

Influenza and congenital anomalies: a systematic review and meta-analysis

J.M. Luteijn*, M.J. Brown, and H. Dolk

Institute of Nursing Research/School of Nursing, University of Ulster, Jordanstown Campus, Shore Road, Newtownabbey BT37 0QB, UK

*Correspondence address. E-mail: jm.luteijn@ulster.ac.uk

Submitted on August 27, 2013; resubmitted on November 12, 2013; accepted on November 21, 2013

STUDY QUESTION: Does first trimester maternal influenza infection increase the risk of non-chromosomal congenital anomalies (CA)?

SUMMARY ANSWER: First trimester maternal influenza exposure is associated with raised risk of a number of non-chromosomal CA, including neural tube defects, hydrocephaly, congenital heart defects, cleft lip, digestive system defects and limb reduction defects.

WHAT IS KNOWN ALREADY: Hyperthermia is a well-established risk factor for neural tube defects. Previous studies suggest influenza may be a risk factor not only for neural tube defects, but also other CA. No systematic review has previously been undertaken.

STUDY DESIGN, SIZE, DURATION: Systematic review and meta-analysis. A search of EMBASE and PUBMED was performed for English and Dutch studies published up to July 2013. A total of 33 studies (15 case–control, 10 cohort and 8 ecological) were included in the systematic review of which 22 studies were included in the meta-analysis.

PARTICIPANTS/MATERIALS, SETTINGS, METHODS: A total of 29 542 babies with congenital anomaly (1112 exposed) from case–control studies and 1608 exposed pregnancies resulting in 56 babies with congenital anomaly from cohort studies were included in the meta-analysis. Maternal influenza exposure was defined as any reported influenza, influenza-like illness or fever with flu, with or without serological or clinical confirmation during the first trimester of pregnancy. Data for 24 (sub)groups with congenital anomaly available from ≥ 3 studies were analysed using the DerSimonian–Laird random effects model. The hypothesis of publication bias was assessed using funnel plots and risk of bias of included studies was assessed using a slightly modified version of the Newcastle–Ottawa scale.

MAIN RESULTS AND THE ROLE OF CHANCE: First trimester maternal influenza exposure was associated with an increased risk of any congenital anomaly [adjusted odds ratio (AOR) 2.00, 95% CI: 1.62–2.48], neural tube defects [odds ratio (OR) 3.33, 2.05–5.40], hydrocephaly (5.74, 1.10–30.00), congenital heart defects (1.56, 1.13–2.14), aortic valve atresia/stenosis (AOR 2.59, 1.21–5.54), ventricular septal defect (AOR 1.59, 1.24–2.14), cleft lip (3.12, 2.20–4.42), digestive system (1.72, 1.09–2.68) and limb reduction defects (2.03, 1.27–3.27). An increased risk for cleft lip (but not for cleft palate) was also reported by ecological studies not included in the meta-analysis. Study outcomes reported for 27 subgroups of congenital anomaly could not be included in the meta-analysis. Visual inspection of funnel plots did not suggest evidence for publication bias.

LIMITATIONS, REASONS FOR CAUTION: This study enrolled observational studies that can be subject to limitations such as confounding, retrospective maternal exposure reports and non-response of intended participants. Influenza exposed pregnancies can also have been exposed to influenza related medication.

WIDER IMPLICATIONS OF THE FINDINGS: Prevention of influenza in pregnant women may reduce congenital anomaly risk, and would be relevant to more than just neural tube defects. More research is needed to determine whether influenza and/or its related medication is teratogenic, to determine the role of hyperthermia in teratogenicity and the role of other environmental factors such as nutritional status in determining susceptibility.

STUDY FUNDING/COMPETING INTERESTS: Funded by the EC, under the framework of the EU Health Programme 2008–2013, Grant Agreement 2010 22 04 (Executive Agency for Health & Consumers). No competing interests.

Key words: influenza / congenital anomalies / meta-analysis / observational studies / public health

Introduction

Both during seasonal influenza and pandemic influenza outbreaks, pregnant women have been at risk of increased morbidity and mortality from influenza infection compared with the general population (Harris, 1919; Neuzil et al., 1998; Dodds et al., 2007; Siston et al., 2010). Women in the later stages of pregnancy are particularly vulnerable to adverse health outcomes after influenza infection (Mak et al., 2008), perhaps because of immunological changes that take place during pregnancy (Jamieson et al., 2006).

Unravelling the question of teratogenicity of influenza is complex. In observational studies, influenza exposure can affect the fetus not only via viral infection of the fetus, influenza-induced hyperthermia and toxic metabolites associated with fever (Edwards, 2006), but also via antiviral and antipyretic use. A recent systematic review found strong evidence of an association between maternal hyperthermia and neural tube defects (Moretti et al., 2005). Evidence on other anomalies with respect to hyperthermia or fever is scarce. Animal models have associated maternal hyperthermia with arthrogryposis (Edwards, 1971a,b), congenital heart defects (Cockcroft and New, 1975, 1978), club foot (Edwards, 1971a,b), microcephaly (Edwards, 1971a,b), microphthalmos (Germain et al., 1985) and others (Edwards, 2006).

The primary method of protecting pregnant women and their unborn child against influenza infection is vaccination. In an increasing number of countries, pregnant women are advised to be vaccinated against seasonal influenza infection (Mak et al., 2008; Mereckiene et al., 2010). However, vaccination policies in European countries vary both for seasonal influenza vaccination and during the 2009 H1N1 influenza pandemic, especially with respect to trimesters eligible for vaccination (Mereckiene et al., 2010; Luteijn et al., 2011). In the absence of consensus, on whether or not to vaccinate first trimester pregnant women, the hypothesis of a causal relationship between congenital anomalies (CA) and influenza virus deserves renewed attention. A better understanding of the possible relationship between influenza and CA will allow for better understanding of the benefit-risk balance of vaccinating first trimester pregnant women and women of childbearing age against influenza.

The objective of our review is to identify and summarize the available epidemiologic evidence regarding the risk of CA associated with first trimester exposure to maternal influenza.

Materials and Methods

Search strategy

This systematic review was informed by PRISMA guidelines (Moher et al., 2009) and the MOOSE group guidelines (Stroup et al., 2000). Two of the authors (J.M.L. and M.J.B.) conducted the various steps of the review and resolved any disagreements by discussion and consensus. The PubMed® and Embase® databases were searched using the MeSH terms ('Influenza, Human') AND ('pregnancy OR congenital abnormality') and ('Influenza') AND ('pregnancy OR congenital abnormality'), respectively, on 1 July 2013. No publication or date restrictions were set. Where the papers' abstract, title or indexed MeSH terms suggested the possibility of reporting any fetal outcomes after maternal exposure to influenza, the full paper was obtained. Reference lists of enrolled papers were reviewed.

Eligibility criteria

Case-control, cohort and ecological studies investigating CA outcomes following maternal exposure to influenza were eligible for inclusion. No quality

criteria were set for inclusion, although risk of bias analysis was performed (see [Supplementary data, Material](#)). Influenza was defined as any reported influenza, influenza-like illness or fever with flu, with or without serological or clinical confirmation.

Only studies reporting influenza exposures during the first trimester of pregnancy were included in the systematic review and meta-analysis. In order to be included in the meta-analysis, case-control and cohort studies needed to allow for the calculation of or to report odds ratios (ORs) or relative risks (RRs). For financial reasons, only English and Dutch language papers were eligible for inclusion. However, no Dutch papers satisfied our inclusion criteria and therefore only English language papers were included.

Data extraction

Study characteristics were extracted by J.M.L. and M.J.B. ([Supplementary data, Tables SI–SIII](#)). Crude and adjusted ORs and RRs and 2 × 2 tables relevant for meta-analysis were extracted by J.M.L. and M.J.B. from cohort and case-control studies. We contacted authors of three studies published ≤ 10 years ago (Czeizel et al., 2008; Oster et al., 2011; Kelly et al., 2012) to obtain core data in order to create aggregate groups such as orofacial clefts and for crude OR calculation and received the complete dataset for one study (Czeizel et al., 2008). In case of studies distinguishing between flu with fever and flu without fever where it was impossible to combine the two, such as the stratified study by Lynberg et al. (1994), the flu with fever dataset was extracted. We extracted data regarding malformed controls, if available, rather than non-malformed controls since use of malformed controls reduces the impact of differential recall bias. We recognize this can lead to underestimation of effect size if CA in the control group are related to influenza exposure. In the meta-analysis, influenza exposures outside of the first trimester were added to the non-exposed cohort and for one cohort study without controls (Doll et al., 1960), this allowed us to form a control group.

CA were classified into European surveillance of Congenital Anomalies (EUROCAT) defined subgroups, excluding minor anomalies as specified by EUROCAT (EUROCAT Central Registry, 2009). EUROCAT is a network of population-based congenital anomaly registries that surveys over 1.7 million births annually in 23 European countries. CA were classified down to the greatest level of precision possible. For example, a study that reported anomalies only as neural tube defects without further specification contributed data to the analysis of neural tube defects but could not contribute data to an analysis of anencephaly or spina bifida specifically. We excluded chromosomal syndromes since their aetiology is not related to maternal exposures.

Risk of bias assessment

In order to assess the validity of included studies' findings we assessed the risk of bias of case-control and cohort studies included in the meta-analysis using a slightly modified version of the Newcastle-Ottawa scale (NOS; Wells et al., 2004). Judgement criteria and deviations from the NOS are summarized in the [Supplementary data, Material](#).

Quantitative data summary and synthesis

Meta-analysis was performed combining adjusted ORs (AORs) and RRs (ARRs) (i.e. adjusted for confounders). Only four case-control studies provided AORs ([Supplementary data, Table SII](#)) and none of the cohort studies made statistical adjustments ([Supplementary data, Table SIII](#)). Where adjusted estimates were not available, crude estimates were used.

