11</sup>, which was done in our case-control study (Paper I).

#### *6.1.1.2.1 Non-response, an issue often forgotten*

Non-response in cases and controls could have distorted calculated measures of association between *H.pylori* and HG, and other considered exposures associated with HG. It has been demonstrated that error in the odds ratio due to non-response is determined by the response rates in each of the four cells of the 2 x 2 table (Austin et al)<sup>83</sup>. Feychting et al. (Feychting et al)<sup>16</sup> has demonstrated that when non-response among controls is associated with exposure the estimated odds ratio from a case-control study could be diluted, while the same situation in cases can overestimate any association. But when non-response is independent of exposure, even when the non-response differs in cases and controls, there is no effect on the risk-estimate. Non-response associated with socioeconomic status has been shown to be an important potential source of bias in population-based case-control studies (Richiardi et al)<sup>84</sup>. In our case-control study (Paper I) an equal non-response of 17% in cases and controls was due to missing information. We could have quantified the direction and magnitude of potential bias due to non-response if characteristics of non-respondents had been known. Information on non-response and the reasons for non-response were not provided in individual studies included in the meta-analysis (Paper II). In the 48 case-control studies critically appraised (Paper III), non-response for cases and controls was reported in 9 studies without discussing potential impact on study findings.



### *6.1.1.3 Confounding bias*

There is a potential problem with residual confounding. There are many risk factors of HG, including maternal age, aspects related to social class, reproductive history, somatic, and psychological disorders (Fell et al)<sup>59</sup>. When associated with the exposure of interest, and not in the causal pathway between that exposure and HG, some of these factors could be potential confounders or effect modifiers of the association between exposure and HG. In our case-control study (Paper I) the confounding effect of maternal age was controlled for in the analysis. We could have considered other major risk factors for HG in the protocol as diabetes mellitus, gastrointestinal, endocrine or psychiatric illness, but as none of them are clearly associated with H.pylori it is not sure they would have had any confounding effect. The pooled estimate of the association between H.pylori and HG in our meta-analysis (Paper II) could be distorted due to confounding, when based on crude estimates from individual studies included. Also among the 48 case-control studies reviewed and critically appraised (Paper III), a high proportion of effect measures of associations could be biased due to confounding simply because potential confounders were not considered and controlled for in the analysis or matching not always followed using a matched statistical test. When the matching variables are important, an unmatched analysis would distort the association between the exposure of interest and HG (Rosner et al)<sup>85</sup> and residual confounding contribute to further distortion.

#### *6.1.1.3.1 Interaction, a word with two meanings*

Statistical interaction is a model-dependent concept. Statistical interaction is considered to be present on a multiplicative scale when the joint effect of risk factors differs from the product of the effect of the individual factors, and present on an additive scale when the joint effect of risk factors differs from the sum of the effects of the individual factors (Ahlbom et al)<sup>86</sup>. Biological interaction is defined using the general causal model presented by Rothman (Rothman)<sup>12</sup>, and refers to the causal responsibility of a joint effect of two or more component causes of a sufficient cause of disease. An additive model is most helpful pinpointing biological interaction (Rothman)<sup>12</sup>, which is important in public health.

Ethnicity could be an effect modifier of the association between H.pylori and HG (Paper I). In the stratified analysis, using the Mantel-Haenzel method, we found a difference in magnitude but not in direction of the OR in the African group (OR = 5.71) as compared to non-Africans

(OR = 1.72). The Breslow and Day test of heterogeneity was non-significant ( $p = 0.15$ ), but the test has low power and larger sample size could have led to significance (Marshall)<sup>87</sup>. We also tried to investigate the existence of interaction using logistic model, but the  $p$ -value for the interaction variable in the model, which included ethnicity and H.pylori, was close to the one of Breslow and Day ( $p = 0.15$ ). These findings indicated non-existence of interaction on the multiplicative level.

Using an additive model, shown by the table and equation below, we found that the effect of H.pylori and African ethnicity acting in combination is stronger than the effect when they operate independently (OR = 6.7 > OR = 1.07), suggesting synergism.

	Both	H.pylori only	African only	Neither (reference)	Total
	H.pylori <sup>yes</sup> African <sup>yes</sup>	H.pylori <sup>yes</sup> African <sup>no</sup>	H.pylori <sup>no</sup> African <sup>yes</sup>	H.pylori <sup>no</sup> African <sup>no</sup>	
HG <sup>yes</sup>	40	65	4	135	244
HG <sup>no</sup>	7	51	4	182	244
Total	47	116	8	317	488
Odds	5.7143	1.2745	1.0000	0.7418	
OR (95% CI)	7.70 (3.35 – 17.73)	1.72 (1.12 – 2.64)	1.35 (0.33 – 5.49)	1.0	

No interaction on the additive scale when

$$\left[ \left( OR_{H.pylori+,Africa+} \right) - 1 \right] = \left[ \left( OR_{H.pylori+,Africa-} \right) - 1 \right] + \left[ \left( OR_{H.pylori-,Africa+} \right) - 1 \right]$$

$$[7.7 - 1] > [(1.72 - 1) + (1.35 - 1)]$$

$$\underline{6.7 > 1.07}$$

We quantified the amount of biological interaction according to three measures presented by Rothman (Rothman)<sup>12</sup>: relative excess risk due to interaction (RERI), attributable proportion due to interaction (AP) and synergy index (S). In our data synergism was indicated showing departure from additivity (S = 6.26), a strong effect of H.pylori and African ethnicity due to interaction (RERI = 5.63) and this interaction could be responsible for 73% of HG cases (AP = 0.73) illustrated in Figure 7.

Measures of additive interaction based on OR:

$$RERI = RR_{H.pylori+,African+} - RR_{H.pylori+,African-} - RR_{H.pylori-,African+} + 1 = 7.7 - 1.72 - 1.35 + 1 = 5.63$$

$$AP_{H.pylori+,African+} = \frac{RERI}{RR_{H.pylori+,African+}} = \frac{5.63}{7.7} = 0.73$$

$$S = \frac{RR_{H.pylori+,African+} - 1}{RR_{H.pylori-,African+} + RR_{H.pylori+,African-} - 2} = \frac{7.7 - 1}{1.35 + 1.72 - 2} = 6.26$$

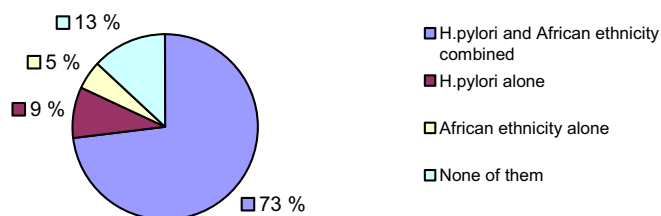
Calculating measures of additive interaction based on OR using R (Källberg et al)<sup>88</sup>, we got approximately identical estimates with its 95% confidence limits.

RERI (95% CI) = 5.64 (-0.94 – 12.21)

AP (95% CI) = 0.73 (0.40 – 1.07)

S (95% CI) = 6.29 (0.77 – 51.45)

**Figure 7. Pie model illustrating the proportion of HG due to interaction between H.pylori and African ethnicity**  
(estimates based on OR)



The use of odds ratios rather than risk ratios tends to exaggerate the interaction. The divergence between results varies with the baseline risk and the magnitude of interaction (Kalilani et al)<sup>89</sup>. For more than additive interaction, the difference is more pronounced for ICR (RERI) and S measures, and for less than additive interaction it is more marked for AP. Even when the outcome is rare, the use of odds ratios may point to interaction (additive or multiplicative) in instances where the use of risk ratios would indicate the absence of interaction (Campbell et al)<sup>90</sup>. When we calculated measures of additive interaction based on

the relative risk, the three measures of additive interaction differed in magnitude as compared to estimates based on the odds ratio, but still indicating synergism (ICR = 0.5, AP = 0.25, S = 2.02). Shown below, the RR is calculated based on the OR's available in our study (Paper I) and substituted in the equations calculating measures of additive interaction according to Kalilani et al (Kalilani et al)<sup>89</sup>:

$$RR_{Africa+,Hp+} = \frac{OR_{Africa+,Hp+} (1 + odd_{Africa-,Hp-})}{1 + odd_{Africa+,Hp+}} = \frac{7.7(1 + 0.7418)}{1 + 5.7143} = 1.9975$$

$$RR_{Africa+,Hp-} = \frac{OR_{Africa+,Hp-} (1 + odd_{Africa-,Hp-})}{1 + odd_{Africa+,Hp-}} = \frac{1.35(1 + 0.7418)}{1 + 1.0} = 1.1757$$

$$RR_{Africa-,Hp+} = \frac{OR_{Africa-,Hp+} (1 + odd_{Africa-,Hp-})}{1 + odd_{Africa-,Hp+}} = \frac{1.72(1 + 0.7418)}{1 + 1.2745} = 1.3172$$

