



SYSTEMATIC REVIEW

Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis [version 1; referees: 2 approved]

Carl Heneghan¹, Jeffrey K. Aronson¹, Elizabeth Spencer¹, Bennett Holman², Kamal R. Mahtani¹, Rafael Perera¹, Igho Onakpoya¹

¹Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

²Underwood International College, Yonsei University, Seoul, South Korea

v1 **First published:** 31 Oct 2018, 7:1725 (<https://doi.org/10.12688/f1000research.16758.1>)
Latest published: 31 Oct 2018, 7:1725 (<https://doi.org/10.12688/f1000research.16758.1>)

Abstract

Background: Oral hormone pregnancy tests (HPTs), such as Primodos, containing ethinylestradiol and high doses of norethisterone, were given to over a million women from 1958 to 1978, when Primodos was withdrawn from the market because of concerns about possible teratogenicity. We aimed to study the association between maternal exposure to oral HPTs and congenital malformations.

Methods: We have performed a systematic review and meta-analysis of case-control and cohort studies that included data from pregnant women and were exposed to oral HPTs within the estimated first three months of pregnancy, if compared with a relevant control group. We used random-effects meta-analysis and assessed the quality of each study using the Newcastle–Ottawa Scale for non-randomized studies.

Results: We found 16 case control studies and 10 prospective cohort studies, together including 71 330 women, of whom 4209 were exposed to HPTs. Exposure to oral HPTs was associated with a 40% increased risk of all congenital malformations: pooled odds ratio (OR) = 1.40 (95% CI 1.18 to 1.66; P<0.0001; I² = 0%). Exposure to HPTs was associated with an increased risk of congenital heart malformations: pooled OR = 1.89 (95% CI 1.32 to 2.72; P = 0.0006; I²=0%); nervous system malformations OR = 2.98 (95% CI 1.32 to 6.76; P = 0.0109 I² = 78%); gastrointestinal malformations, OR = 4.50 (95% CI 0.63 to 32.20; P = 0.13; I² = 54%); musculoskeletal malformations, OR = 2.24 (95% CI 1.23 to 4.08; P= 0.009; I² = 0%); the VACTERL syndrome (Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies, and Limb defects), OR = 7.47 (95% CI 2.92 to 19.07; P < 0.0001; I² = 0%).

Conclusions: This systematic review and meta-analysis shows that use of oral HPTs in pregnancy is associated with increased risks of congenital malformations.

Keywords

pregnancy; congenital malformations, hormones

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 1 published 31 Oct 2018	 report	 report

- 1 **David Healy**, Bangor University, UK
- 2 **Jesse Olszynko-Gryn** , University of Cambridge, UK
- Cyrille Jean**, Sciences Po, France

Discuss this article

Comments (3)

Corresponding author: Carl Heneghan (carl.heneghan@phc.ox.ac.uk)

Author roles: **Heneghan C:** Conceptualization, Formal Analysis, Investigation, Methodology; **Aronson JK:** Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Spencer E:** Formal Analysis, Investigation; **Holman B:** Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Mahtani KR:** Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Perera R:** Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Onakpoya I:** Conceptualization, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: CH, ES, and IO received funding from the NIHR SPCR Evidence Synthesis Working group (NIHR SPCR ESWG). CH is Director of the ESWG and has received financial remuneration from an asbestos legal case. CH is also funded by the NIHR Oxford BRC, is an NIHR Senior Investigator and Editor in Chief of BMJ Evidence-Based Medicine. JKA has published papers and edited textbooks on adverse drug reactions; he has also acted as an expert witness in cases related to adverse drug reactions.

Grant information: The Evidence Synthesis Working Group is funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR) [Project Number 390]. The views expressed are those of the author(s) and not necessarily those of the NIHR, the NHS or the Department of Health.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Heneghan C *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Heneghan C, Aronson JK, Spencer E *et al.* **Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis [version 1; referees: 2 approved]** *F1000Research* 2018, 7:1725 (<https://doi.org/10.12688/f1000research.16758.1>)

First published: 31 Oct 2018, 7:1725 (<https://doi.org/10.12688/f1000research.16758.1>)

Introduction

Oral hormone pregnancy tests (HPTs), such as Primodos (known as Duogynon in Germany), were used from 1958 to 1978, before urine pregnancy tests were available¹. Oral HPTs contained ethinylestradiol and large doses of norethisterone (synthetic forms of estrogen and progesterone respectively), the latter in much larger amounts than those included in current combined oral contraceptives (see [Table 1](#)). The test principle was that menstruation would be induced in those who were not pregnant.

In the UK more than a million women took HPTs². However, evidence that they should not be used in pregnant women because of a risk of fetal malformations³ led the then Committee on Safety of Medicines in 1975 to conclude that a warning should be added to the Data Sheets, stating that HPTs should not be taken during pregnancy. ([Supplementary File 1](#)) In 1978, the manufacturers of Primodos, Schering AG (taken over by Bayer AG in 2008), voluntarily stopped marketing the product.

Since Primodos was withdrawn, the discovery of previously confidential documents has led to renewed concerns about its potential to cause harm. In 2014, therefore, the Medicines and Healthcare products Regulatory Agency (MHRA) initiated a review, which was published in 2017 and reported that the evidence was insufficient, mixed, and too heterogeneous to support an association between oral HPTs and congenital malformations³.

To date, there has been no systematic review and meta-analysis of oral HPTs, using all the available data, to assess the likelihood of an association. We have therefore performed a systematic review to obtain all relevant data on hormone pregnancy tests and congenital malformations, used meta-analytical tools to obtain summary estimates of the likelihood of an association, and assessed the potential biases in these estimates.

Methods

Data sources

Full details of our search strategy are provided in [Supplementary File 2](#). We searched [Medline](#), [Embase](#), and [Web of Science](#)

(which yielded German papers and conference abstracts) and searched for regulatory documents online, including the UK Government's "Report of the Commission on Human Medicines' Expert Working Group on Hormone Pregnancy Tests", which includes the original Landesarchiv Berlin Files¹, and reference lists of retrieved studies from the start of the databases in 1946 to 20 February 2018.

We used the following search terms without date limits or language restrictions: (Primodos OR Duogynon OR "hormone pregnancy test" OR "sex hormones" OR "hormone administration" OR "norethisterone" OR "ethinylestradiol") AND pregnancy AND (congenital OR malformations OR anomalies). Several comparable high-dose HPTs were available at the same time as Primodos; we performed additional searches for evidence relating to these (See [Supplementary File 3](#) for List of HPTs included in evidence search).

Study selection

We included observational studies of women who were or became pregnant during the study and were exposed to oral HPTs within the estimated first three months of pregnancy and compared them with a relevant control group. When a study was described in more than one publication, we chose the publication that contained the most comprehensive data as the primary publication. We excluded studies where the intervention was oral hormones taken for other reasons (e.g., oral contraception) and it was not possible to extract data on hormone pregnancy tests. We did not restrict the language of publication. We checked additional relevant data and extracted them from the secondary publications when necessary.

Data extraction and risk of bias assessment

Two reviewers (CH and ES) applied inclusion and quality assessment criteria, compared results, and resolved discrepancies through discussion with the other authors. We used a review template to extract data on study type, numbers of pregnancies exposed and not exposed to oral HPTs, and types and numbers of outcomes. Where available, we extracted data about the women studied, including ascertainment of cases, age, parity, setting,

Table 1. Doses of ethinylestradiol and norethisterone in various formulations of contraceptive steroids, ordered by increasing dose of norethisterone.

Indication (oral formulation)	Ethinylestradiol dose	Norethisterone acetate dose
Progestogen-only contraception*	-	350 micrograms
Combined oral contraceptive (Loestrin-20)	20 micrograms	1000 micrograms
Combined oral contraceptive (Norimin)	35 micrograms	1000 micrograms
Biphasic combined oral contraceptive (BiNovum)	35 micrograms	500/1000 micrograms
Triphasic combined oral contraceptive (Synphase)	35 micrograms	500/1000/500 micrograms
Combined oral contraceptive (Loestrin-30)	30 micrograms	1500 micrograms
Oral hormone pregnancy test (Primodos)	20 micrograms	10 milligrams
In endometriosis, dysmenorrhoea, dysfunctional uterine bleeding, and menorrhagia, or to delay menstruation*	-	10–15 milligrams/day
Breast cancer*	-	40 milligrams/day

*Unbranded

exposure to other medications, and confounding variables. In case-control studies, if data were reported on more than one control group, we extracted data where possible for non-disease/non-abnormality controls, and combined control groups if necessary.

The primary outcome of interest was all major congenital malformations. We also categorized outcomes for the congenital anomaly in the offspring at any time into congenital cardiac, gastrointestinal, musculoskeletal, nervous system, and urogenital defects, and the VACTERL syndrome (Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies, and Limb defects).

We assessed quality using the Newcastle–Ottawa Scale (NOS) for non-randomized studies included in systematic reviews⁶. The scale assesses the selection of study groups (cases and controls), comparability of study groups, including cases and controls, and ascertainment of the outcome/exposure. Each positive criterion scores 1 point, except comparability, which scores up to 2 points. The maximum NOS score is 9, and we interpreted a score of 1 to 3 points as indicating a high risk of bias⁷. To determine whether the study had controlled for the most important factors, we selected the items reported in the original paper and resolved disagreements through consensus, using a third author (IO). We examined whether there was a linear relation between methodological quality and study results, by plotting the odds ratios against the NOS scores, using Excel, and assessed the correlations of NOS scores with several confounding variables we collected⁸.

Data synthesis and statistical methods

We calculated study-specific odds ratios for outcomes and associated confidence intervals. We meta-analysed the data using a random-effects model. We assessed heterogeneity across studies using the I^2 statistic and publication bias using funnel plots⁹. We performed a sensitivity analysis by removing single studies to judge the stability of the effect and to explore the effect on heterogeneity¹⁰, and we described any sources of variation. We also judged robustness by removing studies of low quality from the analysis. To examine whether the observed heterogeneity could be explained by differences in the NOS score, we also performed meta-regression using the NOS score as the covariate against the log OR as weights for traditional meta-regression using Stata version 14.