We assumed similarity between OR and RR because CA are rare events (Davies et al., 1998). Statistical analysis was performed using Stata version 9.2 (StataCorp, College Station, TX). Meta-analysis was performed on all CA combined and for EUROCAT defined subgroups of CA (if n studies ≥ 3). The DerSimonian and Laird random effects model (DerSimonian and

Laird, 1986) was used since the studies in this meta-analysis involved varying countries, time periods and influenza strains. Subgroup analysis was performed based on study type, publication date, risk of differential recall bias and adjustment for confounders in order to assess the impact of these variables on study outcome. Subgroup analyses combined all studies in the relevant categories, using the estimate for all non-chromosomal CA combined where available and if not available, the estimate for the specific CA subgroup studied.

Owing to scarce numbers and imbalance between some study arms, we used an alternative continuity correction based on the OR of other studies with >0 events in both arms and group ratio imbalance as discussed by Sweeting *et al.* (2004). Heterogeneity between studies was assessed using the I^2 statistic. Values of I^2 equal to 25, 50 and 75% were considered to represent low, moderate and high levels of heterogeneity, respectively. The hypothesis of publication bias was assessed using funnel plots (data not shown).

Results

Selection flow

The PubMed® database search yielded 1369 papers and the Embase® database search yielded 2649 papers (Fig. 1). After removing 1121 duplicates, a total of 2897 potentially relevant papers were identified by the literature search. After screening by MeSH terms, titles and abstracts, 2615 papers were excluded and full papers were retrieved for the remaining 282 papers. Of these, a total of 40 papers covering 27 studies met the inclusion criteria and were included in the systematic review. Six additional eligible papers were detected by reference tracking, leading to a grand total of 46 included papers covering 33 studies.

Study characteristics

The 46 enrolled papers were classified as 25 papers covering 15 case–control studies, 12 papers covering 10 cohort studies and 9 papers covering 8 ecological studies. Enrolled studies are summarized by study type

in [Supplementary data, Tables SI–SIII](#) and evidence provided by enrolled studies that was not included in the meta-analysis is summarized in Table I. For one case–control study information was limited to a conference abstract (Choi and Klaponski, 1970). Included papers were published between 1953 and 2013, with the median year of publication 1971.

Risk of bias assessment

Visual inspection of the funnel plots did not suggest evidence for publication bias (data not shown). Of the 15 case–control studies (25 papers), 10 studies did not take into account possible confounding by maternal age, socioeconomic class or both. In the majority of these 10 studies some form of matching between cases and controls (usually maternal ward, sex and day of birth) took place. Ten papers relied on retrospective maternal reported influenza episodes (or timing of maternal interviews was unknown), making these studies susceptible to differential recall bias. Of the remaining five studies, two used serologic confirmation and three used prospectively collected antenatal records. The last notable source of possible bias was that in five case–control studies over 20% of the cases intended for inclusion were not enrolled, making these studies susceptible to non-response bias ([Supplementary data, Material](#)).

For the 10 cohort studies (12 papers), none took into account possible confounding by maternal age, socioeconomic class or both. For nine cohort studies, infants were not followed up for at least a year (or unclear), making these studies susceptible to misclassification bias for some CA not apparent at birth. Six of the studies used prospectively collected maternal reports for exposure, three serologic confirmations, for one study exposure ascertainment was not described. The last notable source of possible bias was that in five cohort studies the exposed cohort was not drawn from a clearly defined place and time, or failed to enrol $\geq 80\%$ of the population identified in specified place and time, raising questions over representativeness of the enrolled exposed cohort.

Quantitative data summary and synthesis

Meta-analysis was possible for data from 22 studies, forming groups of ≥ 3 independent studies for 24 (sub)groups of EUROCAT defined major CA: any non-chromosomal major CA, neural tube defects, anencephaly, encephalocele, spina bifida, hydrocephaly, congenital heart defects, orofacial clefts (Figs 2–9) and 16 other CA subgroups (Table II).

Overall, our meta-analysis involved 29 542 CA cases of which 1112 were exposed to influenza in the first trimester of pregnancy and 53 089 controls of which 1382 were exposed to influenza in the first trimester of pregnancy from case–control studies. From cohort studies, 1608 exposed pregnancies resulting in 56 CA plus 14 613 non-exposed pregnancies resulting in 347 CA were enrolled. The enrolled cohort studies were relatively small with only the Coffey and Jessop study (Coffey and Jessop, 1959, 1963) reaching over 50 CA (including minor anomalies) following maternal influenza exposure. Case–control studies more readily enrol the large number required for research on CA and 6 out of 15 case–control studies enrolled over 500 cases (Laurence *et al.*, 1968; Saxen, 1975a; Granroth *et al.*, 1978; Botto *et al.*, 2001; Czeizel *et al.*, 2008; Oster *et al.*, 2011). The larger numbers come at a cost and 11 out of 15 case–control studies gathered exposure data by retrospective maternal reports.

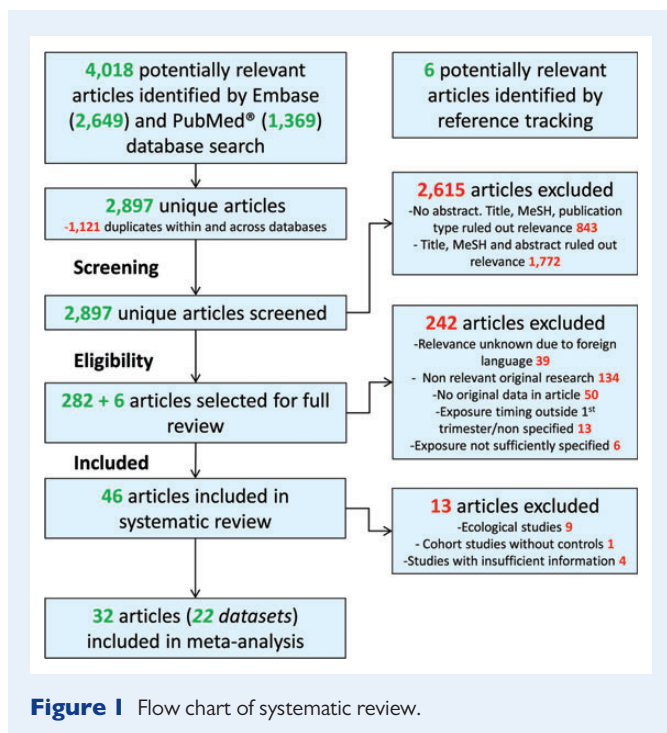


Table 1 Evidence included in the systematic review relating to first trimester influenza exposure and CA not included in meta-analysis.

Anomaly	Evidence reported with respect to first trimester influenza exposure
Any anomaly	<i>Visual inspection:</i> no convincing evidence that infant death rates from CA are associated with first trimester influenza exposure (Buck, 1955); RR 1.10 (Leck, 1963); single defect RR 1.00, multiple defects RR 0.9 (Leck, 1964); single defect RR 1.03, multiple defects RR 1.07 (Leck et al., 1969)
Central Nervous System defects	0.81% in influenza exposed group versus 0.30% and 0.32% in control groups (Hakosalo and Saxen, 1971)
Neural Tube Defects	$P = 0.03$, no data (Choi and Klaponski, 1970)
Anencephaly	RR 0.82 (Leck, 1963); standardized rate ratio (sRR) 1.32 (Leck et al., 1969); RR 1.0, 0.8–1.3 (Saxen et al., 1990). No rise in anencephaly rates following influenza epidemics detected (Record, 1961)
Spina bifida + encephalocele	RR 0.99 (Leck, 1963), sRR 1.13 (Leck et al., 1969)
Hydrocephaly	sRR 0.93 (Leck et al., 1969)
Microcephaly	OR 1.1, 0.3–4.1 (Czeizel et al., 2008); sRR 0.24 (Leck et al., 1969)
Anophthalmos/microphthalmos	OR 1.26, 1.02–1.57 (Busby et al., 2005)
Congenital cataract	OR 0.8, 0.3–2.2 (Czeizel et al., 2008)
Congenital glaucoma	OR 2.0, 0.4–10.9 (Czeizel et al., 2008)
Congenital heart defects	(suggestive but not significant; 0.93% in influenza exposed group versus 0.47% and 0.73% in control groups) (Hakosalo and Saxen, 1971)
Atrioventricular septal defect	OR 2.0, 0.3–15.3 (Botto et al., 2001); AOR 1.29, 0.82–2.01 (Oster et al., 2011)
Tetralogy of Fallot	OR 0.5, 0.1–3.6 (Botto et al., 2001); AOR 0.78, 0.37–1.62 (Oster et al., 2011)
Tricuspid atresia and stenosis	AOR 7.9, 0.3–29.6 (Botto et al., 2001); AOR 6.04, 2.36–15.42 (Oster et al., 2011)
Ebstein's anomaly	OR 3.0, 0.4–23.9 (Botto et al., 2001)
Pulmonary valve stenosis	AOR 1.21, 0.71–2.04 (Oster et al., 2011)
Pulmonary valve atresia	AOR 2.71, 1.16–6.32 (Oster et al., 2011)
Coarctation of aorta	AOR 3.8, 1.6–8.8 (Botto et al., 2001); AOR 0.41, 0.13–1.33 (Oster et al., 2011)
Total anomalous pulm venous	OR 2.2, 0.3–16.9 (Botto et al., 2001)
Cleft lip	RR 1.55, $P < 0.05$ (Leck, 1963); isolated RR 1.4, with other defects RR 6.3 (Leck, 1964); cleft lip without cleft palate sRR 1.47, $P < 0.05$ (Leck et al., 1969); cleft lip without cleft palate RR 1.64, $P < 0.05$ (Leck, 1971); cleft lip without cleft palate sRR 1.08, cleft lip with cleft palate sRR 1.01 (Leck et al., 1969); cleft lip with cleft palate sRR 1.10 (Leck et al., 1969), cleft lip with cleft palate RR 1.35 (Leck, 1971)
Cleft palate	RR 0.81 (Leck, 1963); sRR 0.94 (Leck et al., 1969); sRR 0.93 (Leck et al., 1969)
Digestive system	1 in 63 exposed during first trimester of pregnancy and 4 in 1106 non-exposed, RR 4.4 (Coffey and Jessop, 1959, 1963); 1 in 171 exposed during first trimester of pregnancy and 13 in 6720 non-exposed, RR 3.0 (Hirvensalo and Kinnunen, 1962)
Oesophageal atresia with/without tracheo-oesophageal fistula	OR 1.6, 0.2–11.7 (Czeizel et al., 2008); RR 2.41, $P < 0.01$ (Leck, 1963); isolated RR 1.5, with other defects RR 4.8 (Leck, 1964); sRR 1.12 (Leck et al., 1969)
Atresia/stenosis of the small intestine	OR 2.1, 0.5–8.0 (Czeizel et al., 2008)
Anorectal atresia and stenosis	OR 1.1, 0.4–3.0 (Czeizel et al., 2008); RR 2.12, $P < 0.05$ (Leck, 1963); isolated RR 0.8, with other defects RR 3.6 (Leck, 1964); sRR 0.80 (Leck et al., 1969)
Hirschsprung's disease	OR 0.7, 0.2–3.0 (Czeizel et al., 2008)
Diaphragmatic Hernia	OR 3.2, 1.2–8.8 (Czeizel et al., 2008); 0 in 171 exposed during first trimester of pregnancy and 3 in 6720 non-exposed (Hirvensalo and Kinnunen, 1962); RR 1.05 (Leck, 1963); sRR 1.01 (Leck et al., 1969)
Abdominal Wall Defects	OR 2.8, 1.1–6.9 (Czeizel et al., 2008)
Omphalocele	RR 1.92, < 0.05 (Leck, 1963); isolated RR 1.3, with other defects RR 2.3 (Leck, 1964); sRR 1.16 (Leck et al., 1969)
Urinary	1 in 171 exposed during first trimester of pregnancy and 42 in 6720 non-exposed (Hirvensalo and Kinnunen, 1962)
Bilateral renal agenesis including Potter syndrome	OR 0.8, 0.2–2.8 (Czeizel et al., 2008); RR 1.38 (Leck, 1963)
Congenital hydronephrosis	RR 1.63, $P = 0.05$ (Leck, 1963)
Hypospadias	RR 1.51 (Leck, 1963); sRR 0.93 (Leck et al., 1969)
Limb	sRR 0.75 (Leck et al., 1969)
Limb reduction	Limited to thumbs or radii RR 2.20 (Leck, 1963); both arms or legs RR 1.30 (Leck, 1963); one arm or leg RR 1.37 (Leck, 1963); sRR 1.32 (Leck et al., 1969); RR 1.22 (Leck, 1971)