No interaction on the additive scale when

$$[(RR_{H.pylori+,Africa+}) - 1] = [(RR_{H.pylori+,Africa-}) - 1] + [(RR_{H.pylori-,Africa+}) - 1]$$

$$(1.9975 - 1) > [(1.3172 - 1) + (1.1757 - 1)]$$

$$0.9975 > 0.4929$$

$$1 > 0.5$$

$$ICR(RERI) = RR_{H.pylori+,African+} - RR_{H.pylori+,African-} - RR_{H.pylori-,African+} + 1$$

$$= 1.9975 - 1.1757 - 1.3172 + 1 = \underline{0.5046}$$

$$AP_{H.pylori+,African+} = \frac{ICR(RERI)}{RR_{H.pylori+,African+}} = \frac{0.5046}{1.9975} = \underline{0.25}$$

$$S = \frac{RR_{H.pylori+,African+} - 1}{RR_{H.pylori-,African+} + RR_{H.pylori+,African-} - 2} = \frac{(1.9975 - 1)}{[(1.1757 - 1) + (1.3172 - 1)]} = \underline{2.02}$$

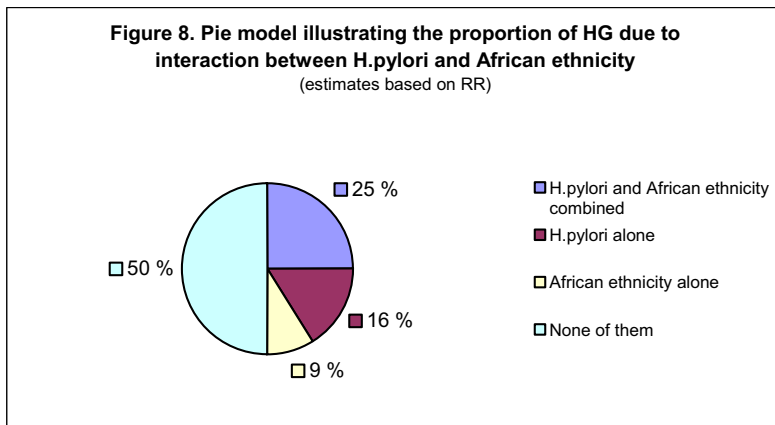
Measures of additive interaction based on RR calculated using ETCETRA Version 2.25:

$$ICR (95\% CI) = 0.53 (-0.71 - 1.20)$$

$$AP (95\% CI) = 0.23 (-0.36 - 0.58)$$

$$S (95\% CI) = 2.12 (0.33 - 13.41)$$

Comparative table		
Measures of additive interaction	Based on OR	Based on RR
<b>RERI or ICR:</b> departure from zero indicates synergism	5.63	0.50
<b>AP:</b> departure from zero indicates synergism	0.73	0.25
<b>S:</b> departure from one indicates synergism	6.26	2.02



If we want to assess the evidence in our data of a possible interaction effect between H.pylori and African ethnicity measured as deviation from additivity, we can test the statistical effect (significance) of this interaction by using the procedure of Hogan (Hogan et al)<sup>91</sup>. In our data Hogan's test was non-significant ( $p = 0.24$ ), but this might be due to small sample size.

An elevated risk of HG in H.pylori infected Africans compared to non-Africans might be explained by socioeconomic differences, as ethnicity and low socioeconomic status are correlated to H.pylori and HG (Vilming et al)<sup>61</sup> (Frigo et al)<sup>92</sup> (Karaca et al)<sup>93</sup> (Jimenez et al)<sup>94</sup> (Segal et al)<sup>95</sup> (Naja et al)<sup>96</sup>. On the other hand, H.pylori infected Africans possibly carries an aggressive variant of the bacterium and host immune mechanisms might be the key to different responses to H.pylori in different populations and geographic areas (Segal et al)<sup>95</sup> (Linz et al)<sup>97</sup>.

Interaction or effect modification was not examined in the other studies included in the meta-analysis (Paper II), except our (Paper I)). Among the 48 case-control studies reviewed for critical appraisal (Paper III) another study investigated interaction by age, parity and smoking

on the relationship between infant sex and HG (Schiff et al)<sup>48</sup>, but the investigators found no evidence of interaction on the multiplicative scale (using the likelihood ratio test).

#### *6.1.1.4 Heterogeneity between studies included in the meta-analysis (Paper II)*

A high heterogeneity indicates genuine differences in the 25 case-control studies put together in the meta-analysis (Paper II). Statistical heterogeneity is a consequence of clinical and methodological differences between studies (Higgins et al)<sup>98</sup>. Perhaps the most obvious diversity could be related to differences in the patient populations studied. The severity of HG may differ because there is no universally accepted definition of HG and a number of pregnancy-related conditions and differential diagnosis might be confused with HG (Goodwin)<sup>36</sup>. Hypothetically an unclear definition of HG could overestimate an association between *H. pylori* and HG, because detection bias in case-control studies usually overestimates the association between exposure and disease (Choi et al)<sup>99</sup>. As far as a stronger effect was pointed to in older studies, the reasons could be either a higher degree of exposition to *H. pylori* earlier or changes in *H. pylori* strains over time (Rupnow et al)<sup>100</sup>. Heterogeneity also may be due to different methods of ascertaining exposure of *H. pylori* infection. Mainly serology tests were used and their positive and negative predictive value varies dependent on the prevalence of *H. pylori* in the populations under study (Ricci et al)<sup>101</sup>, which is known to vary from country to country and between subpopulations within the same country (Malaty et al)<sup>102</sup>.

All studies were case-control designs testing the same hypothesis, but heterogeneity may arise through different control over bias. Information on participation rates and the reasons for non-participation were not provided in individual studies included in the meta-analysis (Paper II), which is a recognized problem in case-control studies (Olson et al)<sup>103</sup>. If the selection of HG cases and control subjects has been dependent on their exposure to *H. pylori* for different reasons (Feychting et al)<sup>16</sup> (Austin et al)<sup>83</sup> (Criqui)<sup>104</sup>, this will introduce a bias in the estimation of the OR's in included studies and contribute further to heterogeneity. Recall bias is unlikely because biological testing of exposure to *H. pylori* was used in all studies for both cases and controls, but non-differential misclassification of exposure would have diluted the effect estimate dependent on the sensitivity and specificity of the test. Different ELISA techniques from different manufacturers were used for ascertaining exposure status, but overall the sensitivity and specificity of available tests was above 90% (Ricci et al)<sup>101</sup>. A

diluted effect was quantified to be 28% in our case-control study (Paper I). Individual studies have failed to control for potential confounders. Well-documented risk factors for HG such as age, smoking and ethnicity (Ismail et al)<sup>37</sup> (Fell et al)<sup>59</sup> (Källén et al)<sup>60</sup> also associated with H. pylori infection (Naja et al)<sup>96</sup> (Suerbaum et al)<sup>105</sup>, are potential hidden confounders, which could have distorted the observed association between H. pylori and HG. In nine studies, matching maternal age, gestational age or parity, did not overcome the problem of heterogeneity. The effect of H. pylori infection could differ dependent on ethnicity for biological, immunological or other reasons (Aatherton)<sup>106</sup>, but assessment of interaction or effect modification was done in our study only (Paper I).