We planned subgroup analyses for the timing of administration of HPTs in relation to pregnancy and organogenesis and study design (case-control versus cohort) using Cochran's Q test. We used RevMan v.5.3 for all analyses, except for meta-regression, for which we used Stata version 14. RevMan and Stata estimate the effects of trials with zero events in one arm by adding a correction factor of 0.5 to each arm (trials with zero events in both arms are omitted). We performed a sensitivity analysis by removing studies with zero events from the analyses.

We followed the reporting guidelines of the Meta-Analysis of Observational Studies in Epidemiology (MOOSE). A completed checklist is available as [Supplementary File 4](#)¹¹

Patient involvement

Members of the Association for Children Damaged by HPTs were involved in the original discussions of this review and provided input to the outcome choices, the search, the location of study articles, and translations. We plan to present the study findings to relevant patient groups and make available lay interpretations.

Results

Description of included studies

We retrieved 409 items for screening. After title and abstract screening and removal of duplicates ($n = 18$), we excluded 354 records as not being relevant to the aim of the review. We assessed the full texts of 37 articles and identified 24 articles for inclusion. [Figure 1](#) shows the PRISMA flow diagram for the inclusion of studies.

The 24 included articles reported on 26 studies (16 case-control studies and ten prospective cohort studies); one article [Nora 78] included two case-control studies and one prospective study. We found no randomized controlled trials. Of these articles, two were unpublished reports (see [Supplementary File 5](#) for full references). The studies included 71 330 women. The case-control studies included 28 761 mothers, 594 of whom were exposed to HPTs; the cohort studies included 42 569 mothers and 3615 exposures to HPTs. The studies were published between 1972 and 2014, and all were performed either in Europe or the USA. They mostly recruited women and their infants at maternity centres or hospital paediatrics wards.

The choices of controls in the case-control studies varied; they included, at one extreme, healthy infants born on a date close to the case infants and, at the other extreme, infants with malformations other than those under investigation. Among the prospective cohort studies, the populations tended to be women recruited at antenatal clinics or birth centres (See [Table 2](#). Characteristics of included studies).

Quality assessment of included studies

Of the 26 included studies, three were assigned a NOS score of 3 or below and were therefore judged as being at high risk of bias. One was a case-control study (Laurence 1971, a published abstract as a letter) and two were cohort studies (Fleming 1978 and Haller 1974, both unpublished). The NOS scores ranged from 2 to 9 (median 5). Twelve of the 26 included studies scored 7 to 9 and were judged to be at low risk of bias (see [Table 3](#) of NOS scores in the data files). Item 5 of the NOS score addresses comparability of cases and controls based on design or analysis. Of the 16 case control studies, 12 controlled for the most important factor (item 5a) and nine controlled for important additional factors (item 5b). Of the ten cohort studies, six controlled for the most important factor (item 5a) and four controlled for important additional factors (item 5b). The mean Newcastle–Ottawa scale score was 6.1, indicating an overall moderate risk of bias. [Table 2](#) also shows that seven studies did not report the confounding variables collected (Laurence 1971; Levy 1973; Tummler 2014; Fleming 1978; Haller 1974; Moire 1978; Rousel 1968). NOS scores correlated with the increasing number of confounding variables collected ($r = 0.83$).

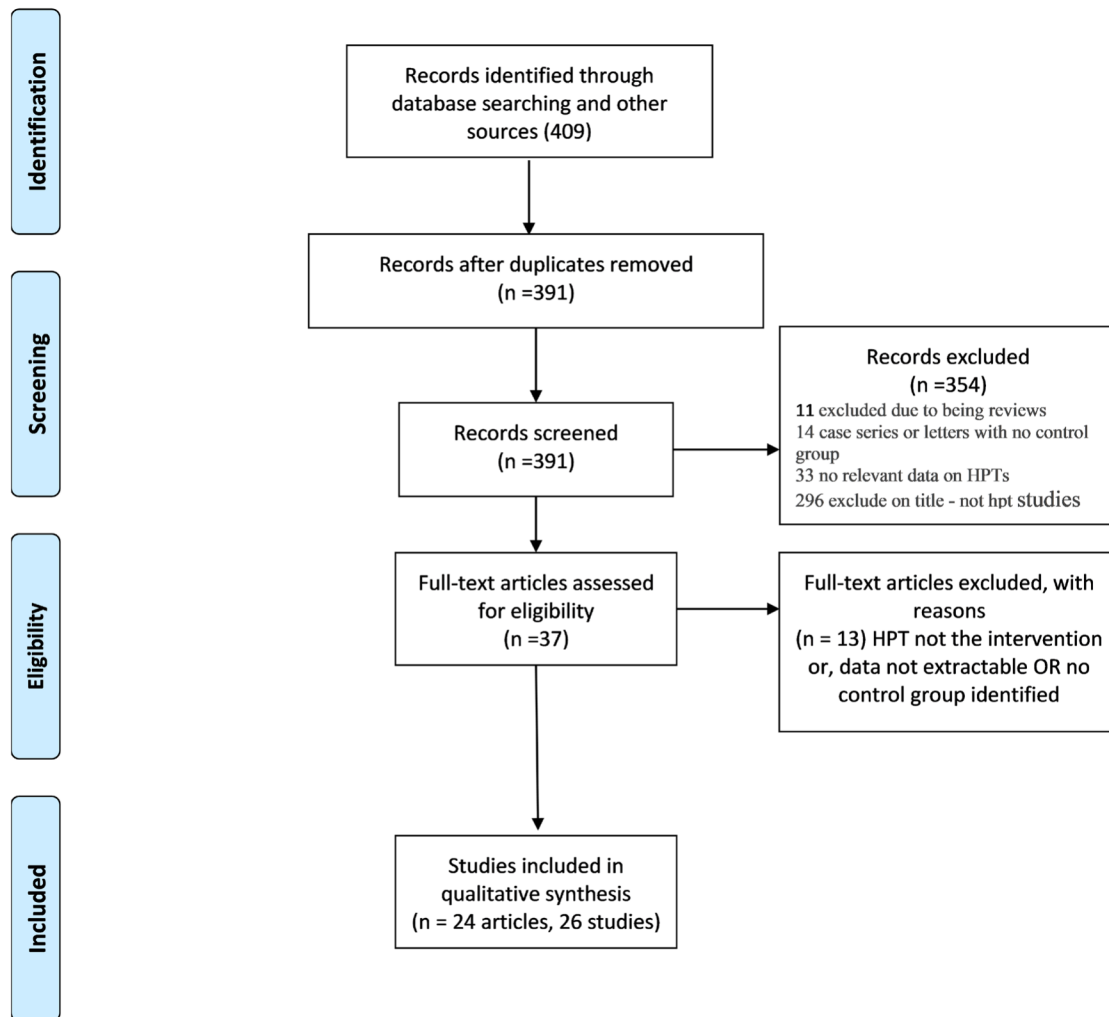


Figure 1. Study flow diagram showing inclusion of relevant studies.

Supplementary File 6 shows the funnel plots for all congenital malformations and congenital heart disease; because of inadequate numbers of included studies, we did not use more advanced statistical methods to assess publication bias.

Association of exposure to HPT with the risks of malformations

Nine studies, including 61 642 mothers of infants and 3274 exposed to HPTs, examined the association in pregnancy with all congenital malformations. Two were case-control studies (Greenberg 1977; Sainz 1987) and seven were cohort studies (Fleming 1987; Goujard 1979; Haller 1974; Kullander 1976; Michaelis 1983; Rumeau-Rouquette 1978; Torfs 1981) (Figure 2). Exposure to oral HPTs was associated with a 37% increased risk of all congenital malformations: pooled odds ratio (OR) = 1.40 (95% CI 1.18 to 1.66; $P < 0.0001$; $I^2 = 0\%$). For the two case-control studies only, pooled OR = 1.70 (95% CI 1.01 to 2.86; $P = 0.04$; $I^2 = 63\%$) and for the seven cohort studies, pooled OR = 1.28 (95% CI 1.05 to 1.56; $P = 0.02$; $I^2 = 0\%$). The test for subgroup differences was not significant ($P = 0.32$). In a post-hoc

sensitivity analysis, removing the studies that collected no confounding variables (Haller 74 and Fleming 78, both of low quality) did not affect the significance of the result (OR 1.44; 95% CI 1.18 to 1.75; $P = 0.0004$, $I^2 = 11\%$). The meta-regression showed no association between total NOS score and increased risk ($P = 0.51$).

Seven studies, including 19 267 mothers of infants and 218 exposed to oral HPTs, analysed congenital heart malformations. Five were case-control studies (Ferencz 1980; Janerich 1977; Levy 1973; Nora 1978-2/3) and two were cohort studies (Hadjigeorgiou 1982; Torfs 1981) (Figure 3). The pooled relative OR = 1.89 (95% CI 1.32 to 2.72; $P = 0.0006$; $I^2 = 0\%$).

In a post-hoc sensitivity analysis, removing one study that collected no confounding variables (Levy 73, a low-quality study) did not affect the significance of the result (OR = 1.88; 95% CI 1.25 to 2.85; $P = 0.003$, $I^2 = 12\%$) For the five case-control studies only, the pooled OR = 1.87 (95% CI 1.23 to 2.85; $P = 0.004$; $I^2 = 9\%$); for the two cohort studies the pooled

Table 2. Characteristics of included studies.