Continued

Table I Continued

Anomaly	Evidence reported with respect to first trimester influenza exposure
Upper limb reductions	sRR 1.91, $P < 0.01$ (Leck et al., 1969)
Lower limb reductions	sRR 1.23 (Leck et al., 1969)
Club foot	sRR 0.89 (Leck et al., 1969)
Hip dislocation and/or dysplasia	sRR 0.77 (Leck et al., 1969)
Disorders of skin	0 in 171 exposed during first trimester of pregnancy and 3 in 6720 non-exposed (Hirvensalo and Kinnunen 1962)
Down's syndrome	sRR 1.16 (Leck et al., 1969)

Solely data involving > 2 CA cases was included and some reported CA could not be translated to EUROCAT defined subgroups. We did not distinguish between 'positive' and 'negative' findings since some of the studies were severely underpowered. Note for some CA, ≥ 3 estimates were available from studies that could not be included in the meta-analysis.

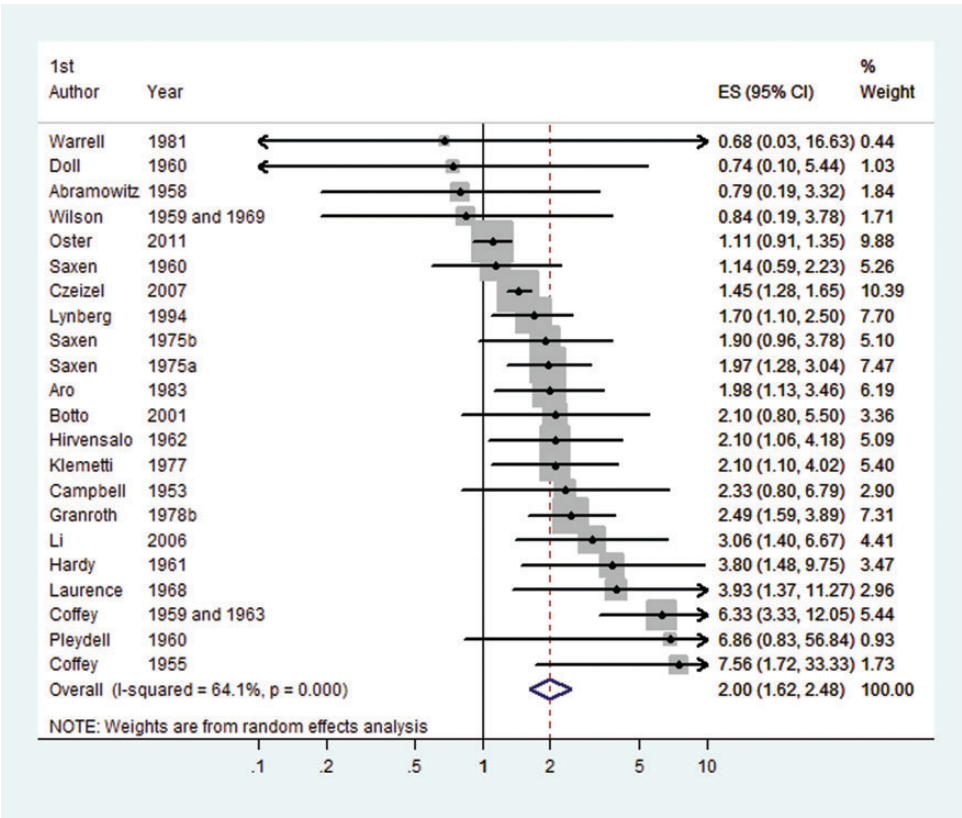


Figure 2 Forest plot of non-chromosomal CA following first trimester influenza exposure. ES, effect size.

Meta-analysis discovered statistically significant associations between first trimester influenza exposure and a large number of CA subgroups (Table II) including all non-chromosomal CA combined (OR 2.00, 95% CI: 1.62–2.48). Medium heterogeneity was detected for the pooled estimate of all non-chromosomal CA (64%). Subgroup analysis showed lower ORs for pooled case–control study outcomes (OR 1.84, 95% CI: 1.49–2.27), than for pooled cohort study outcomes (OR 2.12, 95% CI: 1.20–3.75) while pre-1970 studies reported higher ORs (2.47, 95% CI: 1.50–4.70) than studies published after 1970 (OR 1.71, 95% CI: 1.41–2.08). Overall, studies susceptible to differential recall bias reported a lower risk (OR 1.92, 95% CI: 1.35–2.72) than studies

not susceptible to differential recall bias (OR 2.12, 95% CI: 1.54–2.92). No differences were detected between pooled adjusted and pooled crude estimates (2.15, 1.05–4.42 versus 2.22, 1.78–2.77).

Central nervous system defects

Associations were found for all neural tube defects (OR 3.33, 2.05–5.40) and the neural tube defect subgroups anencephaly (OR 3.52, 1.69–7.32) and spina bifida (OR 2.20, 1.48–3.28). The majority of the 2500 neural tube defects were reported by Czeizel ($n = 1202$, AOR 2.40, 1.30–4.40), Li (344, AOR 3.06, 1.40–6.67), Laurence ($n = 551$, OR 3.93,

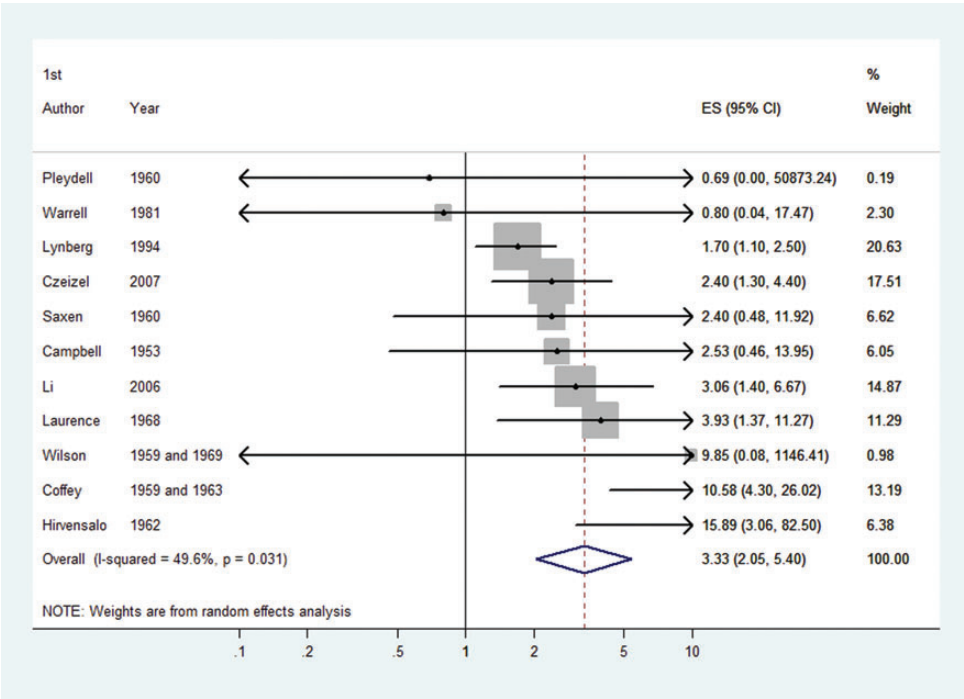


Figure 3 Forest plot of neural tube defects following first trimester influenza exposure. ES, effect size.

