Systematic review with meta-analysis
Li, Cheng Mei; Zhernakova, Alexandra; Engstrand, Lars; Wijmenga, Cisca; Brusselaers, Nele

Published in:
Alimentary Pharmacology & Therapeutics

DOI:
10.1111/apt.15610

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 22-09-2023
Systematic review with meta-analysis: the risks of proton pump inhibitors during pregnancy

Cheng Mei Li$^{1,2,3}$ | Alexandra Zhernakova$^4$ | Lars Engstrand$^{1,2}$ | Cisca Wijmenga$^4$ | Nele Brusselaers$^{1,2}$

1Centre for Translational Microbiome Research, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden
2Science for Life Laboratory, Stockholm, Sweden
3Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China
4Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Correspondence
Nele Brusselaers, Centre for Translational Microbiome Research, Department of Microbiology, Tumor and Cell Biology, Nobelsväg 16, 17177 Stockholm, Sweden. Email: nele.brusselaers@ki.se

Summary
Background: There have been safety concerns considering long-term proton pump inhibitor (PPI) use, also during pregnancy.

Aims: To assess the risk of adverse neonatal outcomes associated with maternal intake of PPIs by means of systematic review and meta-analysis.

Methods: The systematic search included PubMed, Web of Science, Cochrane Database and Embase (inception until June 2019). All studies reporting ≥1 adverse pregnancy outcome comparing PPI users to non-users. Histamine-2 receptor antagonists (H2RA) were also compared to both non-users and PPI users. Outcomes included congenital malformations, abortion, stillbirth, neonatal death, preterm birth, small for gestational age and low birth weight. Pooled odds ratios (OR) and 95% confidence intervals (CI) were obtained by random-effects modelling. PROSPERO study-protocol: CRD42018103320.

Results: In total, 26 observational studies (20 cohort, 6 case-control studies) were identified, of which 19 assessed PPIs and 12 H2RA. PPI use was associated with an increased risk of congenital malformations (OR 1.28, 95% CI 1.09-1.52), especially in case-control studies (OR 2.04, 1.46-2.86). No associations were found between H2RA and congenital malformations. No significant associations were found between PPI use and abortions, stillbirth, neonatal death, preterm birth, small for gestational age and low birth weight. Pooled odds ratios (OR) and 95% confidence intervals (CI) were obtained by random-effects modelling. PROSPERO study-protocol: CRD42018103320.

Conclusions: This meta-analysis suggests an association between maternal PPI use and congenital malformations in humans, yet power was insufficient to assess specific malformations and drugs.

As part of AP&T's peer-review process, a technical check of this meta-analysis was performed by Dr Y. Yuan.

The Handling Editor for this article was Professor Colin Howden, and it was accepted for publication after full peer-review.
1 | INTRODUCTION

Proton pump inhibitors (PPI) are among the most commonly and increasingly used drugs worldwide, with up to one-third using prescribed PPI annually.1-4 PPIs are also accessible because of the over-the-counter availability in many countries. Yet, the safety of PPIs is more and more questioned in particular considering cancer, mortality, chronic kidney disease and other potential long-term effects.5-7 PPIs have also been described as the drug group with the largest effect on the gastro-intestinal microbiome, in terms of decreasing diversity, based on population-based studies, larger than antibiotics and any other drug.8-10 Safety during pregnancy has also been questioned repeatedly,11-13 including a recent meta-analysis showing a 45% increased risk of childhood asthma in offspring,14 and another study showing an increased risk of cholestasis.15 The two previous meta-analyses assessing pregnancy outcomes after PPI use, published more than 10 years ago, showed an increased risk of congenital malformations in women using PPI during pregnancy, yet the results did not reach statistical significance (RR = 1.18, 95% CI 0.72-1.94, 5 studies11; RR = 1.12, 95% CI 0.86-1.45, 7 studies12). Two meta-analyses assessing H2-receptor antagonists (H2RA), the most popular alternative for PPIs, showed no associations with congenital malformations (pooled OR = 1.14, 95% CI 0.89-1.45, 4 studies),15 and a significant increased risk of asthma (pooled RR = 1.57; 95% CI 1.46-1.69, 5 studies).14

