EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

26 April 2019
EMA/245512/2019 - corr 1
Committee for Medicinal Products for Human Use (CHMP)

Assessment report

Procedure under Article 5(3) of Regulation (EC) No 726/2004

Norethisterone and ethinylestradiol

Procedure number: EMEA/H/A-5(3)/1477

Note:

Assessment report as adopted by the CHMP with all information of a commercially confidential nature deleted.

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

© European Medicines Agency, 2019. Reproduction is authorised provided the source is acknowledged.

Table of contents

1. Information on the procedure

On 27 November 2018 the National competent authority in UK, MHRA, triggered a procedure under Article 5(3) of Regulation (EC) No 726/2004, and asked the CHMP to give a scientific opinion on the conclusion of a recently published systematic review and meta-analysis by Heneghan and colleagues (2018; version 1)[1] related to the use of norethisterone and ethinylestradiol containing oral hormone pregnancy tests (HPTs) in pregnancy and a potential associated increased risks of congenital malformations.

Since ethinylestradiol, norethisterone and other progestogens that are found in hormone pregnancy tests are commonly also found in a range of widely used authorised gynaecological medicines across the EU, MHRA considers that there is significant public health interest in a scientific opinion from CHMP on whether these substances would potentially be teratogenic. Despite the fact that these substances are contraindicated in pregnancy, studies show that a substantial number of women continue to take them without realising they are pregnant.

The CHMP's opinion is particularly sought on:

- the suitability and robustness of the methodology, including the selection and application of the data quality score;
- any clinical implications.

As the outcome of this procedure could be of relevance to norethisterone and ethinylestradiol-containing medicinal products in the EU, MHRA also asked CHMP to indicate whether there are public health concerns related to these two substances, and, if so, to recommend whether the concerns need to be investigated further at Community level.

2. Scientific discussion

2.1. Introduction

Ethinylestradiol and norethisterone (synthetic forms of oestrogen and progesterone, respectively) are the active substances of products for post-menopausal use or contraceptive use.

The synthetic oestrogen ethinylestradiol is the oestrogen component present in most of the combined hormonal contraceptives (CHCs) that are available on the EU market, with less predominance for estradiol and estradiol valerate. The dose of ethinylestradiol in CHCs ranges between 0.015 – 0.035 mg. More diversity is found in respect to the progestogen component of CHCs where norethisterone is one example among many other progestogens (desosgestrel, etonogestrel, gestodene, levonorgestrel, norelgestromin, norethindrone, norethisterone norgestimate). The norethisterone dose in CHCs ranges between 0.5 and 1 mg.

Ethinylestradiol and norethisterone can also be found as a monocomponent within authorised medicinal products. Ethylestradiol is approved as monotherapy for hormone replacement therapy in doses ranging from 0.01 mg to 1 mg. Norethisterone is currently approved in doses of 5 – 15 mg/day and 10 mg/day (Primolut N or Primolut Nor) as monotherapy for the treatment of menstrual irregularities and

[1] Heneghan C., Jeffrey K. Aronson, Elizabeth Spencer, Bennett Holman, Kamal R. Mahtani, Rafael Perera, Igho Onakpoya. *Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis [version 1; referees: awaiting peer review]*. F1000Research 2018, 7:1725 Last updated: 31 OCT 2018.

endometriosis, respectively. Only one CHC containing ethinylestradiol and norethisterone (Loestrin 0.02 mg + 1 mg) has been approved in the UK.

Until late 70's early 80's (up to 1978 in UK for Primodos from Schering AG; 1981 in Germany (known as Duogynon) they were also constituents of hormonal pregnancy tests (HPT), first as injections (1950) and later in tablets (since 1956 in UK) and were used before the commercialization of urine pregnancy test. In HPTs, ethinylestradiol was present together with large doses of norethisterone, much higher than found in currently approved combined hormonal contraceptives (CHCs).

Various reports, in particular between 1950s and 1978, when Primodos was withdrawn from the UK market (in Germany, Duogynon was taken off the market in 1981) have been published pointing to a potential association between women that have taken HPT to diagnose their pregnancy and the observation of a variety of congenital anomalies in the offspring.

Due to various public concerns, the MHRA established an Expert Working Group (EWG) in 2015 with the purpose to review the available data and assess the postulated correlation between exposure to HPTs during pregnancy and congenital anomalies or any other adverse outcomes in pregnancy. Their report, which includes the original Landesarchiv Berlin Files and studies from 1946 to 2018, was published in October 2017. The EWG's conclusion was that the available evidence does not support a causal association between the use of HTPs during early pregnancy and adverse outcomes (e.g. miscarriage, stillbirth or congenital anomalies) and also that no implication is envisaged for any currently authorised medicines.