Study ID	Study population	Setting	Confounding variables collected	Information on controls including matching criteria in case-control studies	Outcomes reported
Case-control studies					
Ferencz 1980	Mothers of 110 infants with conotruncal abnormalities of the heart, born 1972–75.	Hospitals served by the Maryland State Intensive Care Neonatal Program, USA.	Maternal health (hospitalisations, illnesses, treatments); past reproductive history; index pregnancy factors including contraception used previously, fertility treatments, symptoms, illnesses and medications during pregnancy including hormones; smoking; alcohol intake; occupational history of mother and father; exposure of mother to fumes, paints and insecticides; family history including history of congenital abnormalities in previous children or in close relatives.	For each case, three unaffected controls were chosen from the birth population: two matched on eight characteristics related to the likelihood of hormone-taking (race, maternal age, parity, foetal losses, gestational age, delivery mode, time of prenatal registration, private service), and one also on the infant's sex and birthweight; the third control was chosen at random.	Congenital heart disease (Conotruncal malformations of the heart)
Gal 1972	100 mothers of infants with spina bifida, and controls	Hospital in London, UK, for cases; unclear where controls were recruited from	Age, parity, reproductive history, illnesses, illegitimacy, bleeding	Controls matched for week of baby's birth; age of mother (5-year bands), reproductive history, course of pregnancy, sex of baby.	Spina bifida
Greenberg 1977	Cases identified via OPCS and matched controls identified from general practices of the cases.	General practices in the UK	Antenatal, personal, and family history and drugs prescribed during the first trimester.	Controls: babies born within 3 months of and based at the same general practice as matched cases. Antenatal, personal, and family history and drugs prescribed during the first trimester.	Neural tube defects, oral clefts, limb malformations and other non-minor abnormalities
Janerich 1974	108 cases of congenital limb defects and 108 unaffected controls	New York State, USA	Age, parity, race	Controls matched on birth date, mother's race and age +/- 2 years; and by default, due to adjacent records for cases and controls these matched well on county of residence of the mothers.	Congenital limb defects
Janerich 1977	104 cases with birth certificate mentioning CHD, 104 matched controls	New York State, USA	Age, country of residence, date of birth, race, medications, infections	From adjacent birth record matched by mother's age, county of residence, date of birth, race	Congenital heart disease
Lammer 1986	1,091 mothers of infants with abnormalities born 1 July 1990 to 20 June 1979, (21% not completed data collection)	Population register	Race, maternal education, family history, socio-economic status, parity, previous foetal loss	Control group was composed of infants with malformations other than the one under investigation. e.g. for spina bifida, controls were those with non-spina bifida abnormalities.	Major malformations, including anencephaly, spina bifida, cleft lip, cleft palate, Down syndrome, oesophageal atresia, small bowel atresia, rectal anal atresia, anterior abdominal wall defects, diaphragmatic hernia, limb reduction.

Study ID	Study population	Setting	Confounding variables collected	Information on controls including matching criteria in case-control studies	Outcomes reported
Case-control studies					
Laurence 1971	1968-1970, UK	3 hospital birth centres in the UK	Non-reported	In London the controls were the next baby with no abnormality born in the same hospital; in Exeter, control mothers were matched for area of birth, parity and month of conception; in Wales the control mothers were those who had had one baby with spina bifida or anencephaly and had a subsequent unaffected birth during the study period; these last were not matched individually.	Spina bifida and anencephaly
Levy 1973	76 cases, 76 controls	Hospital, Montreal, Canada	Non-reported	Controls were infants with Mendelian disorders, matched for date of birth.	Congenital heart defects (transposition of the great vessels)
Nora 1975	15 patients with multiple congenital anomalies, 30 controls (15 with chromosomal anomalies, 15 with functional heart murmurs)	University of Colorado Medical Center, Denver, and affiliated hospitals, USA	Age, race, socioeconomic status, area of residence	Matched for age. 15 controls had chromosomal abnormalities, 15 had functional heart murmurs	VACTERL
Nora 1978 case control 1	32 patients with VACTERL, 60 controls	Hospital	Age, date of birth, sex, gestational age, race, socioeconomic levels, areas of residences, parity	Matched as closely as possible for age, date of birth, sex, gestational age, race, socioeconomic levels, areas of residences, parity	VACTERL
Nora 1978 case control 2 and 3	236 patients with full variety of cardiac lesions, 412 controls with known single mutant gene and chromosomal disorders	Hospital	Sex, race, approximate date of birth, area of residence	Matching was for sex, race, approximate date of birth, area of residence	Congenital heart disease (congenital heart lesions)
Polednak 1983	99 singleton male births with hypospadias and 99 matched controls	New York State, USA	Parity, maternal age, race, area of residence	Most adjacent birth date, matched for maternal age, race, area of residence	Hypospadias
Rothman 1979	390 cases, 1,254 controls. HPTS: 14/388 cases vs 35/1246 controls	State care service for congenital heart disease	Parity, mother's education level, insulin use, alcohol, tobacco	Controls: births within same 3 years of the study period; 1,254 respondents from contacts to births selected randomly from the birth register.	Congenital heart disease
Sainz 1987	244 cases identified via the national collaboration of 42 hospitals registering congenital abnormalities between April 1976 to Sept 1984	Spanish register of congenital abnormalities within 42 participating hospitals	Sex, data and place of birth	Controls: unaffected births at same hospital, matched on sex, date of birth.	Spina bifida and anencephaly

Study ID	Study population	Setting	Confounding variables collected	Information on controls including matching criteria in case-control studies	Outcomes reported
Case-control studies					
Tummler 2014	296 cases, 3,676 infants with abnormalities	Data from the Malformation Monitoring Centre Saxony-Anhalt, Germany.	Non-reported	No information on matching	Congenital bladder exstrophy
Cohort studies					
Fleming 1978	RCGP Outcomes of Pregnancy study 1975: 9,000 women; from this was selected a random sample of 500 pregnancies proceeding to normal outcomes	General practices, UK	Non-reported		Any malformation
Goujard 1979	3,379 women pregnant and attending gynaecology clinics between 1975 to 1977	Obstetrics and gynaecology centres, Paris and Lille, France	Information on current pregnancies including symptoms and medications taken; previous pregnancies and general health backgrounds.		Congenital malformations, also congenital heart defects, skeletal anomalies, microencephaly.
Hadjiigeorgiou 1982	Retrospective cohort, Alexandra Maternity Hospital Greece; births 1975-77. 15,535 live births, 559 exposed to sex hormones of which 112 (20%) exposed to HPTs, 14,976 no hormones, congenital heart disease studied confirmed by cardiologist & lab tests. Diseases and medication reported at admission prior to birth.	Hospital birth centre	Cytomegalovirus, infection, toxoplasmosis, hepatitis, syphilis, rubella, teratogenic drugs		Congenital heart disease
Haller 1974	3588 pregnant women, recruited Oct 1969 to April 1972, University Hospital Göttingen; 617 (17.2 %) with hormonal pregnancy test	Hospital birth centre	Non-reported		Congenital malformations
Kullander 1976	6,376 pregnancies, Malmö, 1963-5, resulting in 5,753 live births, 5,002/5,753 no abnormality, 751/5,753 with abnormality. 156 women took Primodos.	Sweden	Major and minor disease; the woman's age, parity, marital status, and social class. Birth weight, placental weight.		Major and minor malformations

Study ID	Study population	Setting	Confounding variables collected	Information on controls including matching criteria in case-control studies	Outcomes reported
Meire 1978	500 mothers consecutive births in 3 hospitals in Bruges, Belgium, 20 had taken HPTs.	Hospital birth centres	Non-reported		Oesophageal atresia
Michaelis 1983	13,643 pregnancies	Antenatal clinics, Germany	Detailed general and gynaecological history, drug intake, exposure to chemical agents, daily workload, intercurrent diseases, accidents, surgical operations and other factors. NR		Major malformations
Roussel 1968	Pregnancies 1966 to 1967	General practices, UK			Central nervous system malformations including anencephaly, hydrocephaly, microcephaly, meningomyelocele, myelocele, spina bifida.
Rumeau-Rouquette 1978	1963-69, recruitment in 12 gynaecology clinics in Paris; 12,764 women gave birth to 12,895 children in hospitals participating in study; controls were mothers of unaffected infants selected at random among women questioned in same hospital	Hospital birth centres, France	Medical history, course of pregnancy, infectious diseases, inoculations, reproductive history, social and occupational category, use of alcohol, tobacco		Congenital malformations
Torfs 1981	19,906 full term pregnancies, 227 of which exposed to HPTs.	Hospital birth centre	Age, medical and reproductive history, socio-economic information, ethnicity		Severe congenital anomalies including congenital heart defects, neuroblastoma, cleft lip and limb reduction; non-severe congenital anomalies e.g. hypospadias of the first degree, congenital dis-location of the hip, polydactyly.

Table 3. Newcastle-Ottawa scale scores for included studies.