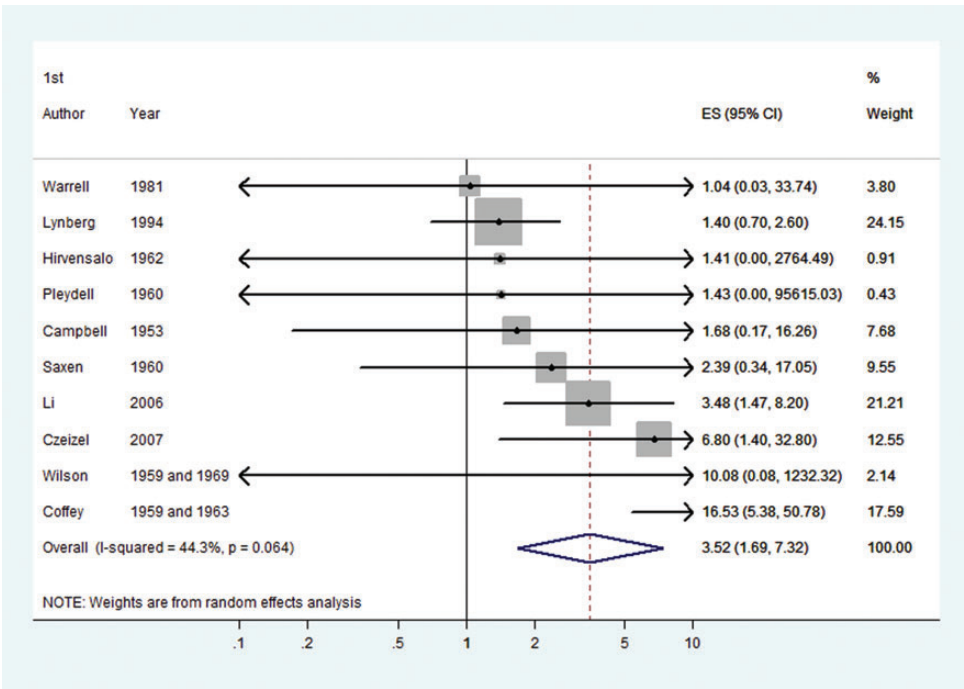


Figure 4 Forest plot of anencephaly following first trimester influenza exposure. ES, effect size.

1.37–11.27) and Lynberg (331, AOR 1.70, 1.10–2.50). The lower OR reported by Lynberg is related to our preference for malformed controls over healthy controls for the study by Lynberg, which lead to more conservative estimates. The study by Lynberg reports higher AOR when using healthy controls for neural tube defects (AOR 3.0, 1.9–4.7). We discovered significant heterogeneity in the aggregate groups

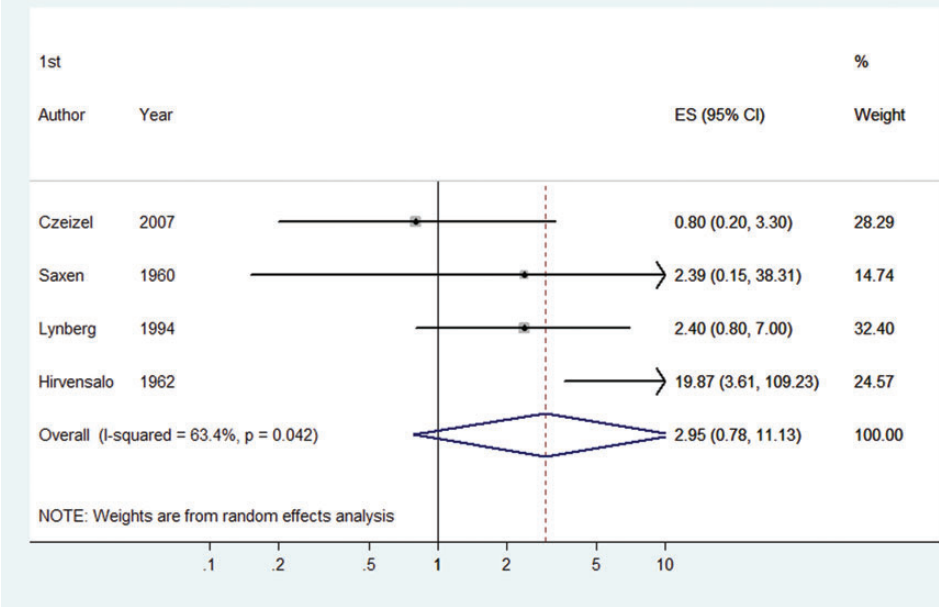


Figure 5 Forest plot of encephalocele following first trimester influenza exposure. ES, effect size.

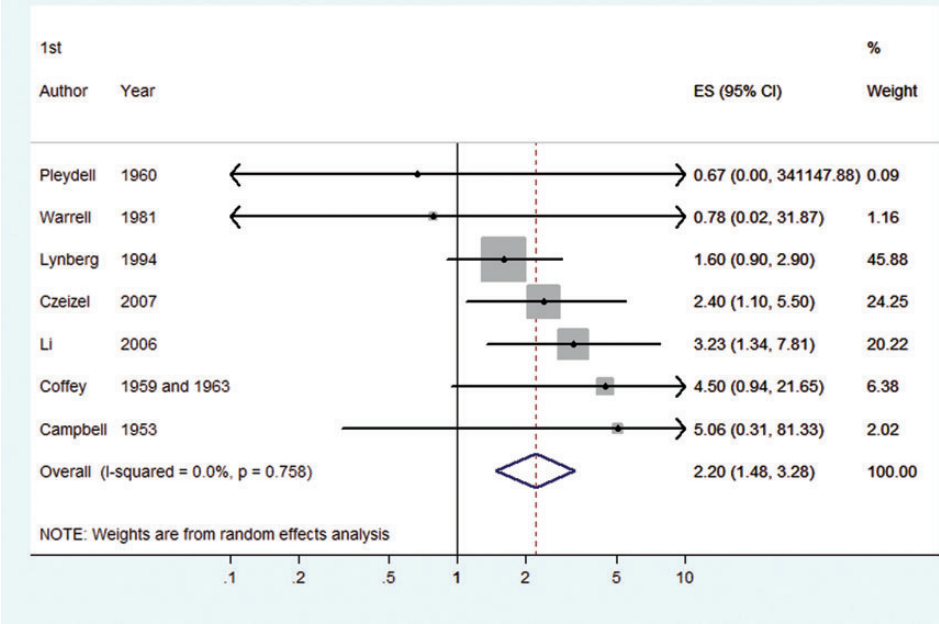


Figure 6 Forest plot of spina bifida following first trimester influenza exposure. ES, effect size.

and neural tube defects, anencephaly, encephalocele and hydrocephaly (Table II). The heterogeneity for neural tube defects, anencephaly, encephalocele and hydrocephaly seems to be driven by the Coffey (Coffey and Jessop, 1959, 1963), Hirvensalo (Hirvensalo and Kinnunen, 1962), Pleydell (Pleydell, 1960), Saxen et al. (1960) and Wilson studies (Wilson et al., 1959; Wilson and Stein, 1969) which all contributed OR

of > 10 in at least one CA subgroup. The ecological study on 1957 pandemic influenza by Hakosalo, enrolling 27 neural tube defects reported suggestive evidence for a relationship between influenza and neural tube defects, but two larger ecological studies by Leck [$n = 2484$ (Leck et al., 1969) and $n = 162$ (Leck, 1963)] did not find such evidence for any neural tube defects subgroup. The study by Hakosalo calculated cases

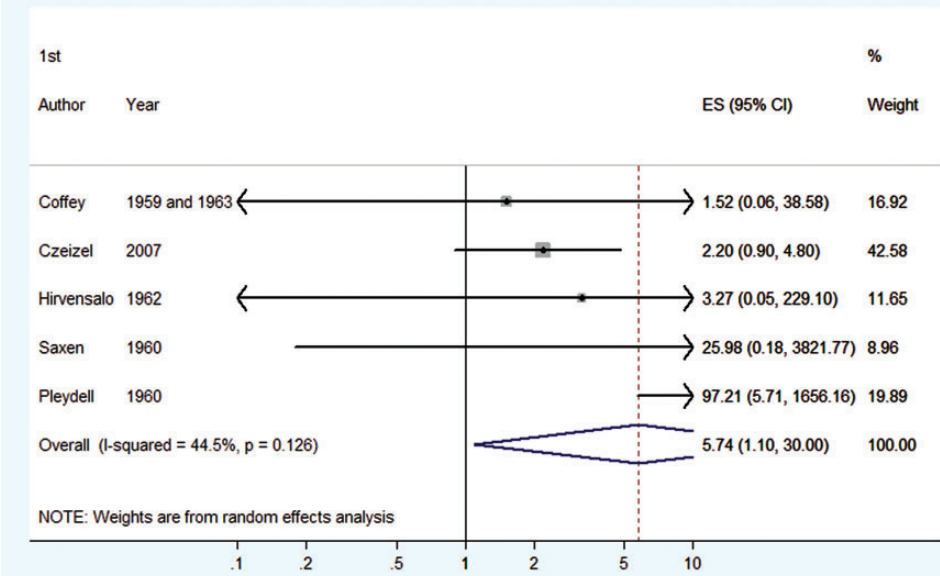


Figure 7 Forest plot of hydrocephaly following first trimester influenza exposure. ES, effect size.