#### 6.1.2 Consideration of positive features of causation on the association between H.pylori and hyperemesis gravidarum

Summarizing non-causal explanations, we believe information bias has been well controlled, but non-response could have distorted our results and there is some non-differential misclassification. Confounding has been adequately controlled by stratification and multivariate methods in the case-control study (Paper I), but there is possibility of residual confounding. In the meta-analysis (Paper II) the problem of confounding is more pronounced because the pooled estimate is based on crude effect from individual studies included. Chance variation can be confidently excluded because the size of the case-control study (Paper I) and meta-analysis (Paper II) gives narrow confidence intervals for the effect estimates.

Presence of a valid statistical association does not imply causality. Hill's criteria are an aid in the judgement concerning causality, including strength of the association, biologic plausibility of the hypothesis, consistency of the findings, as well as other information concerning the temporal sequence and the presence of a dose-response relationship (Hill)<sup>107</sup>.

##### 6.1.2.1 *Is there a correct time relationship?*

Information on exposure to H.pylori probably applies to a relevant time period, before clinical diagnosis of HG. We believe that primary H.pylori infection is uncommon among HG cases in our study (Paper I) because sera taken prior to present pregnancy, available in 25% of Africans and 35% of non-Africans in the study population, were all anti HpIgG positive. Initial H.pylori colonisation usually occurs in childhood, but the immune response rarely

eliminates the organism from the gastric mucosa and a latent, asymptomatic infection can persist in adults in the absence of treatment (Malaty)<sup>102</sup> (Robinson et al)<sup>108</sup>.

#### *6.1.2.2 Is the association strong?*

The association between H.pylori and HG is only moderately strong. For all women in our case-control study the presence of H.pylori increased the risk of HG around two-fold (Paper I), and around three-fold in our meta-analysis (Paper II). Because of a high prevalence of H.pylori in the general population, a risk of HG around two to three in women carrying H.pylori is important in clinical and public health terms. Assuming the observed association between H.pylori infection and HG as causal, one of four HG cases in the total population can be attributable to H.pylori infection (Paper I). On the other hand, such relatively weak associations observed might reflect that H.pylori is a minor contributor to HG risk compared with other factors, as strength of association depends on the prevalence of other factors and confounding (Rothman)<sup>109</sup>. Further it raise the question of whether H.pylori is merely a rather inaccurate estimator of a more fundamental causal factor, which if identified might show a much stronger association.

#### *6.1.2.3 Is there a dose-response relationship?*

A gradient relationship in the risk of HG with increasing levels of HpIgG Ab concentration is an important finding in our case-control study (Paper I). A gradient relationship was found in African and non-African women separately, and in each of these groups the Mantel-Haenszel chi-square test for linear trend was significant. As far the biological plausibility of a dose-response association, it has been suggested that a shift in pH in the gastrointestinal tract during pregnancy, due to an increased accumulation of fluid caused by elevated steroid hormones, could lead to the manifestation of a sub clinical H.pylori infection (Verberg et al)<sup>110</sup>. From biopsies of the gastric antrum and corpus, Bagis et al (Bagis et al)<sup>111</sup> found significantly higher H.pylori density, degree of inflammation and H.pylori activation in their HG patients and concluded that the degree of gastric complaints may be associated with density of H.pylori. Detection of IgG antibodies to various antigens of H.pylori has been used for diagnosis of the infection and also for determining eradication of the infection after antibiotic treatment. A fall of 50% in antibody levels after 6 months (Kosunen et al)<sup>112</sup> or a



40% reduction after 3-6 months (Lerang et al)<sup>113</sup> indicate successful therapy i.e. H.pylori elimination. Thus, high levels of HpIgG Ab may reflect high density of bacteria present.

#### *6.1.2.4 Is there a biological plausible explanation?*

A potential mechanism whereby H.pylori infection may increase the risk of HG might be explained by physiological changes in pregnancy and the virulence of the bacterium (Paper I). First, gastric emptying and intestinal transit time may be prolonged during pregnancy due to elevated serum concentration of steroid hormones and their relaxant effects on gastrointestinal smooth muscle (Zimmermann)<sup>114</sup>. Secondly, it has been assumed that pregnancy increases women's susceptibility to infections due to some form of immune depression in the mother to allow for the survival of the fetal semiallograft (Manyonda)<sup>115</sup>, and latent infections might be reactivated in pregnancy (Stagno et al)<sup>116</sup>. Hypothetically, decreased gastrointestinal motility and a depressed immune defence during pregnancy could result in the manifestation of a sub-clinical H.pylori infection, which in turn irritates the gastrointestinal tract. Finally, the inflammatory response to H.pylori might be determined by the virulence of the infecting strain and the human host (Robinson et al)<sup>108</sup>. As highlighted by Xia et al (Xia et al)<sup>117</sup>, HG cases in that study were more likely infected with H.pylori strains carrying the CagA<sup>+</sup> gene as compared to the controls. Further, it is interesting that antibiotics have relieved symptoms of nausea and vomiting in H.pylori positive HG patients (Strachan et al)<sup>118</sup> (Jacoby et al)<sup>119</sup> (El Younis et al)<sup>120</sup> (Mansour et al)<sup>121</sup>.

#### 6.1.3 Web of causation

Probably the complete causal mechanism of HG involves a multitude of factors of which some unknown. As HG occurs also among H.pylori non-exposed pregnant women, there must be different sufficient causes of which some do not contain H.pylori as one of the component causes. The concept of multi-causality is well known, and implies that diseases have several causes acting together in one or several different mechanisms. Rothman has presented the theoretical causal mechanism for a given disease depicted as a sliced pie (Rothman)<sup>12</sup>. Each "pie" illustrates a sufficient cause made up by many component causes, with the slices being the component causes. Different causal mechanisms correspond to different sufficient causes.

### 6.1.3.1 Endocrine factors

Hormonal theories are related to rapidly increasing and high levels of human chorionic gonadotropin (HCG), estrogens and light increased thyroxine without clinical signs of hyperthyreosis. With the onset of pregnancy thyroid function changes (Blackbourn)<sup>62</sup>, (De Swiet)<sup>63</sup>. HCG has a mild thyroid stimulating activity because  $\beta$ -subunit of HCG and thyroid stimulating hormone (TSH) stimulates the same receptors resulting in a slight increase of thyroxine. Further, increasing levels of estrogens stimulates the formation of thyroxine binding globulin (TBG), which in turn increases circulating thyroxine. Renal iodide clearance increases in pregnancy resulting in higher iodide loss. The prevalence of transient hyperthyroidism among HG patients was estimated to be 66% in one study (Goodwin et al)<sup>122</sup> and 88% in another (Tan et al)<sup>43</sup>. Individual studies included in the 48 case-control studies reviewed for critical appraisal (Paper III) have found increased HCG- and thyroxine-levels, and lower TSH-levels in HG cases compared to controls, suggesting a possible linkage between alterations in pregnancy related physiological mechanisms and risk of HG (Goodwin et al)<sup>44</sup> (Goodwin et al)<sup>122</sup> (Bruun et al)<sup>123</sup> (Kauppila et al)<sup>124</sup> (Swaminathan et al)<sup>125</sup> (Tareen et al)<sup>126</sup> (Leylek et al)<sup>127</sup> (Leylek et al)<sup>128</sup> (Jordan et al)<sup>129</sup> (Al-Yatama et al)<sup>130</sup>. Conversely, some studies found no association between HCG, thyroid hormones and HG (Juras et al)<sup>131</sup> (Wilson et al)<sup>132</sup> (Panesar et al)<sup>133</sup>. Another mechanism discussed by which HCG can cause excessive vomiting in early pregnancy, is through the effect of estrogens. Greater concentrations of estrogens in women with HG compared to controls have been reported (Goodwin et al)<sup>44</sup> (Depue et al)<sup>134</sup>, arguing that HCG stimulates production of estrogens and that the clinical finding of nausea and vomiting in HG might be similar to the common occurrence of nausea and vomiting in some women taking combined oral contraceptive pill compared to those taking the progesterone only pill. Further, it has been hypothesized that severe cases of HG might have higher estrogen levels than less severe cases (Sciff et al)<sup>48</sup>. Theories of metabolic factors associated with HG were also investigated. Such as leptin, produced in fat tissues and placenta during pregnancy and regulates appetite and energy expenditure. Leptin might induce a more severe nausea and vomiting due to a decrease in appetite (Üncel et al)<sup>135</sup> (Aka et al)<sup>136</sup> (Demir et al)<sup>137</sup>. Low levels of lipids, might be due to starvation or hyperthyroidism (Üstün et al)<sup>138</sup>. High intake of saturated fat, might increase circulating estrogens level (Signorello et al)<sup>139</sup>. Pre-pregnancy underweight gives low circulating estrogens prior to pregnancy and might exaggerate the response to an abrupt rise in estrogen in the first trimester (Rochelson et al)<sup>140</sup>.