Newcastle-Ottawa scale case-control studies										
Study ID	Is the case definition adequate?	Are the cases representative?	Selection of controls adequate	Definition of controls adequate	Comparability of cases and controls on the basis of the design or analysis		Ascertainment of exposure adequate	Same method of ascertainment for cases and controls	Non-response rate adequate	Total score /9
					a) Study controls for the most important factor	b) Study controls for important additional factors				
Ferencz 1980	yes	yes	yes	yes	yes	yes	yes	yes	yes	9
Gal 1972	unclear	unclear	yes	yes	yes	yes	yes	yes	unclear	6
Greenberg 1977	yes	yes	yes	yes	yes	yes	unclear	yes	yes	8
Janerich 1974	no	yes	yes	no	yes	no	yes	yes	unclear	5
Janerich 1977	yes	yes	unclear	yes	yes	yes	yes	unclear	yes	7
Hellstrom 1976	yes	unclear	no	yes	yes	no	unclear	yes	unclear	4
Lammer 1986	yes	yes	unclear	yes	yes	yes	yes	yes	no	7
Laurence 1971	yes	unclear	unclear	yes	no	unclear	unclear	unclear	unclear	2
Levy 1973	yes	yes	no	yes	yes	no	unclear	unclear	unclear	4
Nora 1975	yes	yes	yes	yes	yes	no	yes	yes	yes	8
Nora 1978 case control 1	yes	yes	yes	no	yes	yes	yes	yes	yes	8
Nora 1978 case control 2 and 3	yes	yes	yes	no	yes	yes	yes	yes	unclear	7
Polednak 1983	yes	yes	yes	yes	yes	no	unclear	yes	unclear	6
Rothman 1979	yes	yes	no	no	no	yes	yes	yes	unclear	5
Sainz 1987	unclear	yes	yes	unclear	yes	yes	unclear	yes	unclear	5
Tummier 2014	yes	no	no	no	no	no	no	yes	yes	3

Newcastle–Ottawa scale cohort studies										
	Are the participants representative of the exposed cohort?	Is the non-exposed cohort similar to the exposed?	Is the ascertainment of exposure accurate?	Is there evidence that the outcome of interest was not present at start of study?	Comparability of cases and controls on the basis of the design or analysis		Was the assessment of outcome adequate?	Was follow-up long enough for outcomes to occur?	Was there adequacy of follow up of cohorts?	Total score /9
					a) Study controls for the most important factor	b) Study controls for important additional factors				
Fleming 1978	yes	unclear	unclear	yes	unclear	unclear	unclear	yes	unclear	3
Goujard 1979	yes	yes	unclear	yes	yes	no	unclear	yes	yes	6
Hadjigeorgiou 1982	yes	yes	yes	yes	no	yes	yes	yes	unclear	7
Haller 1974	unclear	unclear	unclear	yes	no	no	unclear	yes	unclear	2
Kullander 1976	yes	yes	yes	yes	yes	unclear	yes	yes	yes	8
Meire 1978	yes	no	yes	yes	yes	no	unclear	yes	yes	6
Michaellis 1983	yes	yes	yes	yes	yes	yes	yes	yes	yes	9
Roussel 1968	yes	yes	yes	yes	no	no	unclear	yes	yes	6
Rumeau-Rouquette 1978	yes	unclear	yes	yes	yes	yes	yes	yes	no	7
Torfs 1981	yes	yes	yes	yes	yes	yes	yes	yes	yes	9

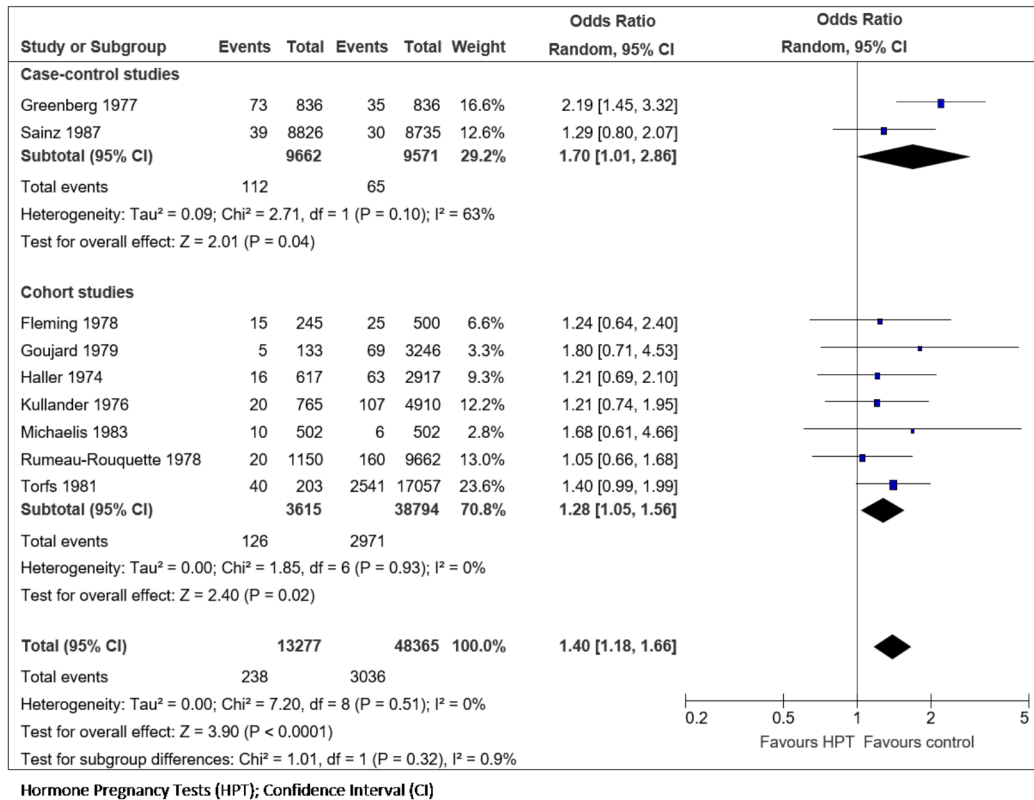


Figure 2. Association of exposure to oral HPTs in pregnancy with all malformations in the offspring.

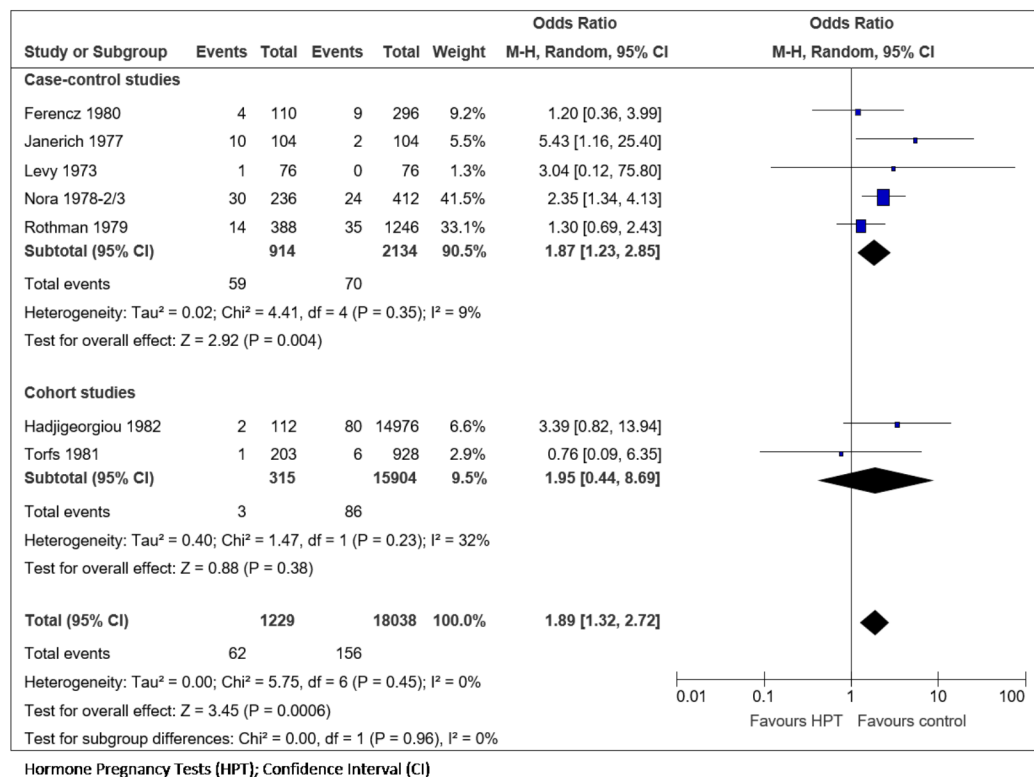


Figure 3. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy with congenital heart disease in the offspring.

OR = 1.95 (95% CI 0.44 to 8.69; P = 0.38; I² = 32%). The meta-regression was not significant (P = 0.94).

For the association between exposure to oral HPTs and nervous system malformations in the offspring, five studies provided data: three case-control studies (Gal 1972; Laurence 1971; Sainz 1987) and two cohort studies (Roussel 1968; Torfs 1981), including 12 486 mothers of infants and 127 exposed (Figure 4). The pooled OR = 2.98 (95% CI 1.32 to 6.76; P = 0.009; I² = 78%). In a post-hoc sensitivity analysis, removing the two studies that collected no confounding variables (Laurence 71; Roussel 68) did not affect the significance of the result and removed the heterogeneity (OR 6.04; 95% CI 3.33 to 10.78; P < 0.00001, I² = 0%).

Gastrointestinal malformations and exposure to oral HPTs were reported in three studies: a case-control study (Lammer 1986) and two cohort studies (Meire 1978 and Torfs 1981), providing data on 2722 mothers of infants, including 79 exposed to HPTs (Figure 5). The pooled OR = 4.50 (95% CI 0.63 to 32.20; P = 0.13; I² = 54%). One case-control study (Polednak 1983) and one cohort study (Torfs 1981) examined the relationship between exposure to oral HPTs in pregnancy and urogenital

malformations: pooled OR = 2.63 (95% CI 0.84 to 8.28; P = 0.10; I² = 0%) (Figure 6).