Omeprazole, the most commonly used PPI was initially categorised as category C by the Food and Drug Administration, implying some teratogenic effects have been described in animal studies, and safety is not absolutely guaranteed for human use during pregnancy.17,18 Other PPIs were categorised as category B, which also warrants some prudence in use during pregnancy.17,18 Current recommendations (not using these categories anymore) for Omeprazole are that “this drug should be used during pregnancy only if the benefit outweighs the risk for the foetus”, and that it is not recommended during breastfeeding.19 Therefore, it may not be surprising that we have not identified any randomised clinical trials on PPI use during pregnancy assessing neonatal adverse events.20,21 Yet, although not recommended for nausea and vomiting during pregnancy nor hyperemesis gravidarum,21 a small studies reported effective treatment of reflux.22 Global statistics on how many women use PPIs during pregnancy are scarce, but it’s use is estimated to be below 1.5% in Sweden.23 Yet, their use may also be underestimated if based on prescription-use,24 and potentially increasing in the many countries with over-the-counter availability. Reported PPIs are also used for gastro-oesophageal reflux and heartburn during pregnancy, usually arising during the 2nd and 3rd trimester due to hormonal changes and mechanistic pressure.25,26 Although outside the critical embryonic period for congenital malformations, drug exposure during later pregnancy may affect other less-studied outcomes such as preterm birth and low-birth weight, and childhood asthma (although timing of exposure could not be assessed in this meta-analysis).14

Since several new studies have been published on PPI use during pregnancy during the last decade, the aim of this systematic review and meta-analysis is to assess the current evidence of PPI use (and H2RA as most popular alternative treatment) during pregnancy and the risk of adverse events, including congenital malformations but also preterm birth and other neonatal adverse events.

2 | MATERIAL AND METHODS

2.1 | Eligibility criteria, information sources, search strategy and study selection

We conducted a comprehensive systematic literature search in PubMed, Web of Science, EMBASE databases and Cochrane Database from inception, which was last updated June 2019 (Supplement 1), and we reported the results according to the PRISMA guidelines for systematic reviews and meta-analyses. The study protocol was registered at PROSPERO (CRD42018103320). The search was conducted independently by two reviewers (CML, NB) and any disagreements were solved by discussion. Backward and forward citation tracking on the identified eligible articles was used to identify articles potentially missed by the search. If no full-text version of eligible conference abstracts was found, we contacted the authors to complement the available data. All study designs (observational and intervention studies) were eligible if the association was assessed between PPI (or H2RA) exposure in pregnancy and the risk of neonatal adverse events compared to a comparison group, providing crude numbers or ratios. Exposure was defined as any type of PPIs (or H2RA), during pregnancy. Although teratogenic effects affecting the risk for congenital malformations only occur because of exposure before or during the first trimester, the critical time-window for exposure is not clear for the other outcomes, and may not always be clearly described. The comparison group included: (a) all or a selection of pregnant women not using PPIs (or H2RA) during pregnancy from the same cohort with or without the same indication, (b) the total population of pregnant women from the same study setting. Since H2RA are used for similar indications, we also assessed the risk of PPI use compared to H2RA use during pregnancy. The studies were eligible if at least one of the following neonatal adverse events were assessed (as defined by the authors): congenital malformations, preterm birth, low birth weight, small for gestational age, spontaneous abortion (miscarriage) and termination of pregnancy (for medical reasons or other), perinatal mortality and stillbirth. Information on mean/median gestational age and weight were collected from the exposed and comparison groups, if available. We excluded case studies, case series, cross-sectional studies, animal studies and articles in other languages than English. As H2RA were commercialised since the 1970 (PPIs since the 1980s) no eligible articles are expected to be published before the 1970, although the databases were searched from inception. When we identified two studies based on the same cohort, describing the same exposures and outcomes, only the largest and/or more recent one was included in these analyses.
2.2 | Data extraction

Data extraction included author, publication year, country, study type, population source and characteristics, control group, exposure (timing, type, duration), outcome definitions and categorisations.

2.3 | Assessment of risk of bias

Two review authors (CML, NB) independently assessed the risk of bias for each included study by means of the Newcastle Ottawa quality assessment scales (NOS) for cohort and case-control studies.27

Each study can be awarded up to nine stars, with a higher amount of stars reflecting a lower risk of bias, and we considered seven stars or more as a low risk of bias. In addition, we assessed in more depth the risk of selection bias (representativeness of the cohorts, cases and controls), information bias (ie recall bias of exposure, attrition bias for outcome), and confounding (including confounding by indication), and potential conflicts of interest.28