Secretion of more thyroid hormone during early pregnancy in response to the thyroid stimulating activity of HCG and estrogens might cause HG in some women. Based on the studies referred to above, it is difficult to judge whether increased hormone levels are causally related to HG. We have not been able to identify cohort studies examining hormonal factors associated with HG. One reason can be that because HG has a low prevalence, around 0.3 – 2% of all pregnancies, the case-control study is the optimal design.

#### *6.1.3.2 Psychological factors*

The temporal relationship between psychological factors and HG is not easily determined. Psychological symptoms might be a result of stress arising from the physical burden of HG rather than being the cause of the condition (Simpson et al)<sup>55</sup>, (Swallow et al)<sup>56</sup>. On the other hand, having a pre-pregnancy psychiatric diagnosis has been found associated with increased risk of HG (Seng et al)<sup>58</sup>.

#### *6.1.3.3 Genetic factors*

There have been some studies investigating the genetic aspect in the etiology of HG (Vikanes et al)<sup>38</sup> (Rodien et al)<sup>49</sup> (Fejzo et al)<sup>50</sup> (Corey et al)<sup>51</sup> (Trogstad et al)<sup>52</sup> (Grijibovski et al)<sup>53</sup>, but those studies were not completely adequate to answer this question. Unfortunately the traditional study designs in genetic epidemiology, as family-studies, twin-studies, adoption studies, and migrant-studies can only assess the relative contribution of genes to disease occurrence (Khoury)<sup>141</sup>. Advances in molecular genetic technology in the 1970<sup>ies</sup> led to the Human Genome Project, a long-term initiative to map and sequence the human genome completed in 2003 ([www.genomics.energy.gov](http://www.genomics.energy.gov)), and has made it increasingly feasible to search for disease susceptibility genes. Genes play an important role in the etiology and pathogenesis of essentially all human diseases and genetic risk factors might interact with the environment causing various diseases. Unlike environmental factors, such as social, physical, chemical, biologic and infectious agents, genetic markers do not change with time and would be stable indicators of host susceptibility. Investigating the etiology of HG further, we suggest a need for well-conducted studies to quantify the impact of genes and interaction between gene-environmental factors that might play a causal role in developing HG.

6.1.4. Need to include gene-environment interaction in the etiology of hyperemesis gravidarum

*The case-control design* is increasingly applied to study etiology and to search for gene-environment interaction (Khoury et al)<sup>142</sup>. To evaluate gene-environmental interaction in the context of a case-control study, one can display data in a 2 x 4 table shown by Khoury et al (Khoury et al)<sup>143</sup>:

Exposure	Susceptibility genotype	Cases	Controls	OR	Main information
-	-	a	b	1	Common reference
-	+	c	d	$OR_g = \frac{bc}{ad}$	Genotype alone versus none
+	-	e	f	$OR_e = \frac{be}{af}$	Environmental factor alone versus none
+	+	g	h	$OR_{ge} = \frac{bg}{ah}$	Joint genotype and environmental factor versus none

There is no interaction when  $[OR_{ge} = (OR_g + OR_e - 1)]$  under the additive model or when  $[OR_{ge} = OR_g * OR_e]$  under the multiplicative model. To assess for interaction on a

multiplicative scale, one can define a synergy index  $SI = \frac{OR_{ge}}{OR_g * OR_e}$

A  $SI > 1$  indicates more than multiplicative effects, and  $SI < 1$  indicates less than multiplicative effects.

Like in other areas, case-control studies of genetic factors are prone to bias. As discussed by Khoury et al (Khoury et al)<sup>142</sup>, the association may be due to confounding by for instance race or ethnicity and we need to stratify on them. The stability of the gene marker in time is protecting us against recall bias, but misclassification might occur and distort the magnitude of the risk estimate. Large sample sizes is needed to detect small to moderate gene-environment interaction effects (Foppa et al)<sup>144</sup>, and thus mandatory to perform adequate power and sample size calculations.

**The case-only design** has also been promoted as an approach to assess gene-environment interaction in disease etiology. In this option, only case subjects are used to assess the magnitude of the association between the exposure of interest and the susceptibility genotype. The basic setup for analysis is a new 2 x 2 table shown by Khoury et al (Khoury et al)<sup>143</sup>:

Exposure	Susceptibility genotype	
	-	+
-	a	b
+	c	d

Under assumption of independence between exposure and genotype in the source population from which cases arose, the case-only odds ratio  $COR = \frac{ad}{bc}$  and becomes simply the synergy index on a multiplicative scale

derived from a regular case-control study as shown above  $SI = \frac{OR_{ge}}{OR_g * OR_e}$

Under the null-hypothesis of no multiplicative interaction the COR and SI are expected to be one.

Major limitations of the case-only approach discussed by Khoury et al (Khoury et al)<sup>143</sup>, includes that only the interaction between exposure and genotype can be evaluated, not their independent effects. Moreover, the measure obtained from the analysis can be interpreted as a departure from multiplicative effects only, while departure from additivity will be missed. However, with sufficient sample size and for many biological interactions rather extreme positive departures from multiplicative effects can be identified (Khoury et al)<sup>145</sup>, which must reflect even greater departure from additivity.

## 6.2 Critical appraisal of case-control studies on the etiology or risk factors of hyperemesis gravidarum

Our review presents a survey of 48 case-control studies examining the etiology or risk factors of HG (Paper III). Critical appraisal concerns in large methods for selecting study participant, methods for measuring main exposure variables, appropriate use of statistical methods and control for confounding. Finally we assessed whether the authors discussed internal and external validity of their study and answered causality. Our results indicate that a substantial proportion of articles contain weaknesses in basic elements of study design and analysis, and show a tendency for obstetric researchers to ignore the potential impact of bias on their results. We suggest that there is a serious risk that some publications reach misleading

conclusions. Our concern is concordant with the conclusion of other investigators who have done critical appraisal of epidemiological papers published in the research fields of psychiatry, paediatrics, obstetrics, genetics and cancer (Pocock et al)<sup>146</sup> (Lee et al)<sup>147</sup> (Bogardus et al)<sup>148</sup> (Skovlund et al)<sup>149</sup> (Smidt et al)<sup>150</sup> (Rushton et al)<sup>151</sup> (Hellems et al)<sup>152</sup> (Mihailovic et al)<sup>153</sup> (Dauphinee et al)<sup>154</sup>. They concluded with great concern about the validity of published articles in mainstream and high impact journals. We are adding our voice to them in the field of HG, but it is important to underline that our survey is limited to case-control studies only. The fact that mislabelling of case-control studies is a problem in medical journals (Hellems et al)<sup>152</sup> (Mihailovic et al)<sup>153</sup> might raise the question whether our sample is representative of all case-control studies on HG. Recently, around one third of case-control studies were assessed as mislabelled in obstetrics and gynaecology journals (Grimes)<sup>20</sup>. Moreover, our judgement of methodological quality of individual studies was based solely on the information contained in the articles and may be open to interpretation.