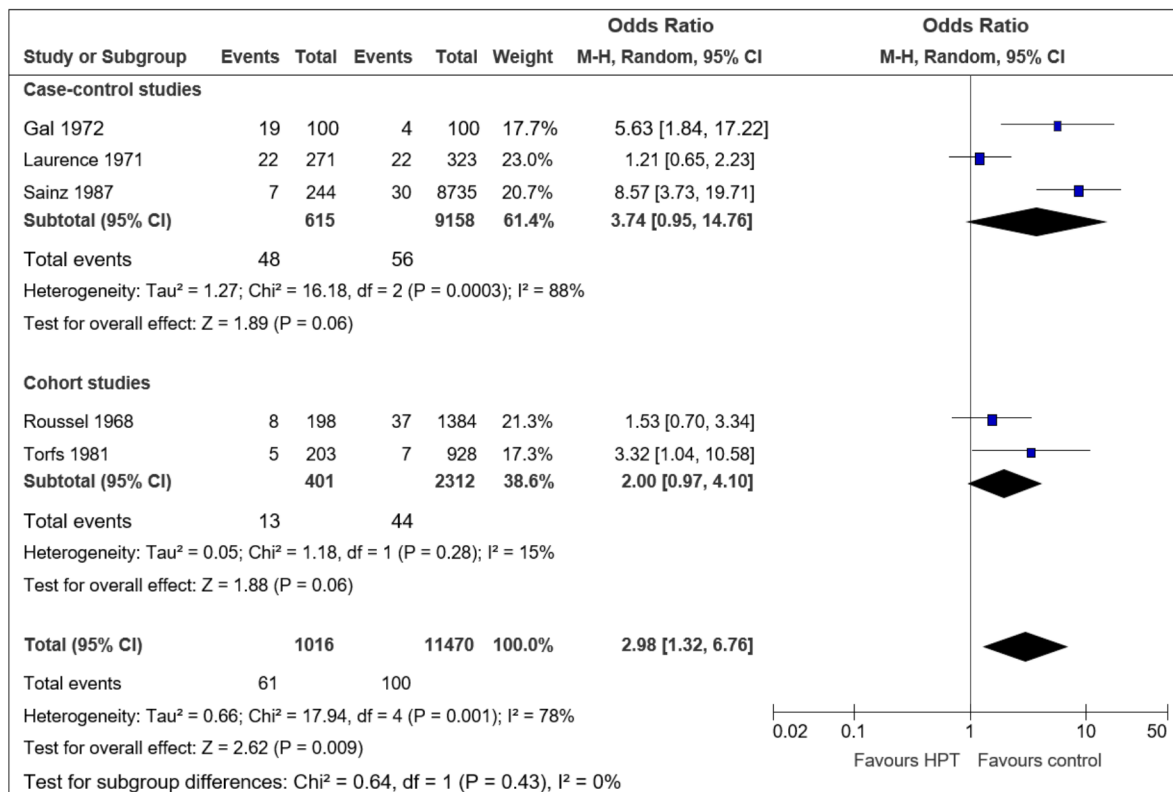
A relation between the exposure to oral HPTs and musculoskeletal malformations was reported in three studies: three case-control studies (Hellstrom 1976; Janerich 1977; Lammer 1986) and one cohort study (Torfs 1981) (Figure 7), based on 2464 women, with 79 exposed to HPTs. The pooled OR = 2.24 (95% CI 1.23 to 4.08; P = 0.009; I² = 0%). Removal of the zero study events (Torfs 1981) did not affect this result. The association of VACTERL with HPT exposure was reported in two case-control studies (Nora 1978-1 and Nora 1975), based on 135 women and infants and 27 exposed to HPTs; the OR was 7.57 (95% CI 2.92 to 19.07; P < 0.0001; I² = 0%) (Figure 8).

Dataset 1. Study extraction sheet

<https://dx.doi.org/10.5256/f1000research.16758.d222937>

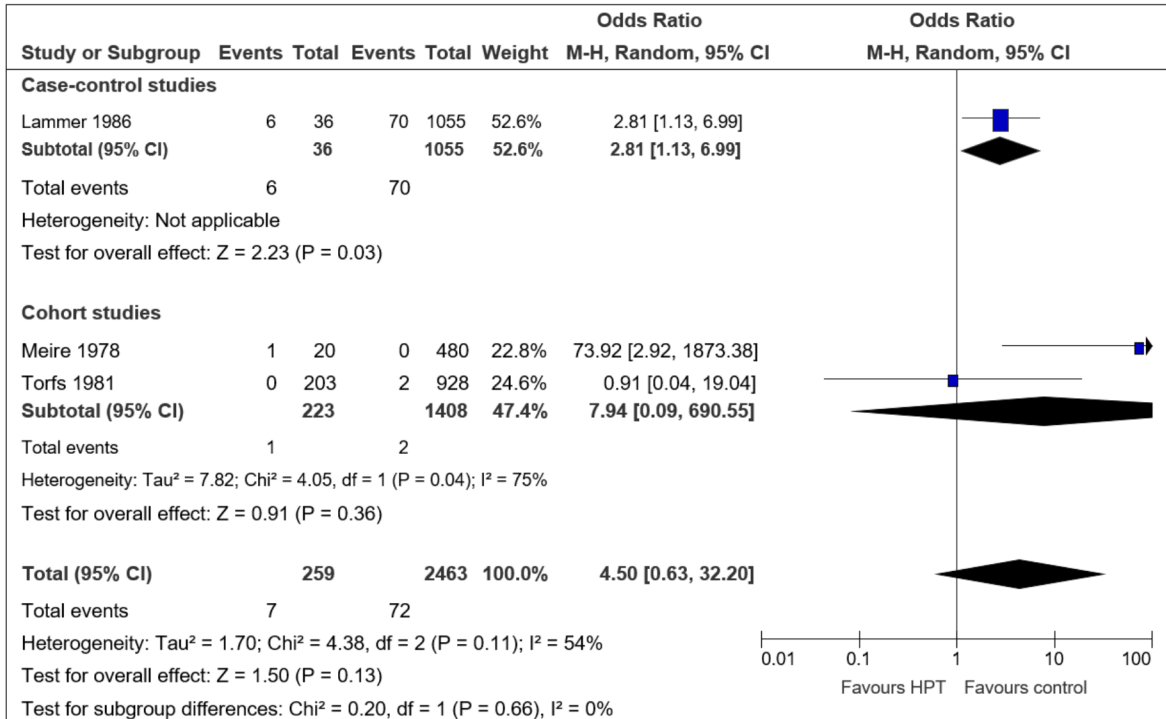
Discussion

We found 24 articles containing 26 studies that reported the association between exposure to oral hormone pregnancy tests in mothers and malformations in their infants: 16 were case-



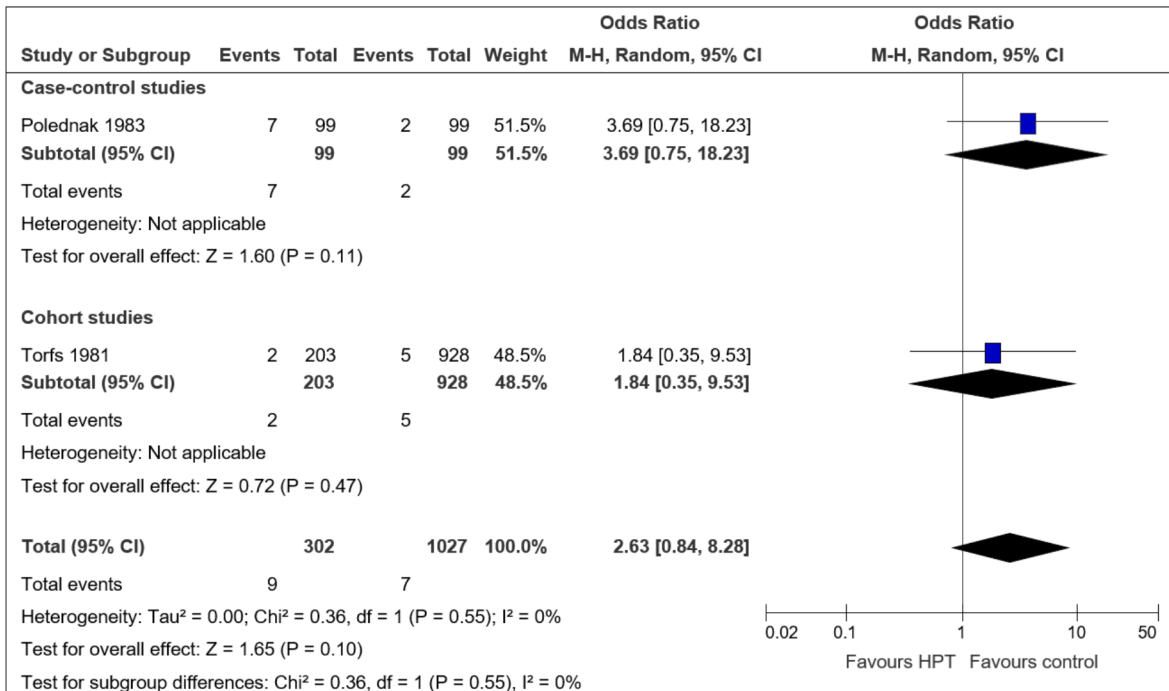
Hormone Pregnancy Tests (HPT); Confidence Interval (CI)

Figure 4. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy and nervous system malformations in the offspring.



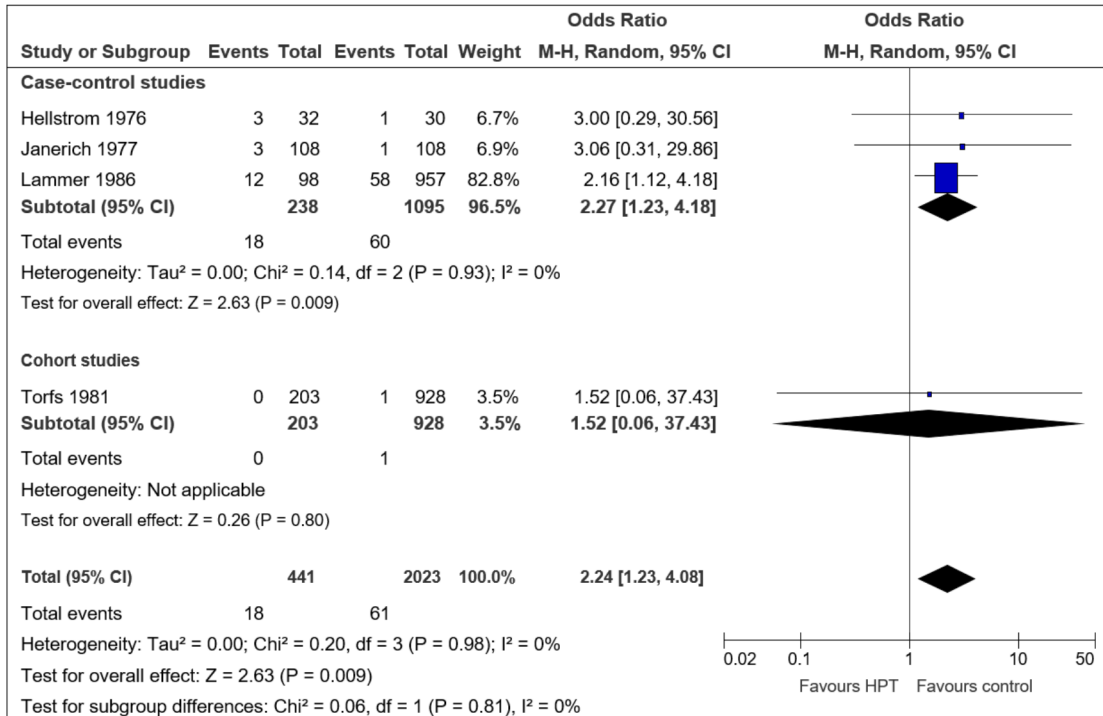
Hormone Pregnancy Tests (HPT); Confidence Interval (CI)

Figure 5. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy and gastrointestinal malformations in the offspring.