Due to the scientific interest concerning these products, and following recently published systematic review and meta-analysis by Heneghan and colleagues (2018; updated version 2 in 2019)[2] MHRA notified the CHMP, on 27 November 2018, of a procedure under Article 5(3) of Regulation (EC) No 726/2004. Indeed, this meta-analysis suggests that use in pregnancy of oral HPTs containing ethinylestradiol with norethisterone or other progestogens are associated with increased risks of congenital malformations (overall odds ratio 1.40 [1.18, 1.66], with significant increases in the risk of congenital heart disease, nervous system malformations and musculoskeletal malformations.

2.2. Assessment of the meta-analysis

The meta-analysis is authored by Carl Heneghan, Jeffrey K. Aronson, Elizabeth Spencer, Bennett Holman, Kamal R. Mahtani, Rafael Perera, Igho Onakpoya. Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis [version 2; referees: 3 approved]. F1000Research 7:1725; Last updated: 29 JAN 2019.

The meta-analysis is uploaded in an online publishing platform, with no editors, prior peer-review and no impact factor. Readers are invited to review and give comments.

A summary of the systematic review and meta-analysis is presented below and assessed in detail with the purpose to determine whether the current information on the risk of teratogenicity with regard to hormone pregnancy tests (HPTs) add to the current knowledge regarding adverse events in early pregnancy in humans. HPTs included Primodos/Duogynon and also other HPTs available at the time.

The publication stated that 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 between hormone pregnancy tests and congenital malformations. The authors have, therefore, performed a systematic

[2] Carl Heneghan, Jeffrey K. Aronson, Elizabeth Spencer, Bennett Holman, Kamal R. Mahtani, Rafael Perera, Igho Onakpoya (2019). *Oral hormone pregnancy tests and the risks of congenital malformations: a systematic review and meta-analysis [version 2; referees: 3 approved]*. F1000Research 7:1725; Last updated: 29 JAN 2019.

review to obtain all relevant data on the topic, 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 the search strategy were provided in the publication. The authors of the meta-analysis 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 Files, and reference lists of retrieved studies from the start of the databases in 1946 to 20 February 2018. The authors 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; the authors therefore performed additional searches for evidence relating to these (See Supplementary File 3 for List of HPTs included in evidence search). The authors indicated that they performed additional searches for evidence relating to other HPTs. As these other hormone products do not contain the same active ingredients, the results based on these other hormonal products should not be taken into account according to the CHMP. However, it is unclear from the information provided in this meta-analysis whether results of other HPTs than Primodos or Duogynon have also been included.

Study selection

The authors 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, they chose the publication that contained the most comprehensive data as the primary publication. They excluded studies where the intervention was oral hormones taken for other reasons (e.g. oral contraception) and where it was not possible to extract data on hormone pregnancy tests.

They did not restrict the language of publication. They checked additional relevant data and extracted them from the secondary publications when necessary.

A comment from the CHMP was that it is unclear which selection criteria were applied to select the studies, and in addition the studies do not seem to have been peer-reviewed.

Data extraction and risk of bias assessment

Two of the authors acted as reviewers and applied inclusion and quality assessment criteria, compared results, and resolved discrepancies through discussion with the other authors. They 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, they extracted data about the women studied, including ascertainment of cases, age, parity, setting, exposure to other medications, and confounding variables.

In case-control studies, if data were reported on more than one control group, they extracted data where possible for non-disease/non-abnormality controls, and combined control groups if necessary.

Regarding the rating of the Meta-analyses Of Observational Studies in Epidemiology (MOOSE) Checklist, several items are unclear or missing:

- whether the two reviewers independently assessed the studies and whether direct contact with the authors has taken place.
- whether a third independent reviewer was consulted in case of disagreement between the two reviewers with regard to the initial inclusion of studies.
- A qualification of the researchers.
- A list of citations included vs excluded, including justification thereof;
- The method of handling letters to the editor and unpublished studies is unclear, while several have been included in the meta-analyses.

The primary outcome of interest was “all major congenital malformations”.

The authors also categorised outcomes for the congenital anomaly in the offspring at any time into congenital cardiac, gastrointestinal, musculoskeletal, nervous system, and urogenital defects, and Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies, and Limb defects (known as VACTERL syndrome).

The CHMP accepted that the primary outcome of interest is ‘all major congenital malformations’. However, a definition of ‘major congenital malformation’ was not provided. Further, it was noted that in several cases also minor malformations have been included.

The authors assessed quality using the Newcastle–Ottawa Scale (NOS) for non-randomised studies included in systematic reviews. 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 they interpreted a score of 1 to 3 points as indicating a high risk of bias. To determine whether the study had controlled for the most important factors, they selected the items reported in the original paper and resolved disagreements through consensus, in consultation with a third author.

The authors examined whether there was a linear relation between methodological quality and study results, by plotting the odds ratios against the NOS scores and assessed the correlations of NOS scores with several confounding variables they collected.

Details on the exact criteria in applying the NOS score have not been provided. For example, in the determination whether or not the "study controls for the most important factors", it is unclear if the reviewers had considered what are the most important factors in advance, or relied on statements about importance in the original paper.