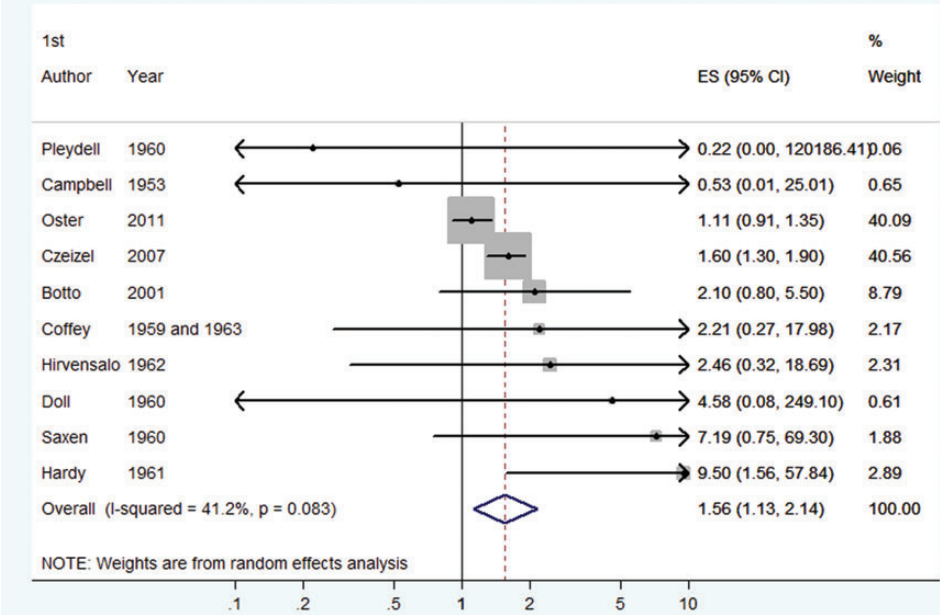


Figure 8 Forest plot of congenital heart defects following first trimester influenza exposure. ES, effect size.

back to last menstrual period, while none of the studies by Leck had access to gestational length, making these studies susceptible to misclassification of exposure introduced by assumptions around gestational age. Of all eight ecological studies included, only three corrected for gestational age (Hakosalo and Saxen, 1971; Saxen et al., 1990; Busby et al., 2005).

Orofacial clefts

Our meta-analysis found an association between orofacial clefts and first trimester influenza exposure (OR 1.96, 95% CI: 1.33–2.91). There was a significant association for cleft lip with or without palate (OR 3.12, 2.20–4.42), but not for cleft palate (OR 1.05, 0.60–1.84) and no heterogeneity

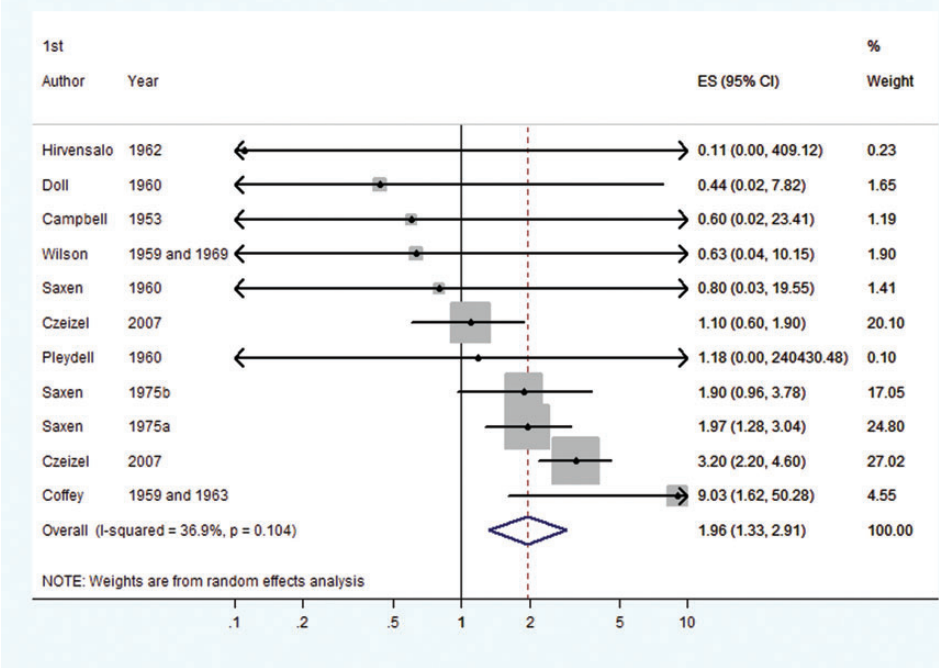


Figure 9 Forest plot of orofacial clefts following first trimester influenza exposure. ES, effect size.

was detected in these pooled groups (Table I). The majority of the orofacial clefts ($n = 2773$) were reported by Czeizel (Czeizel *et al.*, 2008; $n = 1956$) and Saxen (Saxen, 1975a; $n = 591$ and 1975b, $n = 194$) and these studies report ORs between 1.90 and 2.32. Four of the ecological studies, all by Leck, also reported the association between orofacial clefts and influenza and two of these four studies reported associations for cleft lip \pm cleft palate, but not for isolated cleft palate (Leck, 1963; Leck *et al.*, 1969).

Congenital heart defects

Our meta-analysis found an association between congenital heart defects and first trimester influenza exposure (OR 1.56, 95% CI: 1.13–2.14). The vast majority of the congenital heart defects reported in the meta-analysis were reported by Botto ($n = 829$, AOR 2.1, 0.8–5.5), Czeizel ($n = 4479$, OR 1.6, 1.3–1.9) and Oster ($n = 2361$, AOR 1.11, 0.91–1.35) (Botto *et al.*, 2001; Czeizel *et al.*, 2008; Oster *et al.*, 2011). It should be noted the study by Botto suffered from a 2–12 year delay between delivery and maternal interview while influenza exposure in the study by Oster could occur from 3 months before pregnancy to the third month of pregnancy.

With respect to specific types of congenital heart defects (Table II), meta-analysis showed aortic valve atresia/stenosis and ventricular septal defect to be associated with first trimester influenza exposure (OR 2.59, 1.21–5.54 and OR 1.59, 1.24–2.04, respectively). No associations were found for atrial septal defect, hypoplastic left heart and transposition of the great vessels. Four congenital heart defect subtypes were only reported by ≤ 2 studies and therefore not in the meta-analysis (Table I). Of these, two studies showed a consistent absence of association for tetralogy of Fallot, and a consistent and high association for

tricuspid atresia and stenosis (Botto *et al.*, 2001, OR 7.9; Oster *et al.*, 2011, AOR 6.04).

Other anomalies

Limb reductions were associated with first trimester influenza exposure by the meta-analysis (OR 2.03, 1.27–3.27) and this association is supported by several ecological studies (Table I). An association with anophthalmia/microphthalmia is based on a single study (Table I). There was evidence that there is no association with influenza for hypospadias (OR 1.02, 0.75–1.39) and club foot (OR 1.03, 0.83–1.27).

Discussion

This systematic review provides an overview of the published evidence on influenza exposure during the first trimester of pregnancy and CA. Meta-analysis revealed evidence for increases in a wide range of major CA following first trimester influenza exposure. The 2-fold increase in risk of non-chromosomal CA represents an increase in prevalence from 1.8% (EUROCAT Central Registry, 2013) to 3.6% of births among first trimester influenza exposed pregnancies.

Exposure ascertainment

Case–control and cohort studies utilized serologic confirmation and maternal reports (prospective and retrospective) for influenza exposure ascertainment. During the influenza season the positive predictive value of the presentation of an influenza-like illness for influenza is in the order of 66–77% (Monto *et al.*, 2000; Zambon *et al.*, 2001). Serologic confirmation detects clinical and subclinical infections, which might lead to different results since subclinical infections might affect the pregnant women

Table II First trimester maternal influenza exposure and risk of congenital anomalies: studies, total number of cases of congenital anomaly (CA), pooled OR and heterogeneity.

Group	Participating studies (n)	I ² statistic for heterogeneity (%)	Pooled OR (95% CI)	Total number of CA (n)
Any congenital anomaly	22	64	2.00 (1.62–2.48)	29 945 ^c
Susceptible to differential recall bias	9	65	1.92 (1.35–2.72)	5426
Not susceptible to differential recall bias	13	64	2.12 (1.54–2.91)	24 519 ^c
Case–control studies	13	60	1.84 (1.49–2.27)	29 542 ^c
Cohort studies	9	62	2.12 (1.21–3.72)	403
Any type, published 1955–1969	11	58	2.47 (1.50–4.70)	1171
Any type, published 1975–2011	11	55	1.71 (1.41–2.08)	28 774
Adjusted estimates only ^a	4	87	2.15 (1.05–4.42)	3865
Crude estimates only ^a	21	61	2.22 (1.78–2.77)	27 584
Neural tube defects	11	50	3.33 (2.05–5.40)	2500
Anencephaly	10	44	3.52 (1.69–7.32)	608
Encephalocele	4	63	2.95 (0.78–11.13)	225
Spina bifida	7	0	2.20 (1.48–3.28)	1093
Hydrocephaly	5	45	5.74 (1.10–30.00)	323
Congenital heart defects	10	41	1.56 (1.13–2.14)	7715
Aortic valve atresia/stenosis	3	31	2.59 (1.21–5.54)	167
Atrial septal defect	3	0	0.82 (0.45–1.51)	429
Hypoplastic left heart	3	0	1.58 (0.94–2.64)	203
Transposition of the great vessels	3	0	1.40 (0.90–2.17)	321
Ventricular septal defect	4	0	1.59 (1.24–2.04)	1434
Orofacial clefts	10	37	1.96 (1.33–2.91)	2773 ^d
Cleft lip ± palate ^b	7	0	3.12 (2.20–4.42)	1404
Cleft palate ^b	3	0	1.05 (0.60–1.84)	584
Digestive system ^b	4	0	1.71 (1.09–2.69)	1195
Urinary	5	0	1.45 (0.90–2.34)	48
Hypospadias ^b	4	0	1.02 (0.75–1.39)	3041
Limb reduction	3	0	2.03 (1.27–3.27)	1002
Club foot ^b	4	0	1.11 (0.93–1.34)	2430
Hip dislocation/dysplasia	3	0	0.31 (0.00–37.62)	37
Polydactyly ^b	4	0	1.72 (0.85–3.48)	1094
Syndactyly	3	71	1.98 (0.19–20.563)	662
Musculoskeletal	3	0	1.05 (0.16–6.97)	776

^aThree studies (Lynberg *et al.*, 1994; Botto *et al.*, 2001; Li *et al.*, 2007) were able to provide estimates both for adjusted and crude OR.