2.4 | Data synthesis

Data were analysed using Stata (metan package; StataCorp version 14). Random-effects modelling was used to pool the OR with 95% confidence intervals (CIs), based on the adjusted estimates. Since all outcomes are rare (<10%), risk ratios and odds ratios should be equivalent; yet if odds ratios were not reported, these were calculated based on the available crude numbers (so no measurements needed to be transformed).29 A “treatment arm” continuity correction accounting for unbalanced groups was used to include studies with zero events in one arm, which takes into account the ratio of the group sizes,30 yet a sensitivity analyses excluding these trials was presented where appropriate. Studies without events in both arms were excluded (since not always clear which outcomes were investigated in each study). Gestational weight and age could not be pooled, since rarely reported for both the exposed and comparison groups. Results on each outcome were only presented if reported in two or more studies. The I-square test shows the proportion of heterogeneity between studies, and was categorised as low (25%-50%), moderate (51%-75%), or high (>75%) statistical heterogeneity.31 Subgroup analyses were used to test the sources of heterogeneity, and were based on PPI type, study design and different trimester if at least two studies could be included. If different estimates could be used, adjusted results were preferred above unadjusted, and early pregnancy exposure (1st trimester) above late (2nd/3rd exposure) for congenital malformations, and late pregnancy for the other outcomes, unless otherwise specified.

From a developmental perspective, it is not expected that any exposure would increase the risk of all congenital malformations, since in reality a teratogen only affects a limited number of malformations. Yet the overall risk is what is generally reported in many studies, which may show associations if the effect sizes for the specific malformations are sufficiently high. In order to include as many studies as possible in the analyses for congenital malformations (to increase power), studies only reporting on specific malformations were also included unless otherwise specified. If these different malformations could not be combined into one risk estimate, the estimate of the outcome with the highest number of exposed cases was included. Since it is unexpected that teratogens increase the risk of all congenital malformations, specific malformations were assessed separately if feasible. Influence meta-analyses were conducted to investigate the influence of each individual study on the overall meta-analysis summary estimate. For the cumulative meta-analysis, all studies are sorted by the year of publication, and each line represents the accumulated effect of this study and all studies published earlier.32 Symmetry of the funnel plots was assessed visually, and statistically by means of the Egger test, to assess small study effects (publication bias).33

3 | RESULTS

3.1 | Study selection

The systematic literature selection is presented in Supplement 2, and resulted in 26 eligible studies reporting on PPI and/or H2RA use during pregnancy and at least a single type of neonatal adverse events.34-60 The reasons for exclusion of articles excluded during full text evaluation are reported in Supplement 1. Two of the included studies were only identified through backward and forward citation tracking,58,52 and one study59 was only referred to as personal communication in two previously published meta-analyses of which one included the same author.11,12 Two included studies were only included as conference abstracts; we therefore contacted the authors (unsuccessfully) to complement the available data.39,55

3.2 | Study characteristics

Supplement 3 includes all extracted study and population characteristics. All 26 included studies (20 cohort studies, 6 case-control studies) were published between 1985 and 2019, of which more than half were published during the last decade. Of these 26 studies, 14 described the risk for neonatal adverse events among PPI users only,39-41,46,47,51-55,57-60 five on H2RA users only,38,42,44,49,50 and seven reported on both.34,36,37,43,45,48,56 Fourteen different countries contributed with data, with five studies from the United States, four from Denmark,41,52-54 three from Sweden,45,46,60 and Israel,40,50,51 and four international collaborations.40,44,47,56 Two Swedish registry-based studies had an overlap considering the covered time period, so the oldest study was only used to assess H2RA which was not covered in the more recent study.45,46 Two Hungarian studies were also based on the same cohort, yet
<table>
<thead>
<tr>
<th>Pooled odds ratio</th>
<th>95% confidence interval</th>
<th>I²</th>
<th>N included studies</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall, incl. continuity correction</td>
<td>1.28*</td>
<td>1.09-1.52</td>
<td>26.0</td>
<td>18</td>
</tr>
<tr>
<td>Overall</td>
<td>1.29*</td>
<td>1.09-1.53</td>
<td>29.4</td>
<td>17</td>
</tr>
<tr>
<td>Overall, cohort studies only</td>
<td>1.12</td>
<td>0.99-1.27</td>
<td>0.0</td>
<td>11</td>
</tr>
<tr>
<td>Overall, case-control studies only</td>
<td>2.04*</td>
<td>1.46-2.86</td>
<td>0.0</td>
<td>6</td>
</tr>
<tr>
<td>Overall, without 2 conference abstracts</td>
<td>1.18*</td>
<td>1.01-1.38</td>
<td>11.4</td>
<td>15</td>
</tr>
<tr>
<td>Overall, cohorts without conference abstract</td>
<td>1.12</td>
<td>0.98-1.27</td>
<td>0.0</td>
<td>10</td>
</tr>
<tr>
<td>Overall, case-control studies without conference abstract</td>
<td>1.18*</td>
<td>1.01-1.38</td>
<td>0.0</td>
<td>5</td>
</tr>
</tbody>
</table>