Data synthesis and statistical methods

The authors of the meta-analysis calculated study-specific odds ratios for outcomes and associated confidence intervals. They analysed the data using a random-effects model. They also assessed heterogeneity across studies using the I^2 statistic and publication bias using funnel plots.

They performed a sensitivity analysis by removing single studies to judge the stability of the effect and to explore the effect on heterogeneity, and described any sources of variation. They 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, they 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.

The authors 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. They used RevMan v.5.3 for all analyses, except for meta-regression, for which they 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). They performed a sensitivity analysis by removing studies with zero events from the analyses.

They followed the reporting guidelines of the Meta-Analysis of Observational Studies in Epidemiology (MOOSE). A completed checklist was made available.

In general, The CHMP considered acceptable the method for the study selection, data extraction and statistical analysis which are considered standard for a meta-analysis. However, it is noted that some details were lacking. In the NOS score applied, it was unclear what is considered as a "relevant control group". Furthermore, assessing the quality of the studies can be subjective and details on the exact criteria in applying the NOS score are not provided. For example, in the determination whether or not the "study controls for the most important factors", it is unclear if the reviewers had considered what are the most important factors in advance, or relied on statements about importance in the original paper.

With regard to adherence to the MOOSE checklist, several omissions were noted, which were discussed above.

Patient involvement

Members of the Association for Children Damaged by HPTs were involved in the original discussions of the meta-analysis and provided input to the outcome choices, the search, the location of study articles, and translations. The authors planned to present the study findings to relevant patient groups and make available lay interpretations.

The CHMP noted that the involvement of this Association in the original discussion of this meta-analysis and the input on the outcome choices, the search and translations raised questions regarding the independency of the conclusions and whether this analysis is based on independent research.

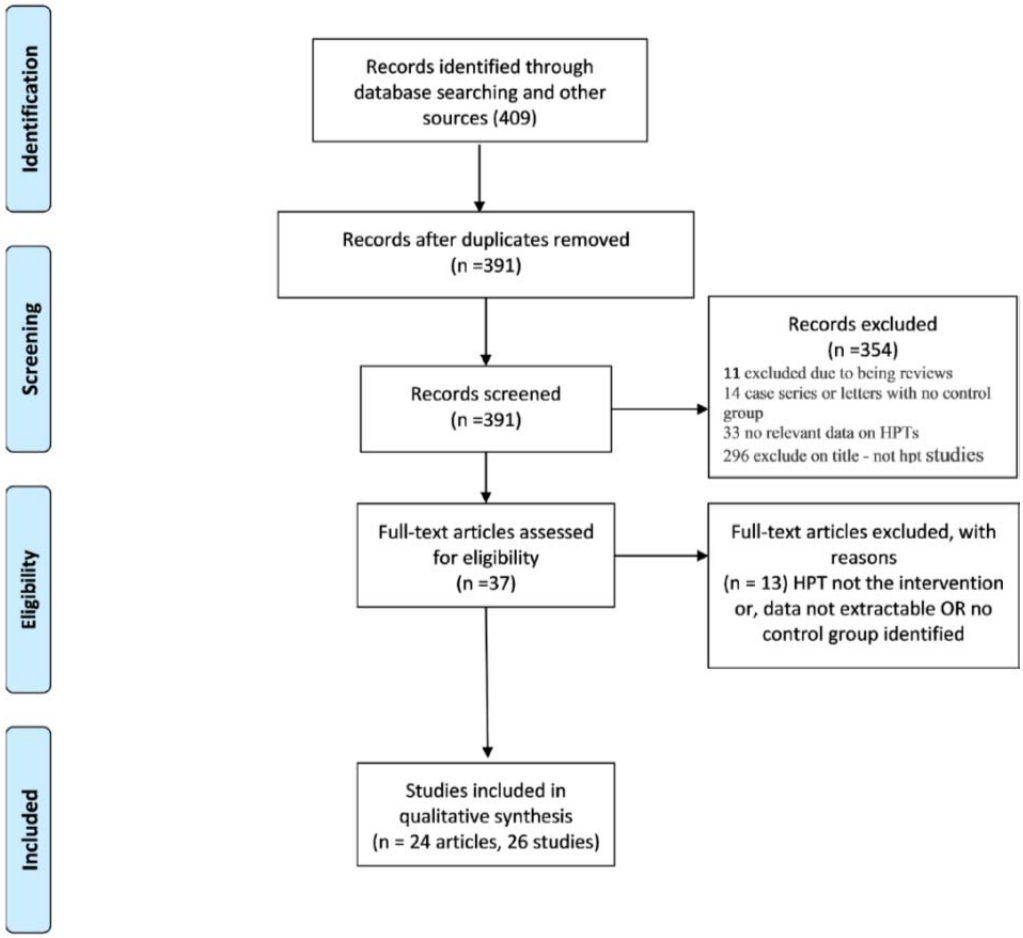
Results

Description of included studies

The authors 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. They 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.