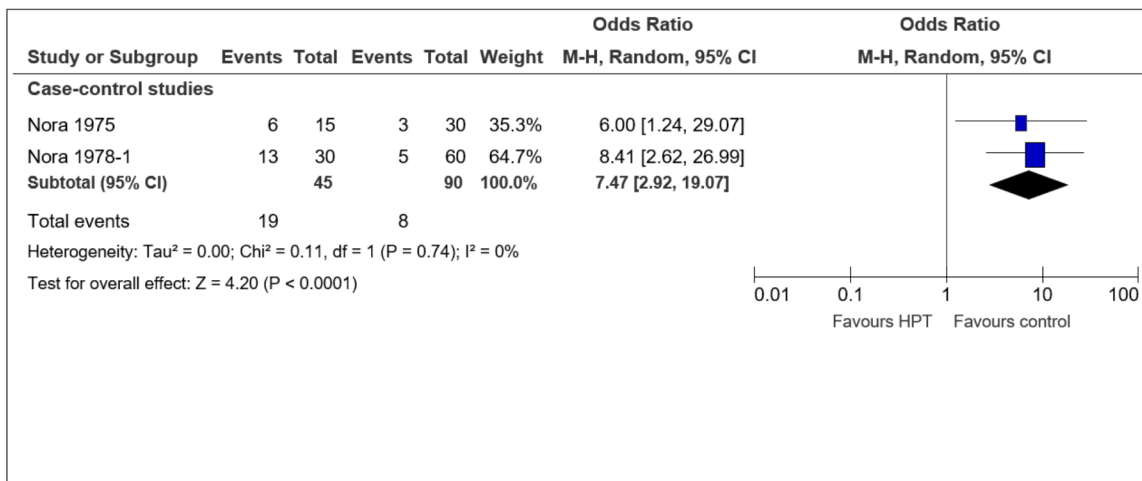
Hormone Pregnancy Tests (HPT); Confidence Interval (CI)

Figure 6. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy and urogenital malformations in the offspring.



Hormone Pregnancy Tests (HPT); Confidence Interval (CI)

Figure 7. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy and musculoskeletal malformations in the offspring.



Hormone Pregnancy Tests (HPT); Confidence Interval (CI)

Figure 8. Association of exposure to oral hormone pregnancy tests (HPTs) in pregnancy with Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies, and Limb defects (VACTERL) syndrome in the offspring.

control studies and ten were prospective cohort studies. The overall quality of the evidence, assessed by the Newcastle–Ottawa Scale, was moderate.

We found significant associations for all congenital malformations pooled and separately for congenital heart malformations, nervous system malformations, musculoskeletal malformations, and the VACTERL syndrome. Many of these pooled analyses had zero heterogeneity, and the direction of effect favoured the controls in 30 of the 32 analyses undertaken (Torfs 81 provided the only effect estimate favouring HPT exposure). The analyses were also robust to sensitivity analyses, and there was no relation between NOS score and increasing risk.

Based on the assumptions that a teratogenic effect of HPTs would be mediated by actions on estrogen and progesterone receptors, and that concentrations of ethinylestradiol and norethisterone in the fetus would be too low to have a significant effect on those receptors, it has been suggested that there is no mechanistic argument for teratogenicity¹. However, other unknown mechanisms might be at play. For example, Isabel Gal first reported concerns of malformations in the children of mothers exposed to HPTs in 1967¹⁴, pointing out that bleeding often occurred in pregnant women soon after exposure and suggesting that that would affect the “equilibrium” of the uterus. Between 5 and 11% of exposed women had bleeding, and the RCGP survey reported induced abortions in about 10% of women¹⁵.

The drugs in Primodos were not tested for animal toxicity and teratogenicity at the time, which, although not unusual, meant that there was a gap in mechanistic understanding. A 2018 study showed that the components in Primodos are associated with dose-dependent and time-related damage in zebrafish embryos, and affect nerve outgrowth and blood vessel patterning in zebrafish^{14,16}. Although it is difficult to compare drug actions between species, and evidence from animal studies is limited, the drugs accumulated in the zebrafish embryos, persisted for some time, and led to rapid embryonic damage^{14,16}. In contrast, other animal studies have shown minimal effects on embryo development¹⁷. There is also evidence that estradiol and progesterone increase the expression of mRNA for isoforms of vascular endothelial growth factor (VEGF) in Ishikawa cells from human endometrial adenocarcinoma¹⁸.

Strengths and weaknesses

Establishing causal associations in the absence of randomization can be difficult. However, the lack of randomized trials in our analysis should not be seen as a hindrance. It would have been unethical to randomize individuals to drugs with known concerns, and randomization, like systematic reviews, was not the norm at the time. Furthermore, for questions about harms, the Oxford CEBM levels of evidence puts systematic reviews of case-control studies on a par with systematic reviews of randomized trials¹⁹.

However, observational methods have limitations²⁰. First, interpretation can be affected by confounding factors. Although most of the studies in this review used matched controls, our analysis was based on raw data from the publications and did not adjust for confounders. Secondly, susceptibility bias can occur,

as women with threatened abortions might be more likely to present and take the medication. Both of these problems can be mitigated by careful matching; 13 of the 16 studies controlled for the most important factor, item 5a on the NOS scale. Thirdly, the severity of malformations studied will have led to differing risk estimates across studies. Fourthly, inappropriate methods of ascertainment of the malformations and exposures could have introduced bias. Finally, incomplete and uneven reporting, along with publication bias (since it is likely that unreported studies exist) could introduce bias and alter the effect estimates.

The use of scoring systems to assess quality has been criticized. However, the NOS scale has been used widely in assessing the quality of non-randomized studies^{21–26}. A NOS score between 0 and 9 has previously been used as a potential moderator in meta-regression²⁷, and has been recommended by the Cochrane Collaboration²⁸. A weakness of the NOS scale is the possible low agreement between assessors²⁹. This was particularly the case when authors had limited experience in doing systematic reviews, but training, even of novices, improves agreement²¹.

The effects were also stable to sensitivity analyses, and changes in NOS score did not affect the risk estimates. The absence of subgroup differences between study designs for the risk estimates supports the robustness of the findings. We also tried to overcome publication bias by translation and assessment of unpublished data. The sample sizes in the studies for all congenital malformations, congenital heart disease, and nervous system malformations were sufficiently large to suggest that small unpublished studies would have little effect on the estimates unless they were highly heterogeneous. The analyses of gastrointestinal, urogenital, musculoskeletal, and VACTERL malformations were limited by their small sample sizes and low number of events: the interpretation of these effects should therefore be treated more cautiously. The significant effect observed for VACTERL should also be treated cautiously, as the confidence intervals for this effect were wide.

A significant strength of this current study is its use of standard systematic review methods. By asking a focused question solely on exposure to HPTs, and excluding exposure to other hormones, we have been able to assess the heterogeneity of the effect estimates. However, as with any observational studies, there is always the possibility that an unknown confounder could be the cause of the observed difference. While such a possibility cannot be ruled out, the lack of heterogeneity means that such a confounder would potentially have to act in the same direction, despite many different confounders being collected and controlled for. Confounding factors with variable effects on the effect estimates would have probably led to a high degree of heterogeneity, which would have prevented pooling; this was not the case.

Conclusion

Regulators were first made aware of the link between exposure to HPTs and congenital malformations in 1967. After 1975, the Primodos label was changed to state that the medication should not be used in pregnancy because of a risk of malformations (see [Figure 9](#)). The evidence of an association has previously

1975 PRIMODOS Data Sheet compendium

PRIMODOS*

Presentation Each orange, sugar-coated tablet contains 10 mg norethisterone acetate and 0.02 mg ethinyl oestradiol.

Uses Primodos is intended for the symptomatic treatment of secondary amenorrhoea of short duration, not due to pregnancy, by the production of a withdrawal bleeding within three to six days of tablet-taking, or in exceptional cases, after up to 10 days.

Dosage and administration One tablet to be taken on each of two consecutive days.

Contra-indications, warnings, etc Rarely, a feeling of nausea may occur.

Special precautions: In the rare case in which bleeding does not follow the administration of Primodos (usually when the amenorrhoea has lasted six months or longer), although the patient is not pregnant, a careful search for organic disease should be made before giving further hormonal treatment, which should be as for primary amenorrhoea.

Pharmaceutical precautions Store in cool, dry conditions, away from strong sunlight: shelf-life five years.

Legal category S4B.

Package quantities Foil strips of 2 and 20 tablets.

Further information Nil.

Product licence number 0053/5027.

Schering Chemicals Limited. Primodos. In: Data Sheet Compendium 1975. London: The Association of the British Pharmaceutical Industry, 1975: 785.