All malformations combined (excluding specific malformations)

| Overall, with continuity correction | 1.08 | 0.95-1.23 | 0.0 | 12 | 34,39,40,43,46,47,51,53,54,56,58,59 |
| Overall | 1.08 | 0.95-1.23 | 0.0 | 11 | 34,39,40,43,46,47,51,53,54,56,59 |
| Only cohort studies | 1.07 | 0.94-1.22 | 0.0 | 9 | 39,40,46,47,51,53,54,56,59 |
| Only case-control studies | 1.79 | 0.66-4.86 | 0.0 | 2 | 34,43 |
| Overall, first trimester | 1.07 | 0.94-1.22 | 0.0 | 8 | 39,40,46,47,51,53,54,56 |
| Overall, first trimester, omeprazole | 1.05 | 0.89-1.24 | 0.0 | 6 | 40,46,47,51,54,56 |
| Overall, first trimester, pantoprazole | 1.28 | 0.83-1.99 | 0.0 | 2 | 40,54 |
| Overall, first trimester, lantoprazole | 1.03 | 0.76-1.40 | 0.0 | 3 | 40,51,54 |
| Overall, second/third trimester | 1.10 | 0.93-2.30 | 0.0 | 2 | 53,54 |
| Overall, any time | 1.18 | 0.74-1.89 | 0.0 | 6 | 34,40,43,47,53,59 |

Specific malformations

| Cleft palate without cleft lip, PPI vs control (1 cohort, 1 case-control) | 1.68 | 0.47-6.05 | 53.9 | 2 | 36,54 |
| Hypospadias, PPI vs control (2 cohorts) | 1.46 | 0.94-2.28 | 34.2 | 2 | 41,48 |
| Cardiovascular malformations, PPI vs control (2 cohorts, 1 case-control) | 1.52 | 0.85-2.73 | 75.4 | 3 | 46,51,55 |
| Ventricular septal defect, PPI vs control (2 cohorts) | 1.07 | 0.74-1.53 | 0.0 | 2 | 51,54 |
| Atrial septal defect, PPI vs control (2 cohorts) | 1.27 | 0.89-1.81 | 0.0 | 2 | 51,54 |
| Clubfoot, PPI vs control (1 cohort, 1 case-control) | 1.29 | 0.74-2.25 | 0.0 | 2 | 54,57 |

H2RA vs Control

| Overall, incl. continuity correction | 1.05 | 0.83-1.34 | 20.5 | 11 | 34,36,38,42-45,47-50,56 |

(Continues)
one assessed the risk in individuals with peptic ulcers, and the other one with chronic dyspepsia, so since the participants did not overlap, both were included in the analyses. Two Danish studies and two American studies reported on hypospadias based on the same nationwide data collection and study period, so the oldest one was excluded from the specific analyses on hypospadias. The total number of pregnancies included in each study ranged between 42,38 and 85,248,352; and the number of congenital malformations between 238 and 22,843.43

### 3.2.1 Risk of bias of included studies

There were 20 cohort studies of which 16 (80.0%) had at least seven stars according to the NOS quality scale, and six case-control studies of which two (33.3%) had at least seven stars (Supplement 4). The risk of bias has been assessed in more detail in Supplement 3: five studies recruited women through a teratogen information service, that is, women who were exposed to potentials teratogens called in for advice, with controls also being selected among callers exposed to non-teratogens. Two studies were prescription-event monitoring studies, one recruited symptomatic individuals (with hyperemesis gravidarum), and 18 studies were based on national or regional data collection systems.

Exposure information (drug use) was collected during pregnancy in all but six studies (including five out of six case-control studies), and 14 studies define exposure as any use, while eight restricted to prescribed use only (unclear in four studies). Three studies looked exclusively at exposure during the first trimester, while four did not report the exposure window of interest. Indications of use were only reported in eight studies, of which five were restricted to one diagnosis (nausea and vomiting, hyperemesis gravidarum, peptic ulcers, chronic dyspepsia). Seven studies reported potential conflicts of interest or received funding from pharmaceutical companies.

### 3.3 Congenital malformations

In total, 23 of the included studies reported at least one type of congenital malformations, 18 assessing PPI use, and 11 H2RA use. Of these, one study reported no events in the exposed group, and one in the non-exposed group. Four studies only reported on specific congenital malformations, including hypospadias, club foot and hydrocephalus, and one additional study only reported on a selected number of congenital malformations, but did not report the overall numbers or risks.