Figure 1. Study flow diagram showing inclusion of relevant studies



The 24 included articles reported on 26 studies (16 case-control studies and ten prospective cohort studies); one article (Nora et al, 1978) included two case-control studies and one prospective study.

The authors found no randomised controlled trials. Of these articles, two were unpublished reports. 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 3,615 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.

It is noted that the authors qualified the three studies by Nora and colleagues (1978) as two case-control studies and one prospective study. However, these were classified as three case-control studies (table 2) and the article itself described three case-control studies and one cohort study. Further, it is noted the study by Nora and Nora in 1975 contained preliminary data, which were also taken into account in the study published by Nora and colleagues in 1978.

The choices of controls in the case-control studies varied; they included, at one extreme, healthy infants born on a date close to the infants of the case 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 (Table 2.).

Table 2. Characteristics of the studies included in the meta-analysis

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 from 110 infants with conotruncal abnormalities of the heart, born 1972-75	Hospitals served by the Maryland State Intensive Care program, USA	Maternal health (hospitalisations, illnesses, treatments); past reproductive history; index pregnancy factors including contraception used previously, fertility treatments, symptoms, illnesses and the medications during pregnancy including hormones; smoking; alcohol intake; occupational history of mother and father; exposure of mother to fumes, paints and insecticides; family 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, 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)
Gall, 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 drug described 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	106 cases of congenital limb defects and 108 unaffected controls	New York State, USA	Age, parity, race	Controls matched on birth date, mothers race and age +/- 2 years; and by default, due to adjacent records for cases and controls these matched well on country 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 from 1 July 190 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
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	VACTERL
Nora 1975	15 patients with multiple congenital anomalies. 30 controls (15 with chromosomal anomalies, 15 with functional heart murmurs)	University of Colorado Medial 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, area of residence, 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 data of birth, area of residence	Matching was for sex, race, approximate date of birth, area of residence	Congenital heart disease (congenital 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
	390 cases, 1,254 controls, HPTS:14/388 cases vs 35/1246 controls	State care service for congenital heart diseae	Parity, mother's education level, insulin use, alcohol, tobacco	Controls: births within the 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
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 extrophy
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 aand general health backgrounds		Congenital malformations, also congenital heart defects, skeletal anomalies, microencephaly
Hadjigeorgiou 1982	Retrospective cohort, Alexandra Maternity Hospital Greece, births 1975-77. 15,535 live births, 559 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 1972, University Hospital Göttingen; 617 (17.2 %) with abnormal pregnancy test	Hospital birth centre	Non-reported		Congenital malformations
Kullander 1976	6,376 pregnancies, Malmo, 1963-5, resulting in 5,753 live births, 5,002/753 no abnormality, 751/5,753 with abnormality. 156 women took Primodos	Sweden	Major and minor disease; the woman's age, parity, maternal 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, 2 had taken HTPs.	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		Major malformations
Roussel 1968	Pregnancies 1966 to 1967	General practices, UK	NR		Central nervous system malformations including anencephaly, microcephaly, meningmyelocele, myelocele, spina bifida
Rumeau-Roquette 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	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

Quality assessment of the 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 et al, 1971, a published abstract as a letter) and two were cohort studies (Fleming et al, 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 (table 3). 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 and nine controlled for important additional factors.

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 et al, 1971; Levy et al, 1973; Tummler et al, 2014; Fleming et al, 1978; Haller, 1974; Meire & Vuylsteek, 1978; Rousel et al, 1968).

NOS scores correlated with the increasing number of confounding variables collected (r = 0.83). The authors indicated that because of inadequate numbers of included studies, the authors did not use more advanced statistical methods to assess publication bias.

Table 3. Newcastle-Ottawa scale scores for included studies

Newcastle-Ottawa scale case control studies										
					Comparability of cases and controls of the design or analysis					
Study ID	**Is the case definition adequate?**	**Are the cases representative**	**Selection of controls adequate**	**Definition of controls adequate**	**a) Study controls for the most important factor**	**b) study controls for important additional factors**	**Ascertainment of exposure adequate**	**Same method of ascertainment for cases and controls**	**Non-response rate adequate**	**Total score /9**
Ferencz 1980	yes	yes	yes	yes	yes	yes	yes	yes	Yes	9
Gal 1972	Unclear	yes	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	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	yes	yes	yes	yes	8
Nora 1978	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
Tummler 2014	yes	no	no	no	no	no	no	yes	yes	3

Newcastle-Ottawa scale cohort studies										
Study ID	**Is the case definition adequate?**	**Are the cases representative**	**Selection of controls adequate**	**Definition of controls adequate**	**a) Study controls for the most important factor**	**b) study controls for important additional factors**	**Ascertainment of exposure adequate**	**Same method of ascertainment for cases and controls**	**Non-response rate adequate**	**Totl score /9**
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
Michaelis 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