### 6.2.1 Problems with the tool for critical appraisal

Unfortunately there is no agreed “gold standard” appraisal tool evaluating quality in observational epidemiologic studies, and there are many tools from which to choose. For the purpose of critically appraising case-control studies we developed a checklist based on recommendations by experienced investigators (Lichtenstein et al)<sup>79</sup> (Elwood)<sup>80</sup>, including core items assessing a) methods for selecting study participant, b) methods for measuring main exposure variables and c) appropriate use of statistical methods and control for confounding. In our opinion, the checklist worked well summarizing characteristics of the studies. On the other hand, when the quality of reporting was poor we had difficulties in extracting from the papers some items on our checklist, leaving room for interpretation by the assessor. The poor reporting on important methodological details may be due to the fact that most papers were published before the STROBE statements (von Elm et al)<sup>78</sup> encouraged higher quality of reporting observational epidemiological studies. Reporting quality has been found associated with methodological quality (Huwiler-Müntener et al)<sup>155</sup>, but similar quality of reporting may also hide differences in methodological quality and well-conducted studies may be reported badly. A clear distinction between reporting quality and methodological quality has been encouraged and scales predominantly measuring quality of reporting should not be used measuring methodological quality. Sanderson et al (Sanderson et al)<sup>81</sup> have evaluated a large number of tools for assessing quality of observational epidemiological

studies, recommending a simple checklist including a small number of key domains being specific to the study design under consideration rather than a scale. Scales resulting in numerical summary of quality scores might look simple to use, but such an approach can introduce a bias when assessing study quality (Greenland et al)<sup>156</sup>. Summary scores involve weighting of items included in the scale and not all of them being necessarily related to validity of study findings, such as sample size and power calculation. Because it might be unclear how weights for different items should be determined, different scales may reach different conclusions on the overall quality of an individual study (Jüni et al)<sup>157</sup>. Similar considerations might apply to checklists, but a qualitative rather than quantitative summary judgement may be less prone to inappropriate analysis. The tool we designed for our purpose represents a checklist approach that concentrate on few, principal, potential sources of bias in case-control studies, as recommended by experienced investigators (Lichtenstein et al)<sup>79</sup> (Elwood)<sup>80</sup> (Sanderson et al)<sup>81</sup>, but obviously the validity and reliability of the tool need be evaluated further.

## **7 Conclusions**

### **7.1 The case-control and case-based designs in obstetrics and gynecology**

Case-control studies are still poorly reported in our field of the obstetric literature and there is urgent need to follow the STROBE guidelines in reporting such studies. Basic elements of the design and analysis of case-control studies are not optimally implemented in our field of research. In the future, incident cases need to be used rather than prevalent cases. Controls need to be randomly selected from the same source population and during the same period of time as the cases derives, rather than consecutively. Matching has been used and abused to control for confounding instead of stratified and multivariate analysis. Alternative case-based designs tend to be used infrequently in the field of obstetrics and gynecology. In the future it would be interesting to use more the case-cohort and nested case-control design or case-control approaches without controls.

## 7.2 Etiology of Hyperemesis gravidarum

The etiology of HG is likely multi-factorial some of which not yet known. H.pylori might be one important factor, but there are other important components too. The findings from individual studies reviewed for my thesis demonstrate complex physiological mechanisms involved in the etiology of the disease, and genetic and environmental factors might affect pregnant women's susceptibility to HG. We reviewed literature published in English only, while access to non-English scientific literature might better serve epidemiology (Fung)<sup>158</sup>. In the future, we need to access the global epidemiological literature in different languages, mainly the official languages of the United Nations: Arabic, Chinese, English, French, Spanish and Russian.

### 7.2.1 Future perspectives

Research on the etiology of HG is limited and conclusive results from high-quality studies uncommon. In the future it would be interesting to study interaction on gene-environmental factors associated with HG, because genetic markers are stable indicators of host susceptibility to HG unlike biologic markers of other types of exposure. The case-control design is well suited for investigating associations of several exposures with a disease outcome. In the future it would be interesting to use more the classical case-control design or case-only design, as described by Khoury et al (Khoury et al)<sup>143</sup>, dealing with genetic components to HG.



## 8 References

1. Paneth N, Susser W, Susser M. Origins and early development of the case-control study: part 1, Early evolution. *Soz-Praventivmed* 2002; 47(5): 282-8
2. Paneth N, Susser E, Susser M. Origins and early development of case-control study: part 2, The case-control study from Lane-Clypon to 1950. *Soz-Praventivmed* 2002; 47(6): 359-65
3. Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *Br Med J* 1952 Dec 13; 2(4797): 1271-86
4. Cornfield J. A method of estimating comparative rates from clinical data. Applications to cancer of the lung, breast, and cervix. *J Natl Cancer Inst* 1951 Jun; 11(6): 1269-75
5. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Nat Cancer Inst* 1959 Apr; 22(4): 719-48
6. Cole P. The evolving case-control study. *J Chronic Dis* 1979; 32(1-2): 15-27
7. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med* 1971 Apr 15; 284(15): 878-81
8. Reingold AL, Hargrett NT, Shands KN, Dan BB, Schmid GP, Strickland BY, Broome CV. Toxic shock syndrome surveillance in the United States, 1980 to 1981. *Ann Intern Med.* 1982 Jun; 96(6 PT 2): 875-80
9. Ziel HK, Finkle WD. Increased risk of endometrial carcinoma among users of conjugated estrogens. *N Engl J Med* 1975 Dec 4; 293(23): 1167-70
10. Smith DC, Prentice R, Thomson DJ, Herrmann WL. Association of exogenous estrogen and endometrial carcinoma. *N Engl J Med* 1975; Dec 4: 293(23) 1164-7
11. Kleinbaum DG, Sullivan KM, Barker ND. *ActivEpi Companion Textbook*. A supplement for use with *ActivEpi CD-ROM*. USA: Springer Science + Business Media; 2003
12. Rothman KJ. *Modern Epidemiology*. USA: Little, Brown and Company; 1986
13. Horwitz RI, Feinstein AR. Alternative analytic methods for case-control studies of estrogens and endometrial cancer. *N Engl J Med* 1978 Nov 16; 299(20): 1089-94

14. Hutchison GB, Rothman KJ. Correcting a bias? *N Engl J Med* 1978 Nov 16; 299(20): 1129-30
15. Lindfors-Harris BM, Eklund G, Adami HO, Meirik O. Response bias in a case-control study: analysis utilizing comparative data concerning legal abortions from two independent Swedish studies. *Am J Epidemiol* 1991 Nov 1; 134(9): 1003-8
16. Feychting M, Ahlbom A, Björk J, Strömberg U. Alarming tendency of increased number of dropouts in epidemiologic case-control studies. *Lakartidningen* 2004 Oct 28; 101(44): 3430-3
17. Madigan MP, Troisi R, Potischman N, Brogan D, Gammon MD, Malone KE, Brinton LA. Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women. *Int J Epidemiol* 2000 Oct; 29(5): 793-8
18. Taubes G. Epidemiology faces its limits. *Science* 1995 Jul 14; 269(5221): 164-9
19. Sørensen HT, Gillman MW. Matching in case-control studies. *BMJ* 1995 Feb 4; 310(6975): 329-30
20. Grimes DA. "Case-Control" confusion: mislabeled reports in obstetrics and gynecology journals. *Obstet Gynecol* 2009 Dec; 114(6): 1284-6
21. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *J Clin Epidemiol* 1999 Dec; 52(12): 1165-72
22. Borgan Ø, Samuelsen SO. A review of cohort sampling designs for Cox's regression model: Potentials in epidemiology. *Norsk Epidemiologi* 2003; 13(2): 239-48
23. Bell EM, Hertz-Picciotto I, Beaumont JJ. Case-cohort analysis of agricultural pesticide applications near maternal residence and selected causes of fetal death. *Am J Epidemiol* 2001 Oct 15; 154(8): 702-10
24. Wernli KJ, Ray RM, Gao DL, Fitzgibbons ED, Camp JE, Astrakianakis G, Seixas N, Li W, De Roos AJ, Feng Z, Thomas DB, Checkoway H. Occupational risk factors for endometrial cancer among textile workers in Shanghai, China. *Am J Ind Med* 2008 Sep; 51(9): 673-9
25. Lubin JH. Extensions of analytic methods for nested and population-based incident case-control studies. *J Chron Dis* 1986; 39(5): 379-88