^bNote that a single study contributed over 90% of the total weight to this pooled estimate.

^cNote that for some studies such as Czeizel *et al.* (2008), the 'any congenital anomaly' group also included CA which were not included in any other analysis.

^dNote that 591 orofacial clefts from the Saxen (1975a) study and 194 orofacial clefts from the Saxen (1975b) study were not specified and therefore solely included in the overall orofacial clefts analysis.

differently from clinical infections. Arguably, serologic confirmation of exposure is more reliable than maternal reports. Five studies based on serologic confirmation were enrolled in the systematic review (Walker and McKee, 1959; Wilson *et al.*, 1959; Hardy *et al.*, 1961; Elizan *et al.*, 1969; Wilson and Stein, 1969; Warrell *et al.*, 1981) of which three were included in the meta-analysis (Wilson *et al.*, 1959; Hardy *et al.*, 1961; Wilson and Stein, 1969; Warrell *et al.*, 1981). The two serologic studies limited to systematic review did not detect an association between influenza and neural tube defects (Elizan *et al.*, 1969) and did not detect an increased prevalence of CA among 1957 H2N2 Asian

pandemic influenza exposed pregnancies (Walker and McKee, 1959). The three serologic studies in the meta-analysis combined contributed 53 out of 27 584 CA, and it was therefore not possible to examine the effect of exposure ascertainment method on effect size. One of these studies reported a possible association between CA and 1957 H2N2 Asian pandemic influenza (Hardy *et al.*, 1961) while a second study did not (Wilson *et al.*, 1959; Wilson and Stein, 1969). The third study did not detect an association between CA and neural tube defects (Warrell *et al.*, 1981). Note that all studies using serologic confirmation were limited to low numbers.

It is apparent that maternal reports lead to misclassification of exposure as women might not recall infection, timing of infection relative to pregnancy or misdiagnose another infection for influenza. However, as long as maternal reports are collected prospective to the mother being aware of the malformation (e.g. from medical records or interviews during pregnancy), there is no reason to believe misclassification of influenza exposure will differ between cases and controls. Therefore, prospective maternal reports will not lead to a spurious association, but rather bias the estimate towards the null as cases and controls are subject to similar misclassification. Retrospective maternal reports (e.g. interviews after birth), which are frequently utilized by case-control studies, are susceptible to recall bias where mothers of cases have a different motivation to recall early pregnancy exposure than mothers of non-cases. Mothers may differ not only in their tendency to remember the infection, but in their tendency to misinterpret an illness as influenza (MacKenzie and Houghton, 1974). Some of the studies included in the systematic review give an estimate of the differential recall bias' effect. For example, in the study by Lynberg, the OR for anencephaly after flu with fever decreased from 3.1 (95% CI: 1.6–6.1) when compared with non-malformed controls to 1.4 (95% CI: 0.7–2.6) when compared with malformed controls (Lynberg *et al.*, 1994). Part of this decrease might also have been related to CA in the control group being related to influenza, thus biasing the OR towards 1. In our meta-analysis, the pooled estimate of OR for studies susceptible to recall bias (1.92) was slightly lower than the pooled estimate for other studies (2.00), contrary to expectation. This is related to the low overall OR (1.11) reported by the large Baltimore–Washington Infant Study (Oster *et al.*, 2011), which was susceptible to differential recall bias and contributed over 20% to the pooled estimate of susceptible studies.

Case-control, cohort and ecological study designs

Cohort studies failed to enrol large numbers of CA and this should not be surprising considering CA only make up for 2–3% of births in a general population and short follow-up time after birth could have led to underascertainment. Investigation of specific CA requires even higher numbers. Owing to the possibility of enrolling larger numbers, case-control studies seem better suited for addressing the hypothesis of teratogenicity of influenza, although this comes at a cost as most case-control studies ascertained exposure by retrospective maternal reports.

Ecological studies are generally considered a weaker study design than case-control or cohort studies (Evans, 2003) due to lack of individual exposure information. It cannot be verified that any excess CA occurred among infected individuals. Correlation between influenza and confounding risk factors at group level may lead to ecological fallacy, for example if influenza and nutritional deficiencies co-occur in winter. Ecological studies base exposure status on timing of pregnancy relative to influenza season (or a proxy thereof) and therefore, the cohort defined as 'exposed' is diluted by pregnancies that did not have influenza. Population influenza exposure in the eight enrolled ecological studies were derived from influenza incidences, counts or deaths ($n = 5$) and sickness absenteeism rates or claims ($n = 3$). For this reason, the distinguishing power of ecological studies is highly dependent on influenza attack rates and precision used to define influenza exposure.

The advantages of ecological studies are that large numbers of patients are enrolled easily and ecological studies are not susceptible to the

exposure misclassification or recall bias inherent in individual level studies. Furthermore, they can be free of individual level confounding, e.g. if those most susceptible to influenza in the population have other risk factors for CA. Ecological studies therefore offer great value for addressing the hypothesis of teratogenicity of infectious diseases and should not be discounted at the bottom of the evidence hierarchy in this area of research. Consistency between study designs lends strength to a causal interpretation (Hofler, 2005).

Associations between first trimester influenza exposure and CA

One of the most striking results of the meta-analysis is the association between first trimester influenza exposure and neural tube defects. Neural tube defects are easily recognized at birth eliminating susceptibility to underascertainment, while a previous meta-analysis reported an association between neural tube defects and hyperthermia (OR 1.92, 95% CI: 1.62–2.29; Moretti *et al.*, 2005). There also is evidence for neural tube defects following hyperthermia exposure in guinea pigs (Smith *et al.*, 1992). In more recent human studies, underascertainment of neural tube defects could have been a problem due to terminations of pregnancy for fetal anomaly. According to a 2004 EUROCAT analysis, 88% of neural tube defects are detected prenatally of which 88% are aborted (Boyd *et al.*, 2008). Neural tube defects were more frequently studied than other CA, possibly a result of interest in the 1950s study by Coffey, suggesting an alarmingly strong link between maternal influenza exposure and neural tube defects (data corresponds to OR 10.58, 4.30–26.02 in the 1963 follow-up; Coffey and Jessop, 1963). This study utilized standardized questionnaires for maternal interview after delivery and therefore was susceptible to differential recall bias, but this limitation would not explain such a high OR. One of the ecological studies found an increase in neural tube defects during the 1957 Asian influenza outbreak in Finland and concluded this might have been caused by either influenza or influenza-related pharmaceuticals (Hakosalo and Saxen, 1971). The study by Li *et al.* (2007) had data available both on antiviral and antipyretic use (and other potential confounders, see Supplementary data, Table SII) and after adjusting for these co-exposures, OR for neural tube defects following maternal influenza exposure dropped slightly but remained statistically significant (OR 3.93, 95% CI: 2.48–6.23).

Hydrocephaly can sometimes be caused by spina bifida, and we could derive estimates for hydrocephaly not associated with neural tube defects for two of the five studies on hydrocephaly (Coffey and Jessop, 1959, 1963; Hirvensalo and Kinnunen, 1962).

A general problem in CA research is that one baby may have more than one anomaly, and some of these multiple malformed babies may have 'sequences' (EUROCAT Central Registry, 2009) that follow from a primary anomaly. The data available for review cannot distinguish different types of diagnoses.

The evidence for an association between first trimester influenza exposure and cleft lip with or without palate is strong due to the lack of heterogeneity in the pooled data and consistent positive associations detected across different study designs. It is well known that cleft lip \pm palate differs aetiologically from cleft palate, so this difference is not surprising (Mossey *et al.*, 2009). The associations detected for congenital heart defects and ventricular septal defects are somewhat contradicted by the results from the BWIS (Oster *et al.*, 2011). On the other hand

congenital heart defects have been associated with hyperthermia in rats (Cockroft and New, 1975, 1978). Club foot has been associated with hyperthermia in guinea pigs (Edwards, 1971a,b), but our meta-analysis did not find evidence for an association between club foot and influenza. A large number of additional associations for other CA types were detected but with more limited underlying evidence.

Pathways for mediation of hypothetical teratogenic effect of influenza

Influenza can mediate a possible teratogenic effect via multiple pathways and there is a risk of confounding due to the intimate linkage between a disease and its cure. As well as antivirals, antipyretics are also often used during influenza infection and the case-control study by Li et al. (2007) reported an AOR for antipyretic drugs and neural tube defects of 4.86 (95% CI: 1.33–17.78). Associations not adjusted for antivirals or antipyretics remain of importance since from a vaccination policy perspective, it is less relevant whether the influenza virus or the antivirals are causing any possible anomalies as vaccination will prevent both exposures. We recognize this puts limits on generalizability of study findings between populations with different use of antipyretics/antivirals.