The CHMP noted that the authors reported that the studies included in the meta-analysis were published between 1972 and 2014. Of note, all but one of the studies included were published before 1988 and some even before 1972. The study published in 2014 was not taken into account. The authors provided a supplementary file 5 from where the list of studies discussed below is taken. This list did not correspond completely with the studies described in tables 2 and 3 above. In table 2, the study by Hellstrom and colleagues (1976) is missing. Both tables (2 and 3) included a study by Janerich (1974), which was not included in the supplementary file and for which it is unclear which article is referred to. The Janerich (1974) and Tummler (2014) studies did not appear to have been used in the analyses. In actual terms this meta-analysis included studies published between 1968 and 1987. Below a discussion on the studies used is presented together with comments on their suitability.

1. Gal I (1972) Risks and benefits of the use of hormonal pregnancy test tablets. Nature 240: 241-242.

The CHMP agreed with the authors of the meta-analysis that the definition and representativeness of the cases are unclear in this study. There is no clear information on the recruitment of the controls, and therefore there are doubts on the classification used for the “selection of controls”. A selection bias cannot therefore be excluded. This position is reinforced by the “Comments” to the article by Laurence and colleagues (1971) that suggested that the selection of controls can be considered highly biased. In addition, there are doubts on the classification of comparability as differences on age (more cases than controls over 35 years) and acute infection are important confounding factors. So, a NOS score of of 4 seems more realistic than 6. As this study is published as a letter and thus contains very limited information, its results cannot be taken into account. The results pointed to an association between hormonal exposure and spina bifida, but selection and confounding bias are possible, which therefore casts doubts on the study conclusion.

2. Levy EP, Cohen A, Fraser FC (1973) Hormone treatment during pregnancy and congenital heart defects. Lancet 1(7803): 611

This study evaluated hormone use as aetiological factor in congenital heart disease. The hormone treatment consisted of progesterone use for threatened abortion, and one woman received a pregnancy test. Although this study described concomitant medications (insulin, phenformin, imipramin, thyroid therapy), it is not clear whether this was balanced between cases and controls, or whether they controlled important factors, like reproductive history. The controls were patients with a mendelian disorder and were matched on birth date. Although the reviewers assessed this as adequate for "Study controls for the most important factor" in the NOS score, it could be considered that this study has a NOS score of 3 rather than 4, which indicates a high risk of bias. However, the study is published as a letter and thus contains very limited information, its results cannot be taken into account.

3. Haller J (1974) Hormontherapie wahrend der graviditat. Deutsches Arzteblatt 14: 1013-1015.

Preliminary results of an ongoing study investigating exposure to hormones during pregnancy and the risk of malformations reported no increased risk. No information on the design and methods has been provided and the type of HPT was not specified. However, the study is published as a letter and thus contains very limited information; its results cannot be taken into account.

4. Nora AH, Nora JJ (1975) A syndrome of multiple congenital anomalies associated with teratogenic exposure. Arch Envir Health. Vol 30.

Preliminary results of an ongoing study. Some recall bias may be present in this study, as pregnancy-test history was assessed in interviews about one year after birth. This may have been exaggerated since the authors describe that considerable probing was needed to find positive pregnancy-test history as the patients did not regard it as medication and not in all cases the type of pregnancy-test (progestogen with or without oestrogen) could be recalled. However, the control group consisted of patients with chromosomal anomalies and patients referred for cardiac disease, who may have a similar recall. Therefore, a NOS score of 8 appears questionable. Furthermore, it appears that the results presented here are preliminary findings of the ongoing study published by Nora and colleagues (1978). Therefore, it is questionable whether these results can be taken into account.

5. Kullander S, Kallen B (1976) A prospective study of drugs and pregnancy. Acta Onstet Gynecol Scand 55: 221-224

This study investigated the possible role of hormones in malformations, with special attention to Primodos. However, case definition of malformation was not clear. No harmful effect on embryonic development could be demonstrated. The reviewers included numbers from both minor and major malformations. This increases heterogeneity in the meta-analysis as most other studies focus on major malformations only.

6. Hellstrom B, Lindsten J, Nilsson K (1976) Prenatal sex-hormone exposure and congenital limb reduction defects. Lancet 2:372-3.

This publication concerns a very small series of cases. These cases were published as a letter with little detailed information, therefore the results cannot be taken into account.