26. White E, Hunt JR, Casso D. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiol Rev* 1998; 20(1): 43-56
27. Wacholder S. Practical considerations in choosing between the case-cohort and nested case-control designs. *Epidemiology* 1991 Mar;2(2): 155-8
28. Ernstein VL. Nested case-control studies. *Prev Med* 1994 Sep; 23(5): 587-90
29. Kapeu AS, Luostarinen T, Jellum E, Dillner J, Hakama M, Koskela P, Lenner P, Löve A, Mahlamaki E, Thoresen S, Tryggvadóttir L, Wadell G, Youngman L, Lehtinen M. Is amoking and independent risk factor for invasive cervical cancer? A nested case-control study within Nordic biobanks. *Am J Epidemiol* 2009 Feb 15; 169(4): 480-8
30. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol* 1991 Jan 15; 133(2): 144-53
31. Mittelman MA, Maclure M, Robins JM. Control sampling strategies for case-crossover studies: an assessment of relative efficiency. *Am J Epidemiol* 1995 Jul 1; 142(1): 91-8
32. Meyer KA, Williams P, Hernandez-Diaz S, Cnattingius S. Smoking and risk of oral clefts: exploring the impact of study designs. *Epidemiology* 2004 Nov; 15(6): 671-8
33. Hutchinson KB, Kip KE, Ness RB. Condom use and its association with bacterial vaginosis and bacterial vaginosis-associated vaginal microflora. *Epidemiology* 2007 Nov; 18(6): 702-8
34. Pang D. A relative power table for nested matched case-control studies. *Occup Environ Med* 1999 Jan; 56(1): 67 – 69
35. Fairwater DV. Nausea and vomiting in pregnancy. *Am J Obstet Gynecol* 1968 Sep 1; 102(1): 135-75
36. Goodwin TM. Hyperemesis gravidarum. *Clin Obstet Gynecol* 1998 Sep; 41(3): 597-605
37. Ismail SK, Kenny L. Review on hyperemesis gravidarum. *Best Pract Res Clin Gastroenterol* 2007; 21(5): 755-69
38. Vikanes A, Grijbovski AM, Vangen S, Magnus P. Variations in prevalence of hyperemesis gravidarum by country of birth: a study of 900,074 pregnancies in Norway, 1967 – 2005. *Scand J Publ Health* 2008 Mar; 36(2): 135-42
39. Bailit JL. Hyperemesis gravidarum: Epidemiologic findings from a large cohort. *Am J Obstet Gynecol* 2005 Sep; 193(3 Pt 1): 811-4

40. Matsuo K, Ushioda N, Nagamatsu M, Kimura T. Hyperemesis gravidarum in Eastern Asian Population. *Gynecol Obstet Invest* 2007; 64(4): 213-6
41. Furneaux EC, Langley-Evans AJ, Langley-Evans SC. Nausea and vomiting of pregnancy: endocrine basis and contribution to pregnancy outcome. *Obstet Gynecol Surv* 2001 Dec; 56(12): 775-82
42. Hershman JM. Physiological and pathological aspects of the effect of human chorionic gonadotropin on the thyroid. *Best Pract Res Clin Endocrinol Metab* 2004 Jun; 18(2): 249-65
43. Tan JY, Loh KC, Yeo GS, Chee YC. Transient hyperthyroidism of hyperemesis gravidarum. *BJOG* 2002 Jun; 109(6): 683-8
44. Goodwin TM, Montoro M, Mestman JH, Pekary AE, Hershman JM. The role of chorionic gonadotropin in transient hyperthyroidism of hyperemesis gravidarum. *J Clin Endocrinol Metab* 1992 Nov; 75(5): 1333-7
45. Basso O, Olsen J. Sex ratio and twinning in women with hyperemesis or pre-eclampsia. *Epidemiology* 2001 Nov; 12(6): 747-9
46. Asklng J, Erlandsson G, Kaijser M, Akre O, Ekblom A. Sickness in pregnancy and sex of child. *Lancet* 1999 Dec 11; 354(9195): 2053
47. del Mar Melero-Montes M, Jick H. Hyperemesis gravidarum and the sex of the offspring. *Epidemiology* 2000 Jan; 12(1): 123-4
48. Schiff MA, Reed SD, Daling JR. The sex ratio of pregnancies complicated by hospitalization for hyperemesis gravidarum. *BJOG* 2004 Jan; 111(1): 27-30
49. Rodien P, Brémont C, Sanson ML, Parma J, Van Sande J, Costagliola S, Luton JP, Vassart G, Duprez L. Familial gastrointestinal hyperthyroidism caused by a mutant thyrotropin receptor hypersensitive to human chorionic gonadotropin. *N Engl J Med* 1998 Dec 17; 339(25): 1823-6
50. Fejzo MS, Ingles SA, Wilson M, Wang W, MacGibbon K, Romero R, Goodwin TM. High prevalence of severe nausea and vomiting of pregnancy and hyperemesis gravidarum among relatives of affected individuals. *Eur J Obstet Gynecol Reprod Biol* 2008 Nov; 141(1): 13-7
51. Corey LA, Berg K, Solaas MH, Nance WE. The epidemiology of pregnancy complications and outcome in a Norwegian twin population. *Obstet Gynecol* 1992 Dec; 80(6): 989-94
52. Trogstad LI, Stoltenberg C, Magnus P, Skjærven R, Irgens LM. Recurrence risk in hyperemesis gravidarum. *BJOG* 2005 Dec; 112(12): 1641-5

53. Grijbovski AM, Vikanes A, Stoltenberg C, Magnus P. Consanguinity and the risk of hyperemesis gravidarum in Norway. *Acta Obstet Gynecol Scan* 2008; 87(1): 20-5
54. Goldberg D, Szilagyi A, Graves L. Hyperemesis gravidarum and *Helicobacter pylori* infection: a systematic review. *Obstet Gynecol* 2007 Sep; 110(3): 695-703
55. Simpson SW, Goodwin TM, Robins SB, Rizzo AA, Howes RA, Buckwalter DK, Buckwalter JG. Psychological factors and hyperemesis gravidarum. *J Womens Health Gend Based Med* 2001 Jun; 10(5): 471-7
56. Swallow BL, Lindow SW, Masson EA, Hay DM. Psychological health in early pregnancy: relationship with nausea and vomiting. *J Obstet Gynaecol* 2004 Jan; 24(1):28-32
57. Munch S. Chicken or the egg? The biological-psychological controversy surrounding hyperemesis gravidarum. *Soc Sci Med* 2002 Oct; 55(7): 1267-78
58. Seng JS, Schrot JA, van De Ven C, Liberzon I. Service use data analysis of pre-pregnancy psychiatric and somatic diagnoses in women with hyperemesis gravidarum. *J Psychosom Obstet Gynaecol* 2007 Dec; 28 (4): 209-17
59. Fell DB, Dodds L, Joseph KS, Allen VM, Butler B. Risk factors for hyperemesis gravidarum requiring hospital admission during pregnancy. *Obstet Gynecol* 2006 Feb; 107(2 Pt 1): 277-84
60. Källén B, Lundberg G, Aberg A. Relationship between vitamine use, smoking, and nausea and vomiting of pregnancy. *Acta Obstet Gynecol Scan* 2003 Oct; 82(10): 916-20
61. Vilming B, Nesheim BI. Hyperemesis gravidarum in a contemporary population in Oslo. *Acta Obstet Gynecol Scand* 2000 Aug; 79(8): 640-3
62. Blackburn ST. *Maternal, Fetal & Neonatal Physiology, A Clinical Perspective*. Third edition. USA: Saunders, an imprint of Elsevier Inc.; 2007
63. De Swiet M. *Medical Disorders in Obstetric Practice*. Fourth edition. USA: Blackwell Science Inc.; 2002
64. 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 Dec; 76(6): 451-5