The overall risk of congenital malformations was 30% higher among PPI users than in the control group (OR = 1.28, 95% CI 1.09-1.52, N = 18), yet no increased risk was found for H2RA users (OR = 1.05, 95% CI 0.83-1.34, N = 11). The risk was highest in the case-control studies (OR = 2.04, 95% CI 1.46-2.86, N = 6) and the confidence intervals were close to reaching statistical significance for the cohort studies (OR = 1.12, 95% CI 0.99-1.27, N = 12). Pooled effect sizes could be calculated for hypospadias, cleft palate without cleft lip, cardiovascular malformations, and clubfoot, all with increased yet not significant effect sizes.
The subgroup analyses on timing of exposure or type of PPI did not reach significance since most studies did not report the necessary stratified information or did not specify the exposure window clearly. Yet, none of these subgroup analyses resulted in decreased effect sizes and statistical heterogeneity was low. The six studies comparing the risk of any congenital malformation for PPIs directly with H2RA, provided a pooled odds ratio of 1.22 (95% CI 0.67-2.23).34,36,43,47,48,56

The cumulative meta-analysis including all studies reporting at least one congenital malformation indicates that if this meta-analysis would have been done any time after the large Danish study in 2010,54 a significantly increased risk of at least 20% would have been found (Figure 2). The influence analysis (Supplement 5) on the same group of studies shows that no single study tips the results into significance, since every single study can be removed from the analyses without losing significance. The funnel plot did not provide evidence for the presence of small study effects (Supplement 6).

### 3.4 | Spontaneous abortion and stimulated termination of pregnancy

The association between PPI47,58 or H2RA44,47,49 use and spontaneous abortion or termination of pregnancy were only reported in three studies each (Table 2). For the PPI studies, heterogeneity was low, with none of the associations reaching statistical significance. For H2RA, lower risk for spontaneous abortion was found (OR 0.66, 95% CI 0.43-1.00).

### 3.5 | Still birth and perinatal mortality

The risk of stillbirth was assessed for PPI users in five studies,40,53,56,58,60 and two studies for H2RA.44,56 Although the risk seems increased for both groups, the confidence intervals were broad and the results were not significant (Table 2). Two studies reported on PPI use and perinatal mortality,46,51 showing no evidence for an association.

### 3.6 | Low birth weight and small for gestational age

Five studies reported on low-birth weight or SGA for women exposed to PPIs.46,51,53,56,60 No analyses reached statistical significance (Table 2).46,51 Insufficient information was available on extreme low birth weight (only reported in one study50), or the average gestational weight for PPI/H2RA users and controls (mean and standard deviation only reported in two studies for PPIs47,60 and one for H2RA49).
3.7 | Preterm birth

Seven studies reported on the association between PPI use and preterm birth,\textsuperscript{40,47,51,53,56,58,60} and five on H2RA use.\textsuperscript{44,47,49,50,56} None of the analyses reached statistical significance, except for a potential increased risk for H2RA (1.25, 95% CI 1.02-1.56) (Table 2). Mean gestational length could not be assessed, since only two studies reported means and standard deviations for PPIs,\textsuperscript{47,60} and one for H2RAs.\textsuperscript{59} The funnel test did not provide evidence for the presence of small study effects (Egger test: \( P = 0.248 \)).

4 | DISCUSSION

This comprehensive meta-analyses based on 26 studies assessing the risk of different pregnancy outcomes among women exposed to PPI or H2RA during pregnancy is the first meta-analysis providing evidence that PPI use during pregnancy may be linked to congenital malformations in humans, with significant pooled results in case-control studies and borderline significant in cohort studies. These results should be without doubt interpreted with caution, since the data are based on published results (no individual patient data were available), originating from a relatively heterogeneous group of observational studies, yet statistical heterogeneity seemed limited. Our main analyses combined all studies reporting on at least one congenital malformation, yet different subgroups were assessed, by study design, type of malformation, type of drug and timing of exposure. Although power was low in these analyses, the results seemed consistent in the direction of effect.