7. Janerich DT, Dugan JM, Standfast SJ, Strite L (1977) Congenital heart disease and prenatal exposure to exogenous sex hormones. BMJ 1(6068):1058-60

This study investigated whether any type of sex hormone exposure (contraceptive pill, HPTs, supportive hormone treatment) had increased risk for congenital heart disease (CHD). Cases were

selected based on CHD malformation recorded on their birth certificate; in 29 (out of 104) patients information was lacking to be classified as definite CHD and in 5 did not have CHD. This questions the case definition. Only two cases were on contraceptive use. Nevertheless, the reviewers scored the case definition as adequate and representative. However, no separate analyses were made for the type of hormonal exposure and the type of HPT was not specified. Of note is a possible recall bias as no information was provided about the time period of the study and as per the information gathered, 91% of cases and 80% of controls refer the use of drugs during pregnancy. It is also not clear if a history of previous congenital malformations was considered for the analysis. Similarly, the specific period of time of the hormone use during the first trimester is not understood. Facing the above, a NOS score of 6 or even 5 instead of 7 as reported by Heneghan and colleagues (2018), would be more adequate. Therefore, this study cannot be taken into account.

8. Greenberg, G, Inman WHW, Weatherall JAC, Adelstein AM, Haskey JC (1977) Maternal drug histories and congenital anomalies. BMJ 2: 853-856.

This study investigated drug use during first trimester in mothers of malformed babies vs mothers of healthy babies. An association was noted between the use of a hormonal pregnancy test and the subsequent birth of a malformed baby. The study was performed adequately, although case definition was unclear.

9. Rumeau-Rouquette et al (1978) Malformations congenitales risques perinatales. L'institut National de la Sante et de la recherce medicale (INSERM) (253/2)

This study describes the percentage of malformations reported in a case series of women exposed to several drugs in the first 3 months of pregnancy, among others to sex hormones (progesterone- or testosterone derivatives, estrogens, and combinations of oestrogen-progestogen). In this study, a total of 2 malformations were reported in 535 women who had used Duogynon, compared to 160 malformations in 9662 non-users. It is not clear what numbers were taken from this publication and used in the analysis. Therefore, it is questionable whether these results can be taken into account.

10. Meire F, Vuylsteek K (1978) Apr 1;1(6116):856. Continued use of hormonal pregnancy tests. BMJ 1:856

As this study is published as a letter and contains limited information, its results cannot be taken into account.

11. Nora JJA, Nora H, Blu J, Ingram J, Fountain A, Peterson M, Lortscher RH, Kimberling WJ (1978) Exogenous progestogen and estrogen implicated in birth defects. JAMA 240(9): 837-843

In study 1, cases were matched to both controls who were referred for evaluation of heart murmurs and when no match could be found, to healthy controls. This results in a mixed control group. Especially in the comparison with healthy controls, recall bias could have occurred, as parents of a baby with VACTERL may remember pregnancy-tests better, especially if they are interviewed more than one year after birth.

In study 2/3, a mix of control patients was used, first matching to patients with a single mutant gene and chromosomal disorders, supplemented with controls from a prospective cohort. It is not clear what numbers were taken from this publication and used in the analysis.

12. Fleming DM (1978) Abnormal outcome of pregnancy after exposure to sex hormones. Personal communication. (Provided by Bayer).

As this study concerned an unpublished personal communication only, the results cannot be taken into account.

13. Goujard J, Rumeau-Rouquette C, Saurel-Cubizolles MJ (1979) Hormonal tests of pregnancy and congenital malformations. Journal de gynecologie obstetrique et biologie de la reproduction 8: 489-196

This study evaluated the teratogenic action of drugs on human beings. Two surveys were performed, part 1 between 1963-1969 and part 2 in 1973-1975. Deficiencies were noted, i.e. only 1 of the 2 study parts was taken into account. Additionally, while the study does not appear to control for any potential confounders, such as reproductive history, the reviewers indicate in their NOS scoring list that the study controls for the most important factor. Further, all types of HPTs were taken into account, i.e. testosterone-based, progesterone-based and unknown. It is not specified if Primodos was one of the HTPs. Therefore, it is questionable whether these results can be taken into account.

14. Rothman KJ, Fyler DC, Goldblatt A, Kreidberg MB (1979) Exogenous hormones and other drug exposures of children with congenital heart disease. Am J Epidemiol 109(4): 433-439.

This concerns a well-designed study. In this study, inadvertent exposure to combined oral contraceptives during, use of HPTs, use of progesterone and other drugs on risk of congenital heart disease are evaluated. No association was found. However, as the type of HPTs was not specified, it is questionable whether the results can be taken into account.

15. Ferencz C, Matanoski GM, Wilson PD, Rubin JD, Neill CA, Gutberlet R (1980) Maternal hormone therapy and congenital heart disease. Teratology 21(2): 225-239.

The study evaluated the effect of several exogenous hormone intakes during the first 3 months of pregnancy, including the use of HPTs (without specifying the type of HPT). The study failed to show an association between maternal hormone intake and congenital heart disease of the conotruncal type. The design is adequate. The study had two control groups, disease controls and healthy controls. In the meta-analysis, these control groups are combined, possibly introducing recall bias due to the healthy controls. However, as the type of HPTs was not specified, it is questionable whether the results can be taken into account.