65. Poursharif B, Korst LM, Fejzo MS, MacGibbon KW, Romero R, Goodwin TM. The psychosocial burden of hyperemesis gravidarum. *J Perinatol* 2008 Mar; 28(3): 176-81
66. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research. Principles and Quantitative Methods*. USA: Van Nostrand Reinhold Company Inc; 1982
67. Mantel N. Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *J Am Stat Assoc* 1963 Sep; 58(303): 690-700
68. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986 Sep; 7(3): 177-88
69. Cochran WG. The Combination of Estimates from Different Experiments. *Biometrics* 1954 Mar; 10(1): 101-29
70. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analysis. *BMJ* 2003 Sep 6; 327(7414): 557-60
71. Thomson SG, Higgins JPT. How should meta-regression analysis be undertaken and interpreted? *Stat Med* 2002 Jun 15; 21(11): 1559-73
72. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for meta-analysis. *Stat Med* 1995 Feb 28; 14(4): 395-411
73. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* 1999 Oct 30; 18(20): 2693-708
74. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001 Jul 14; 323(7304): 101-5
75. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997 Sep 13; 315(7109): 629-34
76. Orwin RG. A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics* 1983; 8(2): 157- 9
77. Gioacchino L. *Meta-analysis in Medical Research. The handbook for the understanding and practice of meta-analysis*. BMJ books. Oxford: Blackwell Publishing Ltd.; 2005
78. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007 Oct 20; 370(9596): 1453-7

79. Lichtenstein MJ, Mulrow CD, Elwood PC. Guidelines for reading case-control studies. *J Chronic Dis* 1987; 40(9): 893-903
80. Elwood M. Critical appraisal of epidemiological studies and clinical trials. Second edition. USA: Oxford University Press; 2003
81. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007 Jun; 36(3): 666-76
82. Kupper LL, Karon JM, Kleinbaum DG, Morgenstern H, Lewis DK. Matching in epidemiologic Studies: validity and efficiency considerations. *Biometrics* 1981 Jun; 37(2): 271-91
83. Austin MA, Criqui MH, Barret-Connor E, Holdbrook MJ. The effect of response bias on the odds ratio. *Am J Epidemiol* 1981 Jul; 114 (1): 137-43
84. Richardi L, Boffetta P, Merletti F. Analysis of nonresponse bias in a population-based case-control study on lung cancer. *J Clin Epidemiol* 2002 Oct; 55(10): 1033-40
85. Rosner B, Hennekens CH. Analytic methods in matched pair epidemiological studies. *Int J Epidemiol* 1978 Dec; 7(4): 367-72
86. Ahlbom A, Alfredsson L. Interaction: A word with two meanings creates confusion. *Eur J Epidemiol* 2005; 20(7): 563-4
87. Marshall SW. Power for tests of interaction: effect of raising the Type I error rate. *Epidemiol Perspect Innov* 2007 Jun; 4: 4
88. Källberg H, Ahlbom A, Alfredson L. Calculating measures of biological interaction using R. *Eur J Epidemiol* 2006; 21(8): 571-3
89. Kalilani L, Astashili J. Measuring additive interaction using odds ratios. *Epidemiol Perspect Innov* 2006 Apr 18; 3: 5
90. Campbell UB, Gatto NM, Schwartz S. Distributional interaction: Interpretational problems when using incidence odds ratios to assess interaction. *Epidemiol Perspect Innov* 2005 Mar 3; 2(1): 1
91. Hogan MD, Kupper LL, Most BM, Haseman JK. Alternatives to Rothman's approach for assessing synergism (or antagonism) in cohort studies. *Am J Epidemiol* 1978 Jul; 108(1): 60-7
92. Frigo P, Lang C, Reisenberger K, Kölbl H, Hirschl AM. Hyperemesis gravidarum associated with *Helicobacter pylori* seropositivity. *Obstet Gynecol* 1998 Apr; 91(4): 615-7

93. Karaca C, Güler N, Yazar A, Camlica H, Demir K, Yildirim G. Is lower socio-economic status a risk factor for Helicobacter pylori infection in pregnant women with hyperemesis gravidarum? Turk J Gastroenterol 2004 Jun;15(2): 86-9
94. Jimenez V, Marleau JD. Is hyperemesis gravidarum related to country of origin? Short report. Can Fam Physician 2000 Aug; 46: 1607-8
95. Segal I, Ally R, Mitchell H. Helicobacter pylori – an African perspective. QJM 2001 Oct; 94(10): 561-5
96. Naja F, Kreiger N, Sullivan T. Helicobacter pylori infection in Ontario: prevalence and risk factors. Can J Gastroenterol 2007 Aug; 21(8): 501-6
97. Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnac P, Falush D, Stamer C, Prugnolle F, Van der Merwe SW, Yamaoka Y, Graham DY, Perez-Trallero E, Wadstrom T, Sauerbaum S, Achtman M. An African origin for the intimate association between humans and Helicobacter pylori. Nature 2007 Feb 22; 445(7130): 915-8
98. Higgins J, Thomson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy. 2002 Jan; 7(1): 51-61
99. Choi BC, Noseworthy AL. Classification, direction, and prevention of bias in epidemiologic research. J Occup Med 1992 Mar; 34(3): 265-71
100. Rupnow MF, Shachter RD, Owens DK, Parsonnet J. A dynamic transmission model for predicting trends in Helicobacter pylori and associated diseases in the United States. Emerg Infect Dis 2000 May-Jun; 6(3): 228-37
101. Ricci C, Holton J, Vaira D. Diagnosis of Helicobacter pylori: invasive and non-invasive tests. Best Pract Res Clin Gastroenterol 2007; 21(2): 299-313
102. Malaty HM. Epidemiology of Helicobacter pylori infection. Best Pract Res Clin Gastroenterol 2007; 21(2): 205-14
103. Olson SH, Voigt LF, Begg CB, Weiss NS. Reporting participation in case-control studies. Epidemiology 2002 Mar; 13(2): 123-6
104. Criqui MH. Response bias and risk ratios in epidemiologic studies. Am J Epidemiol 1979 Apr; 109(4): 394-9
105. Suerbaum S, Michetti P. Helicobacter pylori infection. N Engl J Med 2002 Oct 10; 347(15): 1175-86



106. Atherton JC. The clinical relevance of strain types of *Helicobacter pylori*. *Gut* 1997 Jun; 40(6): 701-3
107. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965 May; 58: 295-300
108. Robinson K, Argent RH, Atherton JC. The inflammatory and immune response to *Helicobacter pylori* infection. *Best Pract Res Clin Gastroenterol* 2007; 21(2): 237-59
109. Rothman KJ. *Epidemiology. An Introduction*. USA: Oxford University Press; 2002
110. Verberg MF, Gillott DJ, Al-Fardan N, Grudzinskas JG. Hyperemesis gravidarum, a literature review. *Hum Reprod Update* 2005 Sep-Oct; 11(5): 527-39
111. Bagis T, Gumurdulu Y, Kayaselcuk F, Yilmaz ES, Kilicadag E, Tarim E. Endoscopy in hyperemesis gravidarum and *Helicobacter pylori* infection. *Int J Gynaecol Obstet* 2002 Nov; 79(2): 105-9
112. Kosunen TU, Seppälä K, Sarna S, Sipponen P. Diagnostic value of decreasing IgG, IgA and IgM antibody titres after eradication of *Helicobacter pylori*. *Lancet* 1992 Apr 11; 339(8798): 893-5
113. Lerang F, Haug JB, Moum B, Mowinckel P, Berge T, Ragnhildstveit E, Bjørnklett A. Accuracy of IgG serology and other tests in confirming *Helicobacter* eradication. *Scand J Gastroenterol* 1998; 33: 710-5
114. Zimmermann EM, Christman GM. Approach to gastrointestinal and liver diseases in the female patient. In: Yamada T. *Textbook of Gastroenterology*. 4<sup>th</sup> edn. USA: Lippincott Williams & Wilkins; 2003
115. Manyonda IT. The immune system. In: Chamberlain G, Pipkin FB. *Clinical Physiology in Obstetrics*. 3<sup>rd</sup> edn. Oxford: Blackwell Science Ltd.; 1998
116. Stagno S, Reynolds DW, Huang ES, Thames SD, Smith RJ, Alford CA. Congenital cytomegalovirus infection. *N Engl J Med* 1977 Jun 2; 296(22): 1254-8
117. Xia LB, Yang J, Li AB, Tang SH, Xie QZ, Cheng D. Relationship between hyperemesis gravidarum and *Helicobacter pylori* seropositivity. *Chin Med J (Engl)*. 2004 Feb; 117(2): 301-2
118. Strachan BK, Jokhi RP, Filshie GM. Persistent hyperemesis gravidarum and *Helicobacter pylori*. *J Obstet Gynaecol* 2000 Jul; 20(4): 427