1978 PRIMODOS

Data Sheet compendium

Added Warning

PRIMODOS*

Presentation Each orange, sugar-coated tablet contains 10 mg norethisterone acetate and 0.02 mg ethinyl oestradiol.

Uses Primodos is intended for the symptomatic treatment of secondary amenorrhoea of short duration, not due to pregnancy. In the absence of pregnancy it is possible to produce a withdrawal bleeding within three to six days or, in exceptional cases, after 10 days.

Dosage and administration One tablet to be taken on each of two consecutive days.

Contra-indications, warnings, etc

Contra-indication: Pregnancy.

Warning: A possibility exists of an association between the use of Primodos during early pregnancy and an increased incidence of congenital abnormalities.

Because of this possible hazard, Primodos must not be taken unless it is certain that the patient is not pregnant.

Special precautions: In the rare case in which bleeding does not follow the administration of Primodos (usually when the amenorrhoea has lasted six months or longer), although the patient is not pregnant, a careful search for organic disease should be made before giving further hormonal treatment, which should be as for primary amenorrhoea.

Side-effects Rarely, nausea may occur.

Overdosage: There have been no reports of ill-effects from overdosage and treatment is generally unnecessary. If overdosage is discovered within two or three hours and is so large that treatment seems desirable, gastric lavage can safely be used.

There are no special antidotes, and further treatment should be symptomatic.

Pharmaceutical precautions Store in cool, dry conditions, away from strong sunlight: shelf-life five years.

Legal category POM.

Package quantities Foil strips of 2 tablets.

Further information Nil.

Product licence number 0053/5027.

Schering Chemicals Limited. Primodos. In: Data Sheet Compendium 1978. London: The Association of the British Pharmaceutical Industry, 1978: 888.

Figure 9. Primodos label 1975 and 1978.

been deemed weak, and previous litigation and reviews have been inconclusive. However, we believe that this systematic review shows an association of oral HPTs with congenital malformations.

Our results show the benefit of undertaking systematic reviews, a study type not in routine use when most of these studies were done. For example, only one study (Greenberg 1997) out of nine reported a significant effect for all congenital malformations; the pooled estimate was significant. Much of the discussion over the associations of HPTs with congenital malformations at the time these studies were published focused on the lack of significance of individual studies¹², although it was also recognized that the numbers involved were insufficient to reject the hypotheses¹³.

Declarations

Data availability

Dataset 1: Study extraction sheet [10.5256/f1000research.16758.d222937](https://doi.org/10.5256/f1000research.16758.d222937)³⁰

Grant information

The Evidence Synthesis Working Group is funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR) [ProjectNumber 390]. The views expressed are those of the author(s) and not necessarily those of the NIHR, the NHS or the Department of Health

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Supplementary File 1. Committee on Safety of Medicines report 1975-colour pdf.

[Click here to access the data](#)

Supplementary File 2. Search Strategy.

[Click here to access the data](#)

Supplementary File 3. List of hormone pregnancy tests (HPTs) included in evidence search.

[Click here to access the data](#)

Supplementary File 4. Completed Meta-analysis Of Observational Studies in Epidemiology (MOOSE) checklist.[Click here to access the data](#)**Supplementary File 5.** Full references for included studies.[Click here to access the data](#)**Supplementary File 6.** Funnel Plots.[Click here to access the data](#)**References**

- Website 1: Medicines and Healthcare Products Regulatory Agency. (accessed 1 Jun 2018).
[Reference Source](#)
- Website: Medicines and Healthcare products Regulatory Agency. 2015 Call for Evidence. (accessed 1 Jun 2018).
[Reference Source](#)
- Report of the Commission on Human Medicines' Expert Working Group on Hormone Pregnancy Tests. GOV.UK. (accessed 24 May 2018).
[Reference Source](#)
- Anger after report finds birth defects not caused by hormone pregnancy tests. accessed 2/8/18.
[Reference Source](#)
- Review launched to respond to patient concerns about NHS treatments. (accessed 2/08/2018).
[Reference Source](#)
- Ottawa Hospital Research Institute. (accessed 1 Jun 2018).
[Reference Source](#)
- Lunny C, Knopp-Sihota JA, Fraser SN: Surgery and risk for multiple sclerosis: a systematic review and meta-analysis of case-control studies. *BMC Neurol.* 2013; **13**: 41.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ernst E, Pittler MH: Re-analysis of previous meta-analysis of clinical trials of homeopathy. *J Clin Epidemiol.* 2000; **53**(11): 1188.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Egger M, Smith GD: meta-analysis bias in location and selection of studies. *BMJ.* 1998; **316**: 61–6.
[Publisher Full Text](#)
- Higgins JP, Thompson SG: Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002; **21**(11): 1539–58.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stroup DF, Berlin JA, Morton SC, et al.: Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000; **283**(15): 2008–12.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Torfs CP, Milkovich L, van den Berg BJ: The relationship between hormonal pregnancy tests and congenital anomalies: a prospective study. *Am J Epidemiol.* 1981; **113**(5): 563–74.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gal I, Kirman B, Stern J: Hormonal pregnancy tests and congenital malformation. *Nature.* 1967; **216**(5110): 83.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gal I: Risks and benefits of the use of hormonal pregnancy test tablets. *Nature.* 1972; **240**(5378): 241–2.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Brown S, Fraga LR, Cameron G, et al.: The Primodos components Norethisterone acetate and Ethinyl estradiol induce developmental abnormalities in zebrafish embryos. *Sci Rep.* 2018; **8**(1): 2917.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Maier WE, Herman JR: Pharmacology and toxicology of ethinyl estradiol and norethindrone acetate in experimental animals. *Regul Toxicol Pharmacol.* 2001; **34**(1): 53–61.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Archer DF, Navarro FJ, Leslie S, et al.: Effects of levonorgestrel, medroxyprogesterone acetate, norethindrone, and 17beta-estradiol on vascular endothelial growth factor isomers 121 and 165 in Ishikawa cells. *Fertil Steril.* 2004; **81**(1): 165–70.
[PubMed Abstract](#) | [Publisher Full Text](#)
- OCEBM Levels of Evidence. CEBM. 2016; (accessed 29 May 2018).
[Reference Source](#)
- Sackett DL: Bias in analytic research. *J Chronic Dis.* 1979; **32**(1–2): 51–63.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Oremus M, Oremus C, Hall GB, et al.: Inter-rater and test-retest reliability of quality assessments by novice student raters using the Jadad and Newcastle-Ottawa Scales. *BMJ Open.* 2012; **2**(4): pii: e001368.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wallis CJ, Mahar AL, Choo R, et al.: Second malignancies after radiotherapy for prostate cancer: systematic review and meta-analysis. *BMJ.* 2016; **352**: i851.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Croles FN, Nasserinejad K, Duvekot JJ, et al.: Pregnancy, thrombophilia, and the risk of a first venous thrombosis: systematic review and bayesian meta-analysis. *BMJ.* 2017; **359**: j4452.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hackshaw A, Morris JK, Boniface S, et al.: Low cigarette consumption and risk of coronary heart disease and stroke: meta-analysis of 141 cohort studies in 55 study reports. *BMJ.* 2018; **360**: j5855.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Boland MV, Ervin AM, Friedman DS, et al.: Comparative effectiveness of treatments for open-angle glaucoma: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2013; **158**(4): 271–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Haraus EP, Garcia-Prats AJ, Law S, et al.: Treatment and outcomes in children with multidrug-resistant tuberculosis: A systematic review and individual patient data meta-analysis. *PLoS Med.* 2018; **15**(7): e1002591.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Veronese N, Cereda E, Solmi M, et al.: Inverse relationship between body mass index and mortality in older nursing home residents: a meta-analysis of 19,538 elderly subjects. *Obes Rev.* 2015; **16**(11): 1001–1015.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bae JM: A suggestion for quality assessment in systematic reviews of observational studies in nutritional epidemiology. *Epidemiol Health.* 2016; **38**: e2016014.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hartling L, Milne A, Hamm MP, et al.: Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *J Clin Epidemiol.* 2013; **66**(9): 982–93.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Heneghan C, Aronson JK, Spencer E, et al.: Dataset 1 in: Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis. *F1000Research.* 2018.
<http://www.doi.org/10.5256/f1000research.16758.d222937>

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 29 November 2018

<https://doi.org/10.5256/f1000research.18318.r40291>

 **Jesse Olszynko-Gryn** ¹, **Cyrille Jean**²

¹ Department of History and Philosophy of Science, University of Cambridge, Cambridge, UK

² Sciences Po, Paris, France

This is a timely and much-needed paper that deserves to be widely read and cited. It provides the first systematic review and meta-analysis of old epidemiological data pointing towards a long-acknowledged association between HPTs and birth defects. Most of the paper is devoted to apparently rigorous statistical analysis. We leave constructive criticism of the statistics to other, more appropriately qualified reviewers. Instead, we confine our comments to the historical context and factual details presented in the paper. These, on the whole, are entirely satisfactory. But some minor errors — that do not significantly detract from the overall argument — should be amended:

1. 'Oral hormone pregnancy tests (HPTs), such as Primodos, containing ethinylestradiol and high doses of norethisterone, were given to over a million women from 1958 to 1978' (p. 1).

It is worth clarifying that HPTs were available as injections from 1950 and in tablet form (e.g., Schering's Orasecron, Roussel's Amenorone Forte), in the UK, from at least 1956. See, for example, Britton (1956¹); and

<https://archive.org/details/b19974760M4180/page/n45?q=amenorone+1956>. For an extended discussion, see Olszynko-Gryn (2014), available for download [here](#). Furthermore, not all HPTs contained norethisterone; different companies used other types of synthetic progesterone, and the same goes for ethinylestradiol.

2. 'Oral hormone pregnancy tests (HPTs), such as Primodos (known as Duogynon in Germany), were used from 1958 to 1978, before urine pregnancy tests were available' (p. 3).