16. Torfs CP, Milkovich L, Van den Berg BJ (1981) The relationship between hormonal pregnancy tests and congenital anomalies: A prospective study. Am J Epidemiology 113(5): 563-574

This article concerns a well-designed study. The HPTs included several preparations, both with and without estrogens and with different progestagens (the reviewers scored this as accurate exposure ascertainment). Primodos was among these HTPs. Both severe and mild anomalies were included. No association was found. However, as no separate results were provided for Primodos, the results cannot be taken into account.

17. Hadjigeorgiou E, Malamitsi-Puchner A, Lolis D, Lazarides P (1982) Cardiovascular birth defects and antenatal exposure to female sex hormones. Developmental pharmacology and therapeutics 5(1-2): 61-67.

This study evaluated the possible association between congenital malformations and the exposure to

oestrogens, progestogens, or a combination of oestrogen/progestogen. Little information on the methodology is provided, nevertheless, the reviewers gave this study a high NOS score of 7. However, the type of HPT (which progestogen) was not specified, as the exposure was divided into oestrogens, progestogens, or a combination of oestrogen/progestogen. Further, exposure was up to 4 months and from 4 months pregnancy. Therefore, it is questionable whether these results can be taken into account.

18. Michaelis J, Michaelis H, Gluck E, Koller S (1983) Prospective study of suspected associations between certain drugs administered during early pregnancy and congenital malformations. Teratology 27: 57-64

The study investigated the influence of anti-emetic drugs and sex hormones, including HPTs, among others Primodos. The study was performed adequately. Primodos users were specified, but exposure included in the meta-analysis referred to both tablets and injections. The use of a hormonal pregnancy test was not significantly associated with an increase of major malformations.

19. Polednak AP, Janerich DT (1983) Maternal characteristics and hypospadias: a case-control study. Teratology 28(1):67-73

Study which investigated the influence of hormonal exposure and maternal characteristics (maternal disease, menstrual cycle pattern, age, weight, age at menarche) during the first 3 months pregnancy and risk of hypospadias. Case definition poor (based on birth certificates only). The exposure information was collected by interview after birth and compared to healthy controls, which may have introduced recall bias. Maternal use of hormones in pregnancy did not differ significantly between cases and controls. It was not specified what type of hormone pregnancy-test was used. Therefore, it is questionable whether the results of this the study can be taken into account.

20. Lammer EJ, Cordero JF (1986) Exogenous sex hormone exposure and the risk for major malformations. JAMA 255(22): 3128-3132.

This case-control study of first-trimester sex hormone exposure among mothers of 1,091 infants with Down syndrome or at least one of 11 major malformations. For each malformation category, the infants with other malformations served as the control group. As the primary objective consisted of “all malformations”, this type of control is not suitable. Further, in this study multiple analyses were performed, only the statistically significant ones were reported and could be included in the meta-analysis. This selective reporting may be based on chance findings and may thus have introduced bias.

21. Sainz MP, Rodriguez Pinilla E, Martinez-Frias ML (1987) Progestogens and estrogens in high doses (hormone pregnancy tests): the risk of appearance of spina bifida with anencephaly. Medicina clinica (Barcelona) 89: 272-274

Controls were matched on sex and time of birth. Maternal history was not recorded which can be an important confounding factor. As it was not specified what hormonal pregnancy-test was used, it is questionable whether the results of this study can be taken into account.

22. Tümmler G, Rißmann A, Meister R, Schaefer C (2014) Congenital bladder exstrophy associated with Duogynon hormonal pregnancy tests-signal for teratogenicity or consumer report bias? Reprod Toxicol ;45:14-9

This study presents a case reports series compared to historical malformation data from a birth defect registry. As there is no adequate control group, the results from this study were not used in the meta-analysis.

After taking into account the above studies, the CHMP considered that the methodology of this meta-analysis is standard and in principle acceptable. However the performance of the study selection and data extraction were questioned. Two studies were unpublished and eight studies were published as a letter. The amount of information in these publications is too limited to properly ascertain the quality of the studies and the results presented. For the other studies, the quality ascertainment, using NOS scoring, appears to have been overestimated, e.g. case definition considered adequate and representative while the definition was based on birth certificates only (which in Jannerich, 1977 was confirmed in 70% of cases only); control selection was considered adequate, while the study had to rely on a mix of patients (diseased or healthy) to complete the set; several studies were scored as controlling for the most important factor, which was sex and date of birth only; exposure ascertainment by (late) interview, possibly influenced by recall bias, or involved studies were scored as adequate, while norethisterone/ethinylestradiol containing HPTs (Primodos/Diogynon) were mixed with other HPTs or other hormonal preparations used to prevent miscarriage.

Therefore, the quality of most studies used is questioned and, as a result, the conclusions of the meta-analysis cannot be considered reliable.