Direct pathways via which influenza infection can possibly lead to CA are toxic metabolites caused by fever, hyperthermia and the influenza virus crossing the placenta. Hyperthermia has been associated with causing neural tube defects as discussed above (Moretti et al., 2005). It should be noted this meta-analysis involved a large number of possible causes of hyperthermia, while influenza causes high fever that might be different from general hyperthermia. This could explain the higher OR reported in this meta-analysis for influenza exposure and neural tube defects (Table II). Several of the included studies distinguished between first trimester influenza and first trimester fever (Saxen, 1975a,b; Klemetti, 1977; Aro, 1983; Lynberg et al., 1994; Botto et al., 2001; Oster et al., 2011). One study found an association for influenza, but not for fever (Klemetti, 1977), another found an association for influenza with fever and neural tube defects (OR 1.7, 1.1–2.5), which was lowered for influenza without fever (OR 1.3, 0.7–2.5; Lynberg et al., 1994). A study on congenital heart defects reported associations for fever (OR 1.8, 1.4–2.4) and influenza (OR 2.1, 0.8–5.5) (Botto et al., 2001), while a second study on congenital heart defects reported very similar low and non-significant excesses for fever (OR 1.14, 0.89–1.46) and influenza (OR 1.11, 0.91–1.35; Oster et al., 2011). The two studies by Saxen (1975a,b) on orofacial clefts reported comparable associations for influenza (RR 2.00 for the first study) and fever (RR 1.96 for the first study), and the study by Aro on limb reduction defects reported an OR of 1.6 for fever and 1.9 for influenza (Aro, 1983). It can be concluded that included studies generally reported equivalent or stronger associations for influenza than for fever with respect to CA following first trimester exposure.

Another possible pathway by which the influenza virus can mediate a teratogenic effect is placental transmission. Placental transmission has been documented, but appears to be rare (McGregor et al., 1984; Gu et al., 2007). Toxic metabolites associated with fever as a cause of CA have also been suggested (Edwards, 2006).

Limitations of the study

The systematic review results should be interpreted in the light of the findings that all of the included studies are observational studies, many

were susceptible to several types of bias and ascertainment of exposure might not have been reliable. Although the funnel plots did not provide evidence for publication bias, it should be noted that most of the included studies reported a wide range of positive associations raising the question of whether studies reporting negative results remained unpublished. Reference tracking identified six new studies and these studies were missed because they were not indexed as influenza studies. This leaves the possibility open that some, particularly negative, studies were missed due to poorly indexed terms. A possible reason for this is that some case-control studies investigate a wide range of possible causes of CA and may tend to be selectively indexed for the positive associations. For older studies, we could not be sure whether chromosomal CA were excluded from analysis.

A weakness of the meta-analysis was that adjustment for confounders was not performed in most included studies. Adjustment for confounders showed a moderate effect on OR within studies that did report both crude OR and adjusted OR (Granroth et al., 1978; Acs et al., 2005; Li et al., 2007). Subgroup analysis comparing adjusted versus crude estimates did not detect differences, but this could be related to the fact that 50% of the adjusted estimate was composed of neural tube defects data since studies generally reported higher estimates for neural tube defects than for other CA. Owing to the limited number of studies reporting adjusted OR, comparison of crude and adjusted OR for subgroups of the same CA was not possible. Some very large datasets involved matched and/or stratified controls (Botto et al., 2001; Czeizel et al., 2008), and most other datasets were matched by one or more variables (Laurence et al., 1968; Karkinen-Jaaskelainen and Saxen, 1974; Saxen, 1975a,b; Granroth, 1978; Granroth et al., 1978; Warrell et al., 1981; Aro, 1983; Lynberg et al., 1994; Li et al., 2007) lowering the impact of confounding on the meta-analysis.

Statistical limitations

The study was susceptible to statistical limitations unique to meta-analysis of scarce events. In the context of scarce events like CA and complicated exposure like first trimester influenza, cohort studies can enrol large numbers of exposed and unexposed, but have very few exposed CA outcomes. Current statistical methods will assign these studies a lot of weight compared with case-control studies with many cases and greater numbers of exposed cases. An example from this systematic review is the cohort study by Pleydell, which reported 1 case of hydrocephaly among 12 first trimester influenza exposed pregnancies and 1 case among 1071 unexposed pregnancies leading to an OR of 97.27. The heterogeneity model favours outliers and smaller studies and provided this study with 20% weight in the overall hydrocephaly estimate, compared with 43% weight for the case-control study by Czeizel which enrolled 314 cases of hydrocephaly (16 exposed). The weight allocated by the heterogeneity model to the Pleydell study is clearly disproportionate.

A second problem lies in the continuity correction that is used to address zero events in one of both arms of a study when pooling ORs. For rare events like CA, zero events in one or both arms are not uncommon. The standard value for continuity correction is 0.5, but this arbitrary value causes problems in studies with uneven arms and can even dominate the other arm. We addressed this problem as proposed by Sweeting et al. (2004) by letting the continuity correction depend on the OR of (other) studies with >0 events in both arms and group ratio imbalance.

However, this still led to $OR > 1$ for studies reporting 0 events in the exposed arm such as the hydrocephaly data by Hirvensalo and Kinnunen (1962).

Conclusions and implications for CA prevention

Given the risk of congenital anomaly associated with influenza we show here, prevention of influenza by vaccinating women who are planning to get pregnant may reduce congenital anomaly risk. However, before evidence-based policy can be implemented, further safety data on use of influenza vaccines in pregnancy with respect to CA is required (Kallen and Olausson, 2012; Pasternak et al., 2012). Other methods for preventing influenza in early pregnancy include improving nutritional and general health status and adopting behaviours that prevent interpersonal spread.

In conclusion, prevention of influenza in pregnant women may reduce congenital anomaly risk, and would be relevant to more than just neural tube defects. More research is needed to determine whether influenza and/or its related medication is teratogenic, to determine the role of hyperthermia in teratogenicity and the role of other environmental factors such as nutritional status in determining susceptibility.

Supplementary data

Supplementary data are available at <http://humrep.oxfordjournals.org/>.

Acknowledgements

We would like to thank Dr I. Barisic for assistance with case classification and Dr Julia Métneki, Dr Erzsébet Puhó, Professor Andrew Czeizel and Dr Annukka Ritvanen for sending additional data. We would like to thank Professor Lolkje de Jong-van den Berg and Dr. Gordon Marnoch for their helpful comments.

Authors' roles

J.M.L. was involved in study design, data collection and synthesis, manuscript preparation and revision, construction of tables and figures, statistical analysis and submission of the manuscript. M.J.B. was involved in data collection and manuscript revision. H.D. was involved in study design, writing and revising the manuscript.

Funding

Funded by the EC, under the framework of the EU Health Programme 2008–2013, Grant Agreement 2010 22 04 (Executive Agency for Health & Consumers).

Conflict of interest

No competing interests.

References

- Acs N, Banhidy F, Puho E, Czeizel AE. Maternal influenza during pregnancy and risk of congenital abnormalities in offspring. *Birth Defects Res A Clin Mol Teratol* 2005;**73**:989–996.
- Aro T. Maternal diseases, alcohol consumption and smoking during pregnancy associated with reduction limb defects. *Early Hum Dev* 1983;**9**:49–57.
- Botto LD, Lynberg MC, Erickson JD. Congenital heart defects, maternal febrile illness, and multivitamin use: a population-based study. *Epidemiology* 2001;**12**:485–490.
- Boyd PA, Devigan C, Khoshnood B, Loane M, Garne E, Dolk H. EUROCAT Working Group. Survey of prenatal screening policies in Europe for structural malformations and chromosome anomalies, and their impact on detection and termination rates for neural tube defects and Down's syndrome. *BJOG* 2008;**115**:689–696.
- Buck C. Exposure to virus diseases in early pregnancy and congenital malformations. *Can Med Assoc J* 1955;**72**:744–746.
- Busby A, Dolk H, Armstrong B. Eye anomalies: seasonal variation and maternal viral infections. *Epidemiology* 2005;**16**:317–322.
- Choi NW, Klapanski FA. On neural-tube defects: an epidemiological elicitation of etiological factors. *Neurology* 1970;**20**:399–400.
- Cockroft DL, New DA. Effects of hyperthermia on rat embryos in culture. *Nature* 1975;**258**:604–606.
- Cockroft DL, New DA. Abnormalities induced in cultured rat embryos by hyperthermia. *Teratology* 1978;**17**:277–283.
- Coffey VP, Jessop WJ. Maternal influenza and congenital deformities: a prospective study. *Lancet* 1959;**2**:935–938.
- Coffey VP, Jessop WJ. Maternal influenza and congenital deformities. A follow-up study. *Lancet* 1963;**1**:748–751.
- Czeizel AE, Puho EH, Acs N, Banhidy F. Use of specified critical periods of different congenital abnormalities instead of the first trimester concept. *Birth Defects Res A Clin Mol Teratol* 2008;**82**:139–146.
- Davies HT, Crombie IK, Tavakoli M. When can odds ratios mislead? *Br Med J* 1998;**316**:989–991.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177–188.
- Dodds L, McNeil SA, Fell DB, Allen VM, Coombs A, Scott J, MacDonald N. Impact of influenza exposure on rates of hospital admissions and physician visits because of respiratory illness among pregnant women. *CMAJ* 2007;**176**:463–468.
- Doll R, Hill AB, Sakula J. Asian influenza in pregnancy and congenital defects. *Br J Prev Soc Med* 1960;**14**:167–172.
- Edwards MJ. The experimental production of Arthrogryposis multiplex congenita in guinea-pigs by maternal hyperthermia during gestation. *J Pathol* 1971a;**104**:221–229.
- Edwards MJ. The experimental production of clubfoot in guinea-pigs by maternal hyperthermia during gestation. *J Pathol* 1971b;**103**:49–53.
- Edwards MJ. Review: Hyperthermia and fever during pregnancy. *Birth Defects Res A Clin Mol Teratol* 2006;**76**:507–516.
- Elizan TS, Ajero-Froehlich L, Fabiyi A, Ley A, Sever JL. Viral infection in pregnancy and congenital CNS malformations in man. *Arch Neurol* 1969;**20**:115–119.
- EUROCAT Central Registry. EUROCAT Guide 1.3 and reference documents. Instructions for the Registration and Surveillance of Congenital Anomalies. Revised: May 2009.
- EUROCAT Central Registry. EUROCAT Prevalence Tables 2007–2011. **2013**.
- Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs* 2003;**12**:77–84.
- Germain MA, Webster WS, Edwards MJ. Hyperthermia as a teratogen: parameters determining hyperthermia-induced head defects in the rat. *Teratology* 1985;**31**:265–272.