119. Jacoby EB, Porter KB. Helicobacter pylori infection and persistent hyperemesis gravidarum. *Am J Perinatol* 1999; 16(2): 85-8
120. El Younis CM, Abulafia O, Sherer DM. Rapid marked response of severe hyperemesis gravidarum to oral erythromycin. *Am J Perinatol* 1998; 15(9): 533-4
121. Mansour GM, Nashaat EH. Helicobacter pylori and hyperemesis gravidarum. *Int J Gynaecol Obstet* 2009 Jul; 106(1): 63-4
122. Goodwin TM, Montoro M, Mestman JH. Transient hyperthyroidism and hyperemesis gravidarum: clinical aspects. *Am J Obstet Gynecol* 1992 Sep; 167(3): 648-52
123. Bruun T, Kristoffersen K. Thyroid function during pregnancy with special reference to hydatiform mole and hyperemesis. *Acta Endocrinol (Copenh)* 1978 Jun; 88(2): 383-9
124. Kauppila A, Huhtaniemi I, Ylikorkala O. Raised serum human chorionic gonadotrophin concentrations in hyperemesis gravidarum. *Br Med J* 1979 Jun 23; 1(6179): 1670-1
125. Swaminathan R, Chin RK, Lao TT, Mak YT, Panesar NS, Cockram CS. Thyroid function in hyperemesis gravidarum. *Acta Endocrinol (Copenh)* 1989 Feb; 120(2): 155-60
126. Tareen AK, Baseer A, Jaffry HF, Shafiq M. Thyroid hormone in hyperemesis gravidarum. *J Obstet Gynaecol (Tokyo)* 1995 Oct; 21(5): 497-501
127. Leylek OA, Cetin A, Toyaksi M, Erselcan T. Hyperthyroidism in hyperemesis gravidarum. *Int J Gynaecol Obstet* 1996 Oct; 55(1): 33-7
128. Leylek OA, Toyaksi M, Erselcan T, Dokmetas S. Immunologic and biochemical factors in hyperemesis gravidarum with or without hyperthyroxinemia. *Gynecol Obstet Invest* 1999; 47(4): 229-34
129. Jordan V, Grebe SK, Cooke RR, Ford HC, Larsen PD, Stone PR, Salmond CE. Acidic isoforms of chorionic gonadotropin in European and Samoan women are associated with hyperemesis gravidarum and may be thyrotropic. *Clin Endocrinol (Oxf)* 1999 May; 50(5): 619-27
130. Al-Yatama M, Diejomaoh M, Nandakumaran M, Monem RA, Omu AE, Al Kandari F. Hormone profile of Kuwaiti women with hyperemesis gravidarum. *Arch Gynecol Obstet* 2002 Aug; 266(4): 218-22

131. Juras N, Banovac K, Sekso M. Increased serum reverse triiodothyronine in patients with hyperemesis gravidarum. *Acta Endocrinol (Copenh)* 1983 Feb; 102 (2): 284-7
132. Wilson R, McKillop JH, MacLean M, Walker JJ, Fraser WD, Gray C, Dryburgh F, Thomson JA. Thyroid function tests are rarely abnormal in patients with severe hyperemesis gravidarum. *Clin Endocrinol (Oxf)* 1992 Oct; 37 (4): 331-4
133. Panesar NS, Li CY, Rogers MS. Are thyroid hormones or hCG responsible for hyperemesis gravidarum? A matched paired study in pregnant Chinese women. *Acta Obstet Gynecol Scand* 2001 Jun; 80(6): 519-24
134. 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 May; 156 (5): 1137-41
135. Unsel N, Benian A, Erel CT. Leptin levels in women with hyperemesis gravidarum. *Int J Gynaecol Obstet* 2004 Feb; 84(2): 162-3
136. Aka N, Atalay S, Sayharman S, Kilic D, Köse G, Küçüközkan T. Leptin and leptin receptor levels in pregnant women with hyperemesis gravidarum. *Aust NZJ Obstet Gynaecol* 2006 Aug; 46(4): 274-7
137. Demir B, Erel CT, Haberal A, Oztürk N, Güler D, Kocak M. Adjusted leptin level (ALL) is a predictor for hyperemesis gravidarum. *Eur J Obstet Gynecol Reprod Biol* 2006 Apr 1; 124(2): 193-6
138. Ustün Y, Engine-Ustün Y, Dökmeci F, Söylemez. Serum concentrations of lipids and apolipoproteins in normal and hyperemetic pregnancies. *J Matern Fetal Neonatal Med* 2004 May; 15(5): 287-90
139. Signorello LB, Harlow BL, Wang S, Erick MA. Saturated fat intake and risk of severe hyperemesis gravidarum. *Epidemiology* 1998 Nov; 9(6): 636-40
140. Rochelson B, Vohra N, Darvishzadeh J, Pagano M. Low pregnancy ideal weight : height ratio in women with hyperemesis gravidarum. *J Reprod Med* 2003 Jun; 48(6): 422-4
141. Khory MJ, Liittle J, Burke W. *Human Genome Epidemiology: A Scientific Foundation for using Genetic Information to Improve Health and Prevent Disease*. USA: Oxford University Press; 2004

142. Khoury MJ, Beaty TH. Applications of case-control method in genetic epidemiology. *Epidemiol Rev* 1994; 16 (1): 134-50
143. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Am J Epidemiol* 1996 Aug 1; 144(3): 207-13
144. Foppa I, Spiegelman D. Power and sample size calculations for case-control studies of gene-environment interactions with a polytomous exposure variable. *Am J Epidemiol* 1997 Oct !; 146(7): 596-604
145. Khoury MJ, Beaty TH, Hwang SJ. Detection of genotype-environment interaction in case-control studies of birth defects: how big a sample size? *Teratology* 1995 May; 51(5): 336-43
146. Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE, McCormack VA. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004 Oct 16; 329(7471): 883
147. Lee W, Bindman J, Ford T, Glozier N, Moran P, Stewart R, Hotopf M. Bias in psychiatric case-control studies. *Br J Psychiatry* 2007 Mar; 190: 204-9
148. Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 1999 May 26; 281(20): 1919-26
149. Skovlund E. A critical review of papers from clinical cancer research. *Acta Oncol* 1998; 37(4): 339-45
150. Smidt N, Rutjes AW, van der Windt DA, Ostelo RW, Reitsma JB, Bossuyt PM, Bouter LM, de Vet HC. Quality of reporting of diagnostic accuracy studies. *Radiology* 2005 May; 235 (2): 347-53
151. Rushton L. Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods. *Occup Environ Med* 2000 Jan; 57(1): 1-9
152. Hellems MA, Kramer MS, Hayden GF. Case-Control Confusion. *Ambul Pediatr* 2006 Mar-Apr; 6(2): 96-9
153. Mihailovic A, Bell CM, Urbach DR. Users' guide to the surgical literature. Case-control studies in surgical journals. *Can J Surg* 2005 Apr; 48(2): 148-51
154. Dauphinee L, Peipert JF, Phipps M, Weitzen S. Research methodology and analytic techniques used in the *Journal Obstetrics & Gynecology*. *Obstet Gynecol* 2005 Oct; 106(4): 808-12

155. Huwiler-Müntener K, Jüni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA* 2002 Jun 5; 287(21): 2801-4
156. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics* 2001 Dec; 2(4): 463-71
157. Jüni P, Witchi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999 Sep 15; 282(11): 1054-60
158. Fung ICH. Seek, and ye shall find: accessing the global epidemiological literature in different languages. *Emerg Themes Epidemiol* 2008 Sep 30; 5: 21