For any teratogenic exposure, it is unexpected that the risk of all congenital malformations would be increased (complicating detection\textsuperscript{61}), yet the low risk of each individual malformation combined with the low prevalence of exposure (<1% in many of the included studies) makes it even more challenging to detect any potential association. Therefore, potential associations with congenital malformations based on case-control studies are considered closer to reality than cohort studies.\textsuperscript{61}

Although quality was assessed to be good for most studies, most studies do have some methodologic limitations, in particular relating to selection bias, since only 10 studies were population-based covering
Therefore, some selection process occurred when selecting exposed and non-exposed women in the cohort studies, or cases but especially controls in case-control studies. An important limitation for assessing the risk of congenital malformations, is that timing of exposure was not always clearly mentioned, although generally known to be only of importance if exposure occurred during the critical organ-forming period (1st trimester). Therefore, even if exposure was not clearly defined, it could be assumed that the majority of the exposure occurred during the first trimester since most drug-safety studies during pregnancy focus on congenital malformations. For assessing the other outcomes, the critical exposure window has not yet been established. Recall bias seemed to be of importance in the case-control studies, were information on drug exposure was collected in five out of six studies after the outcome had occurred (potentially contributing to higher pooled risk estimates for case-control studies compared to cohort studies).

### TABLE 2  Risk of preterm birth and other neonatal adverse events

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Pooled odds ratio</th>
<th>95% CI</th>
<th>I²</th>
<th>Number of studies</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous abortion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control, with continuity correction</strong></td>
<td>1.27</td>
<td>0.85-1.90</td>
<td>0.0</td>
<td>3</td>
<td>40,47,58</td>
</tr>
<tr>
<td><strong>PPI vs control</strong></td>
<td>1.29</td>
<td>0.84-1.98</td>
<td>6.6</td>
<td>2</td>
<td>40,47</td>
</tr>
<tr>
<td><strong>PPI vs control, with continuity correction</strong></td>
<td>1.27</td>
<td>0.85-1.90</td>
<td>0.0</td>
<td>3</td>
<td>40,47,58</td>
</tr>
<tr>
<td><strong>PPI vs control, omeprazole only</strong></td>
<td>1.37</td>
<td>0.89-2.11</td>
<td>0.0</td>
<td>2</td>
<td>40,47</td>
</tr>
<tr>
<td><strong>H2RA vs control</strong></td>
<td>0.66*</td>
<td>0.43-1.00</td>
<td>5.7</td>
<td>3</td>
<td>44,47,49</td>
</tr>
<tr>
<td>Stimulated termination of pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control</strong></td>
<td>0.90</td>
<td>0.32-2.58</td>
<td>0.0</td>
<td>2</td>
<td>47,58</td>
</tr>
<tr>
<td><strong>H2RA vs control</strong></td>
<td>1.38</td>
<td>0.30-6.23</td>
<td>89.2</td>
<td>3</td>
<td>44,47,49</td>
</tr>
<tr>
<td>Stillbirth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control, with continuity correction</strong></td>
<td>2.57</td>
<td>0.77-8.65</td>
<td>63.6</td>
<td>5</td>
<td>40,53,56,58,60</td>
</tr>
<tr>
<td><strong>H2RA vs control</strong></td>
<td>1.76</td>
<td>0.67-4.64</td>
<td>26.9</td>
<td>2</td>
<td>44,56</td>
</tr>
<tr>
<td>Low birth weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control, any or preferably early</strong></td>
<td>0.84</td>
<td>0.64-1.09</td>
<td>0.0</td>
<td>3</td>
<td>46,51,53</td>
</tr>
<tr>
<td><strong>PPI vs control, any or preferably late</strong></td>
<td>0.79</td>
<td>0.57-1.11</td>
<td>0.0</td>
<td>3</td>
<td>46,51,53</td>
</tr>
<tr>
<td><strong>PPI vs control, first semester only</strong></td>
<td>1.29</td>
<td>0.59-2.82</td>
<td>55.6</td>
<td>2</td>
<td>46,53</td>
</tr>
<tr>
<td><strong>PPI vs control, 2nd/3rd semester only</strong></td>
<td>0.78</td>
<td>0.55-1.10</td>
<td>0.0</td>
<td>3</td>
<td>46,51,53</td>
</tr>
<tr>
<td>Small for gestational age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H2RA vs control</strong></td>
<td>0.27</td>
<td>0.06-1.16</td>
<td>48.9</td>
<td>2</td>
<td>49,56</td>
</tr>
<tr>
<td>Low birth weight and SGA combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control, incl. continuity correction, any or preferably early</strong></td>
<td>1.05</td>
<td>0.56-1.97</td>
<td>85.4</td>
<td>5</td>
<td>46,51,53,56,60</td>
</tr>
<tr>
<td><strong>PPI vs control, incl. continuity correction, any or preferably late</strong></td>
<td>1.08</td>
<td>0.53-2.24</td>
<td>83.8</td>
<td>4</td>
<td>46,51,53,56,60</td>
</tr>
<tr>
<td>Preterm birth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control (priority any time)</strong></td>
<td>1.04</td>
<td>0.86-1.25</td>
<td>25.9</td>
<td>7</td>
<td>40,47,51,53,56,58,60</td>
</tr>
<tr>
<td><strong>PPI vs control (priority late pregnancy)</strong></td>
<td>0.98</td>
<td>0.87-1.11</td>
<td>4.7</td>
<td>7</td>
<td>40,47,51,53,56,58,60</td>
</tr>
<tr>
<td><strong>PPI vs control, omeprazole only</strong></td>
<td>1.28</td>
<td>0.91-1.79</td>
<td>0.0</td>
<td>4</td>
<td>40,47,56,58</td>
</tr>
<tr>
<td><strong>PPI vs control, first semester only</strong></td>
<td>1.30</td>
<td>0.77-2.22</td>
<td>0.0</td>
<td>3</td>
<td>53,56,58</td>
</tr>
<tr>
<td><strong>PPI vs control, 2nd/3rd semester only</strong></td>
<td>1.27</td>
<td>0.49-3.30</td>
<td>73.5</td>
<td>2</td>
<td>51,53</td>
</tr>
<tr>
<td><strong>H2RA vs control (priority late pregnancy)</strong></td>
<td>1.22</td>
<td>0.92-1.64</td>
<td>47.1</td>
<td>5</td>
<td>44,47,49,50,56</td>
</tr>
<tr>
<td><strong>H2RA vs control (priority early pregnancy)</strong></td>
<td>1.25*</td>
<td>1.02-1.56</td>
<td>24.9</td>
<td>5</td>
<td>44,47,49,50,56</td>
</tr>
<tr>
<td><strong>H2RA vs control, first semester only</strong></td>
<td>1.09</td>
<td>0.89-1.33</td>
<td>0.0</td>
<td>2</td>
<td>50,56</td>
</tr>
<tr>
<td><strong>PPI vs H2RA</strong></td>
<td>0.82</td>
<td>0.45-1.50</td>
<td>16.1</td>
<td>2</td>
<td>47,56</td>
</tr>
<tr>
<td><strong>PPI + H2RA vs control</strong></td>
<td>1.20</td>
<td>0.94-1.55</td>
<td>0.0</td>
<td>3</td>
<td>37,47,56</td>
</tr>
<tr>
<td>Perinatal mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI vs control</strong></td>
<td>0.79</td>
<td>0.16-3.94</td>
<td>75.5</td>
<td>2</td>
<td>46,51</td>
</tr>
</tbody>
</table>