- Granroth G. Defects of the central nervous system in Finland: III. Disease and drugs in pregnancy. *Early Hum Dev* 1978;**2**:147–162.
- Granroth G, Haapakoski J, Saxen L. Defects of the central nervous system in Finland: V. Multivariate analysis of risk indicators. *Int J Epidemiol* 1978;**7**:301–308.
- Gu J, Xie Z, Gao Z, Liu J, Korteweg C, Ye J, Lau LT, Lu J, Gao Z, Zhang B et al. H5N1 infection of the respiratory tract and beyond: a molecular pathology study. *Lancet* 2007;**370**:1137–1145.
- Hakosalo J, Saxen L. Influenza epidemic and congenital defects. *Lancet* 1971;**2**:1346–1347.
- Hardy JM, Azarowicz EN, Mannini A, Medearis DN Jr, Cooke RE. The effect of Asian influenza on the outcome of pregnancy, Baltimore, 1957–1958. *Am J Public Health Nations Health* 1961;**51**:1182–1188.
- Harris J. Influenza occurring in pregnant women: a statistical study of thirteen hundred and fifty cases. *J Am Med Assoc* 1919;**72**:978–980.
- Hirvensalo M, Kinnunen O. Influenza and pregnancy. *Duodecim* 1962;**78**:740–747.
- Hofler M. The Bradford Hill considerations on causality: a counterfactual perspective. *Emerg Themes Epidemiol* 2005;**2**:11.
- Jamieson DJ, Theiler RN, Rasmussen SA. Emerging infections and pregnancy. *Emerg Infect Dis* 2006;**12**:1638–1643.
- Kallen B, Olausson PO. Vaccination against H1N1 influenza with Pandemrix(R) during pregnancy and delivery outcome: a Swedish register study. *BJOG* 2012;**119**:1583–1590.
- Karkinen-Jaaskelainen M, Saxen L. Maternal influenza, drug consumption, and congenital defects of the central nervous system. *Am J Obstet Gynecol* 1974;**118**:815–818.
- Kelly S, Keuhl K, Leffredo C. Tricupsid Atresia: an epidemiologic investigation in the Baltimore-Washington area. *Am J Epidemiol* 2012;**175**:S59.
- Klemetti A. Definition of congenital malformations and detection of associations with maternal factors. *Early Hum Dev* 1977;**1**:117–123.
- Laurence KM, Carter CO, David PA. Major central nervous system malformations in South Wales. II. Pregnancy factors, seasonal variation, and social class effects. *Br J Prev Soc Med* 1968;**22**:212–222.
- Leck I. Incidence of malformations following influenza epidemics. *Br J Prev Soc Med* 1963;**17**:70–80.
- Leck I. Examination of the incidence of malformations for evidence of drug teratogenesis. *Br J Prev Soc Med* 1964;**18**:196–201.
- Leck I. Further tests of the hypothesis that influenza in pregnancy causes malformations. *HSMHA Health Rep* 1971;**86**:265–269.
- Leck I, Hay S, Witte JJ, Greene JC. Malformations recorded on birth certificates following A2 influenza epidemics. *Public Health Rep* 1969;**84**:971–979.
- Li Z, Ren A, Liu J, Pei L, Zhang L, Guo Z, Li Z. Maternal flu or fever, medication use, and neural tube defects: a population-based case-control study in Northern China. *Birth Defects Res A Clin Mol Teratol* 2007;**79**:295–300.
- Luteijn JM, Dolk H, Marnoch GJ. Differences in pandemic influenza vaccination policies for pregnant women in Europe. *BMC Public Health* 2011;**11**:819.
- Lynberg MC, Khoury MJ, Lu X, Cocian T. Maternal flu, fever, and the risk of neural tube defects: a population-based case-control study. *Am J Epidemiol* 1994;**140**:244–255.
- MacKenzie JS, Houghton M. Influenza infections during pregnancy: association with congenital malformations and with subsequent neoplasms in children, and potential hazards of live virus vaccines. *Bacterial Rev* 1974;**38**:356–370.
- Mak TK, Mangtani P, Leese J, Watson JM, Pfeifer D. Influenza vaccination in pregnancy: current evidence and selected national policies. *Lancet Infect Dis* 2008;**8**:44–52.
- McGregor JA, Burns JC, Levin MJ, Burlington B, Meiklejohn G. Transplacental passage of influenza A/Bangkok (H3N2) mimicking amniotic fluid infection syndrome. *Am J Obstet Gynecol* 1984;**149**:856–859.
- Mereckiene J, Cotter S, D'Ancona F, Giambi C, Nicoll A, Levy-Bruhl D, Lopalco PL, Weber JT, Johansen K, Dematte L et al. Differences in national influenza vaccination policies across the European Union, Norway and Iceland 2008–2009. *Euro Surveill* 2010;**15**:19700.
- Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Br Med J* 2009;**339**:b2535.
- Monto AS, Gravenstein S, Elliott M, Colopy M, Schweinle J. Clinical signs and symptoms predicting influenza infection. *Arch Intern Med* 2000;**160**:3243–3247.
- Moretti ME, Bar-Oz B, Fried S, Koren G. Maternal hyperthermia and the risk for neural tube defects in offspring: systematic review and meta-analysis. *Epidemiology* 2005;**16**:216–219.
- Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. *Lancet* 2009;**374**:1773–1785.
- Neuzil KM, Reed GW, Mitchel EF, Simonsen L, Griffin MR. Impact of influenza on acute cardiopulmonary hospitalizations in pregnant women. *Am J Epidemiol* 1998;**148**:1094–1102.
- Oster ME, Riehle-Colarusso T, Alverson CJ, Correa A. Associations between maternal fever and influenza and congenital heart defects. *J Pediatr* 2011;**158**:990–995.
- Pasternak B, Svanstrom H, Molgaard-Nielsen D, Krause TG, Emborg HD, Melbye M, Hviid A. Risk of adverse fetal outcomes following administration of a pandemic influenza A(H1N1) vaccine during pregnancy. *J Am Med Assoc* 2012;**308**:165–174.
- Pleydell MJ. Anencephaly and other congenital abnormalities. An epidemiological study in Northamptonshire. *Br Med J* 1960;**1**:309–315.
- Record RG. Anencephalus in Scotland. *Br J Prev Soc Med* 1961;**15**:93–105.
- Saxen I. The association between maternal influenza, drug consumption and oral clefts. *Acta Odontol Scand* 1975a;**33**:259–267.
- Saxen I. Epidemiology of cleft lip and palate. An attempt to rule out chance correlations. *Br J Prev Soc Med* 1975b;**29**:103–110.
- Saxen L, Hjelt L, Sjostedt JE, Hakosalo J, Hakosalo H. Asian influenza during pregnancy and congenital malformations. *Acta Pathol Microbiol Scand* 1960;**49**:114–126.
- Saxen L, Holmberg PC, Kurppa K, Kuosma E, Pyhala R. Influenza epidemics and anencephaly. *Am J Public Health* 1990;**80**:473–475.
- Siston AM, Rasmussen SA, Honein MA, Fry AM, Seib K, Callaghan WM, Louie J, Doyle TJ, Crockett M, Lynfield R et al. Pandemic 2009 influenza A(H1N1) virus illness among pregnant women in the United States. *J Am Med Assoc* 2010;**303**:1517–1525.
- Smith MS, Upfold JB, Edwards MJ, Shiota K, Cawdell-Smith J. The induction of neural tube defects by maternal hyperthermia: a comparison of the guinea-pig and human. *Neuropathol Appl Neurobiol* 1992;**18**:71–80.
- Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *J Am Med Assoc* 2000;**283**:2008–2012.
- Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;**23**:1351–1375.
- Walker WM, McKee AP. Asian influenza in pregnancy; relationship to fetal anomalies. *Obstet Gynecol* 1959;**13**:394–398.
- Warrell MJ, Tobin JO, Wald NJ. Examination for influenza IgA and IgM antibodies in pregnancies associated with fetal neural-tube defects. *J Med Microbiol* 1981;**14**:159–162.
- Wells GA, Shea B, O'Connell D. The Newcastle–Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2004, <http://www.medicine.mcgill.ca/rtamblyn/Readings/The%20Newcastle%20-%20Scale%20for%20assessing%20the%20quality%20of%20non%20randomised%20studies%20in%20meta-analyses.pdf>.

- Wilson MG, Stein AM. Teratogenic effects of Asian influenza. An extended study. *J Am Med Assoc* 1969;**210**:336–337.
- Wilson MG, Heins HL, Imagawa DT, Adams JM. Teratogenic effects of Asian influenza. *J Am Med Assoc* 1959;**171**:638–641.
- Zambon M, Hays J, Webster A, Newman R, Keene O. Diagnosis of influenza in the community: relationship of clinical diagnosis to confirmed virological, serologic, or molecular detection of influenza. *Arch Intern Med* 2001;**161**:2116–2122.