Contrary to popular belief, urine pregnancy tests were in fact widely though unevenly available between 1958 and 1978 and HPTs were never the dominant method of pregnancy testing. For a detailed timeline of pregnancy testing in the UK, please see Olszynko-Gryn *et al.* (2018²), esp. pp. 35-36. It would also be helpful to clarify that HPTs were removed from UK market in 1978, but earlier and later elsewhere. See Olszynko-Gryn *et al.* (2018²) for details (pp. 41-42).

3. 'The test principle was that menstruation would be induced in those who were not pregnant' (p.3).

At the time HPTs were variously described as 'clinical', 'hormonal', or 'withdrawal bleeding' pregnancy tests and it would be more precise to refer to their effect as inducing *menstrual-like withdrawal bleeding*, which is not identical to menstruation.

4. Worth mentioning that Gal 1967, though a highly significant intervention, was not the first published warning against HPTs; these began to appear as early as 1956, in response to marketing literature aimed at GPs. See Britton (1956¹) and Olszynko-Gryn *et al.* (2018²), p. 36.
5. 'However, we believe that this systematic review shows an association of oral HPTs with congenital malformations' (p. 17).

More optionally, the authors might consider reflecting on the extent to which the association they identify implies a *causal* association. An association between the use of HPTs and birth defects has long been recognised and was rarely in dispute. Many experts explained the association in terms of a suspected though as unknown direct mechanistic effect of HPTs on the developing human embryo. Others, however, preferred to explain the association in terms of underlying factors, e.g., a patient history of miscarriage or birth defects. This view, which still has traction in some quarters, is discussed to some extent in Olszynko-Gryn *et al.* (2018²) (pp. 39-41). The authors might usefully offer a fresh perspective based on their findings, in the Conclusion and/or in the interesting discussion of unknown mechanisms on p. 16.

References

1. Britton H: Pregnancy Test. *BMJ*. 1956; **2** (4989). [Publisher Full Text](#)
2. Olszynko-Gryn J, Bjørvik E, Weßel M, Jülich S, Jean C: A historical argument for regulatory failure in the case of Primodos and other hormone pregnancy tests. *Reprod Biomed Soc Online*. 2018; **6**: 34-44 [PubMed Abstract](#) | [Publisher Full Text](#)

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: No competing interests were disclosed.

Referee Expertise: History of medicine

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 05 November 2018

<https://doi.org/10.5256/f1000research.18318.r40053>



David Healy

School of Medical Sciences, Bangor University, Bangor, UK

I am fully supportive of this article on the effects of hormone pregnancy tests as it stands. I have no substantive criticism of the content or methods.

I am of course interested in why the regulator (MHRA) did not find comparable results but this is not a matter that should be addressed in this article.

There is one extra point that this article may speak to which is that from some time it was thought that teratogens caused signature defects - such as the phocomelia of thalidomide. This may now be a minority position (I'm not sure of this point). The findings here do not support that point of view. I can understand if the authors may think that commenting on this point is a matter for others or for another article; I mention it for consideration.

I have one very minor point about the layout which is that in the column where the numbers of women recruited to various studies is mentioned, the right justification of paragraphs leads to an odd spacing between 28 thousand and 671 - this doesn't apply when the page is resized.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Yes

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Discuss this Article

Version 1

Reader Comment 19 Nov 2018

Nick Brain, Association for Children Damaged by Hormone Pregnancy Tests, UK

As Prof. Heneghan and colleagues note, previous studies of the association of Hormone Pregnancy Tests (HPTs) with congenital malformations in children have been acknowledged as being too small to definitively refute the hypothesis. This pattern is seen often in the epidemiological literature as investigators design studies that either do not make use of statistical power calculations or are over-optimistic in their modelling of effect sizes, leading to studies that may not detect an effect but also provide little confidence that there is no effect.

Heneghan *et al.* also note that it has previously been argued that there is no known mechanism to explain teratogenicity of HPTs. Of course, the lack of a known mechanism does not mean that there is not one. It is well recognised that our knowledge of gene function and biological pathways is very incomplete. It is also very well recognised that response to drugs is heterogeneous - indeed, understanding genetics and heterogeneity of drug response is a major focus for the pharmaceutical industry today as it seeks to identify patient populations in whom new drugs are effective and safe.

This review and meta-analysis by Heneghan *et al.* is, therefore, a particularly important paper for our understanding of risks associated with HPT exposure including, as it does, 26 smaller studies combining to give sample sizes as large as 61,642 mothers with 3,274 exposed to HPTs during pregnancy.

The strength of the effects indicated by the Odds Ratios in these analyses, and the levels of statistical significance that are stated by the authors, surely now give pause for thought and reinforce the concerns that the previous studies were individually under-powered.

These results are robust across not only risk of "any" malformation (37% higher risk for children of mothers exposed to HPTs) but also risks of specific types of malformations, with risks doubling, tripling or higher compared to mothers not exposed to HPTs.

Of course, any meta-analysis of observational studies has its limitations and the authors discuss these, and their study design strategies to mitigate them, in detail. In the absence of large, randomised HPT trials (not only atypical of the era but also unethical in pregnant women) this meta-analysis is an essential and important step forward.

The authors' search strategy was designed to minimise the risk of missing published HPT studies. That 409 articles were found and only 24 (less than 6%) included after review underlines the breadth of the search.

The authors note that non-randomised study quality scoring systems such as the Newcastle-Ottawa Scale (NOS) have their critics but also note that the NOS scale has been widely used (from a simple Title/Abstract search of PubMed, I see >1,200 citations in the last three years). Their methodology was rigorous to minimise assessor discrepancies using a standard review template, two assessors and resolving discrepancies in consultation with the other authors. They also note that a previous study showed scoring concordance which is improved by training and experience, even for novice assessors. Given the level of expertise of the authors, this bodes well for the study.

In their discussion the authors also consider the likely impact of other limitations such as confounding factors, susceptibility bias, ascertainment bias and publication bias.

Additional checks and balances during the statistical analyses sought to uncover any hidden biases or heterogeneity but did not indicate such issues and demonstrated that the results were robust to NOS score and to sensitivity analyses.

Following this detailed discussion of their work, the authors conclude that an association of oral HPTs with congenital malformations exists. This work must now be considered very seriously given the doubts that have been expressed regarding previous studies and the significance of these findings for the mothers who were exposed to HPTs during pregnancy and their children.

Competing Interests: Relative of member of ACDHPT. I also hold a Ph.D. in genetic epidemiology of complex traits.

Reader Comment 07 Nov 2018

Sarah-Jane Richards, Fortitude Law, UK

Heneghan *et al.* 2018 report on a meta-analysis of all available studies on the historical risks of congenital malformations attributable to the oral hormone pregnancy tests (HPTs) – principally Primodos – given to more than a million women between 1958 and 1978. This is a really nice example of how the rigorous application of the modern statistical methods of systematic review allows a clear picture of the side effects of a pharmaceutical product to be demonstrated when the traditional approach of many small studies conducted at the time do not have the power individually to reveal significant effects. As the authors conclude “*Much of the discussion over the associations of HPTs with congenital malformations ... focused on the lack of significance of individual studies, although it was also recognized that the numbers involved were insufficient to reject the hypotheses*”.

The optimal design of a clinical trial to avoid all influences of bias is the randomised control trial, in which subjects are randomly allocated to the different treatment groups. In other types of trial there is always the possibility that biases in the allocation of subjects to groups can influence the outcome. In the case of Primodos, it would however have been unethical to allocate women to receive drugs with known concerns. Consequently the study authors focussed on cohort and case-control trials reporting on the risk of congenital malformations after administration of HRTs, and then applied a rigorous analysis of trial design and power to assess risks of bias according to the well-recognised Newcastle-Ottawa scale for rating the risk of bias in non-randomised trials.

Importantly, an increased risk of all congenital malformations was found consistently among subsets of trials at high or low risk of bias, providing strong corroboration for the view that the revealed increase in congenital malformations could not be attributed to biases in patient selection. The analysis went on to break down the data for specific risks, showing that the significant overall incidence of malformations after HPT was not just due to a single target, but was raised significantly for malformations in heart, nervous, and musculoskeletal systems when analysed separately. A raised risk was seen similarly in rarer targets, such as gastro-entorhinal system, but in these cases the numbers of observed cases were too low to reach a conventional level set for significance.

It is not sufficient simply to undertake a formal systematic review to make old data definitive. What makes this study convincing is the experience and expertise of the team at the Oxford Centre for Evidence-Based Medicine, the quality of the evidence, the extent of the search criteria to include all relevant studies, reported in multiple languages, not just English, with a judicious balanced exclusion of the least relevant or objective reports, and the rigorous attempt to undertake objective analysis of the risks and control of selection bias in different studies. As a result, we have here, the most convincing and compelling evidence that Primodos in particular, and related HPTs more generally, did cause unforeseen congenital

malformations in some children of mothers administered these tests throughout the 1950s, 60s and 70s. The growing fears both at the time and in subsequent litigations, can no longer be simply rejected on the basis of lack of evidence of significant risk or explained away in terms of the selection bias in the trials designed to assess that risk that were conducted at the time.

Dr. S-J Richards, Senior Consultant Fortitude Law
Prof. SB Dunnett, School of Biosciences, Cardiff University

Competing Interests: None

Reader Comment 04 Nov 2018

Leonard Lofts, The Northam Care Trust, UK

Excellent research. Looking forward to observing peer review. The research is so important for so many families.

Dr Len Lofts

Competing Interests: No competing interests were disclosed.

Reader Comment 02 Nov 2018

Marie Lyon, Association for Children Damaged by Hormone Pregnancy Tests, UK

Prof. Henegan's systematic analyses of epidemiological studies, is a scientific review which members of the Association for children damaged by HPT's have waited over 45 years for. The findings are incredible and mirror the congenital abnormalities suffered by our members. It is a scandal that this epidemiological study was not commissioned by the Government Health Authorities and we cannot thank Prof. Heneghan and his colleagues enough, for the comprehensive and utterly compelling review.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research