Note: Associations reaching statistical significance (as shown by the 95% confidence intervals) are marked with *. 

---

a well-described geographical region. Therefore, some selection process occurred when selecting exposed and non-exposed women in the cohort studies, or cases but especially controls in case-control studies. An important limitation for assessing the risk of congenital malformations, is that timing of exposure was not always clearly mentioned, although generally known to be only of importance if exposure occurred during the critical organ-forming period (1st trimester). Therefore, even if exposure was not clearly defined, it could be assumed that the majority of the exposure occurred during the first trimester since most drug-safety studies during pregnancy focus on congenital malformations. For assessing the other outcomes, the critical exposure window has not yet been established. Recall bias seemed to be of importance in the case-control studies, were information on drug exposure was collected in five out of six studies after the outcome had occurred (potentially contributing to higher pooled risk estimates for case-control studies compared to cohort studies).
and over-the-counter use may have been missed when based on pre-
scribed records (13 out of 25 studies were based on any use), while
compliance to prescriptions (actual use) is not reported. Insufficient in-
formation was provided to allow assessment of (accumulated) dosage
or indication of use. The effect of pre-gestational exposure could not be
assessed in the current study because only reported in sufficient
detail in one study, although five studies considered the 30 days
before pregnancy as gestational exposure. Congenital mal-
formations were without doubt the most commonly reported adverse
event. Yet, taking into account the rarity of congenital malformations
(approximately 2.5%) combined with the low prevalence of PPI use
during pregnancy (<1% in several of the included studies), many of the
studies were clearly underpowered to detect differences. Although
we restricted our inclusion criteria to English articles only, no eligible
articles were identified in other languages, so the effect of language
bias should be negligible. One additional yet important limitation of this
meta-analysis is that a relatively large proportion of the included stud-
ies (7/25) reported some conflict of interest through pharmaceutical
sponsoring or connections; which may contribute to the fact that none of
the single studies showed strong evidence for an association with
these outcomes before. We cannot rule out a negative publication bias
for sponsored studies that did find an association, also since they would
go against the common belief that these drugs are safe. Therefore, we
searched any unpublished trials through clinicaltrials.gov, NIH Clinical
Research Studies and the EU Clinical trial registry, but did not identify
any relevant studies. Nevertheless, a reporting bias may also occur,
where positive associations were more likely to have reported.

To assess the clinical importance, it remains important to note
that the absolute risk of congenital malformations is low. Globally,
only one in 200–300 boys develops hypospadias, and even with a
46% risk as estimated in this paper (although not significant), this
would correspond to only 0.5% of all cases of hypospadias being
related to PPI exposure (population attributable fraction), and 31% of
all cases of hypospadias occurring to mothers exposed to PPIs
(attribution fraction), assuming causality and a 1% exposure to PPIs
in the pregnant population. Since this risk would still “only” be one
in approximately 200, this is clearly less than high-risk teratogens
affecting at least 25% of all exposed foetuses, and moderate terato-
gens with a relative risk between 2 and 10.

Previous meta-analyses showed results in line with the pre-
sented findings, but failed to reach statistical significance. Yet, we identified 13 studies which have been published since, in-
creasing power and enabling subgroup analyses (although power
was still limited). Yet, power was still limited for the different sub-
group analyses, which may also explain why no significant increased
risk of congenital malformations was found when comparing PPIs
with H2RA directly. Power calculations (assuming alpha = 0.05
and power = 0.80) showed that, assuming a 1% use of PPIs among
pregnant women, 0.85 million pregnancies are needed to detect a
45% increased risk of hypospadias (assuming −1/200) or 0.48 mil-
lon for detecting a 30% increase for all congenital malformations
assuming −2/100). With only two studies over 800000 pregnan-
cies, and three more over 200000, it is not surprising

that most individual studies did not reach statistical significance,
and that case-control studies are usually preferred for investigat-
ing congenital malformations. Although the described potential
associations are unlikely due to direct effects of PPIs, indirect ef-
fects may occur through changes in the maternal microbiome, in
the gut, vagina or other niches. The increasing amount of stud-
ies linking all sorts of outcomes to long-term PPI use in adults (incl.
cancer, kidney disease, mortality, cardiac events, ...), combined
with potential associations with other outcomes described for in-
utero exposure (including asthma), we believe a risk-benefit as-
essment for use during pregnancy is warranted, and more active
discouragement of unsupervised over-the-counter use.

Future research is needed to establish or narrow down the harm-
ful window of exposure, in particular for non-teratogenic outcomes
such as preterm birth; and the effect of pre-gestational exposure.
This would ideally be based on sufficiently powered nationwide com-
plete data on PPI use (both prescribed and over-the-counter), with
exact information on compliance, dosage and timing of exposure,
enabling assessment of competing effects of the different outcomes
such as spontaneous abortions and termination of pregnancy (for
medical reasons).

To conclude, this meta-analysis is the first to show an increased risk
of congenital malformations associated with PPI use during pregnancy,
in particular when based on case-control studies, although power was
low in all subgroup analyses. These findings suggest that PPIs should
ideally be restricted to well-described indications if no other available

treatment options are available and if the benefits outweigh the risks.
No conclusions could be drawn for the other pregnancy outcomes.

ACKNOWLEDGEMENT
Declaration of personal interest: None.

AUTHORSHIP
Guarantor of the article: Nele Brusselaers.

Author contributions: Ms Cheng Mei Li and Associate Professor
Nele Brusselaers conceptualised and designed the study, conducted
the systematic literature search, quality assessment and meta-analy-

teses; drafted the initial manuscript and reviewed and revised the man-
uscript. Professors Alexandra Zhernakova, Cisca Wijmenga and Lars
Engstrand provided important feedback on the proposed study design,
contributed significantly to the interpretation of the results, critically
reviewed the manuscript for important intellectual content and revised
the manuscript. All authors approved the final manuscript as submitted
and agree to be accountable for all aspects of the work.

ORCID
Nele Brusselaers https://orcid.org/0000-0003-0137-447X

REFERENCES
therapy with proton pump inhibitors and risk of gastric cancer: a
2017;7:e017739.


59. Moretti ME. Personal communication on use of proton pump inhibitors during pregnancy and rates of major malformations. 2002.


SUPPORTING INFORMATION

Additional supporting information will be found online in the Supporting Information section.

How to cite this article: Li CM, Zhernakova A, Engstrand L, Wijmenga C, Brusselaers N. Systematic review with meta-analysis: the risks of proton pump inhibitors during pregnancy. Aliment Pharmacol Ther. 2020;00:1-11. https://doi.org/10.1111/apt.15610