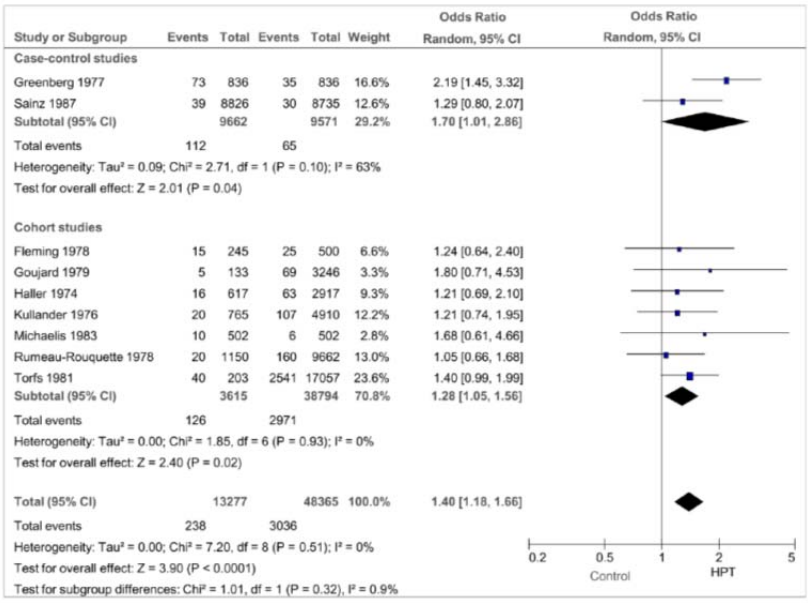
Association of exposure to HPT with the risks of malformations

Exposure to HPT and all congenital malformations

Nine studies, including 61,642 mothers of infants and 3,274 exposed to HPTs, examined the association in pregnancy with all congenital malformations. Two were case-control studies (Greenberg et al, 1977; Sainz et al, 1987) and seven were cohort studies (Fleming et al, 1987; Goujard et al, 1979; Haller, 1974; Kullander et al, 1976; Michaelis et al, 1983; Rumeau-Rouquette et al, 1978; Torfs et al, 1981) (Figure 2).

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; 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, 1974 and Fleming et al, 1978; 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).

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

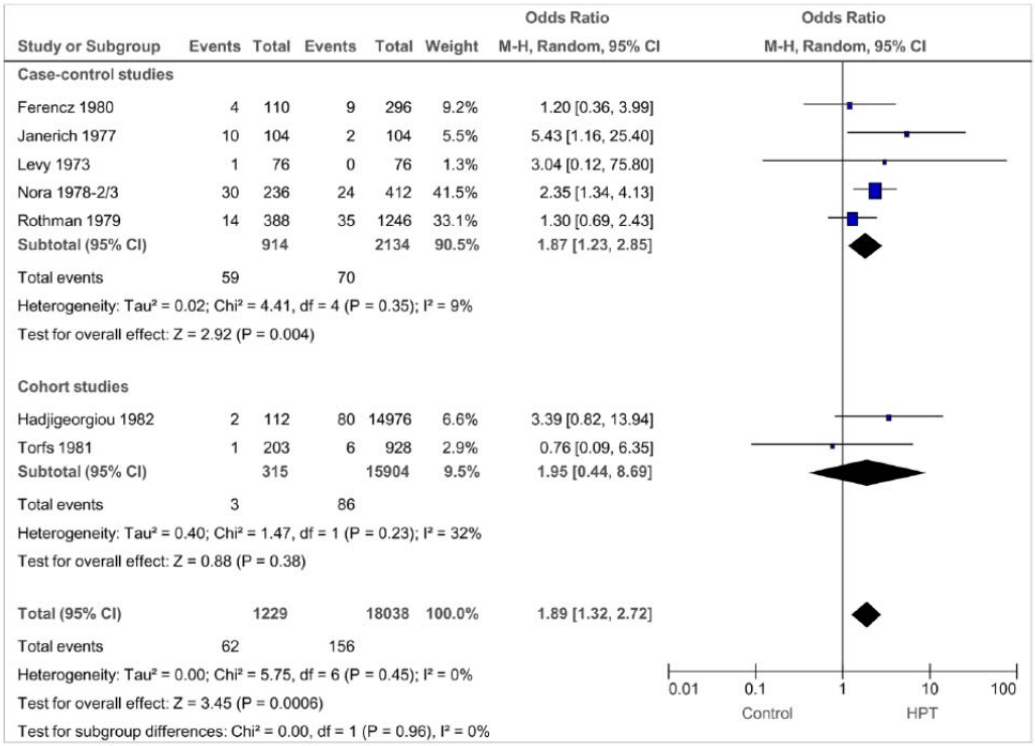


Of the studies included, one study was unpublished (Greenberg et al, 1977), and one concerned a letter (Haller, 1974), which contains limited information and can therefore not be taken into account. Further, for 3 studies (Goujard et al, 1979, Rumeau-Rouquette et al, 1978, Torfs et al, 1981) it was questionable whether the results could be taken into account. For further details, see assessment of individual studies above.

HPT and congenital heart malformations

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

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

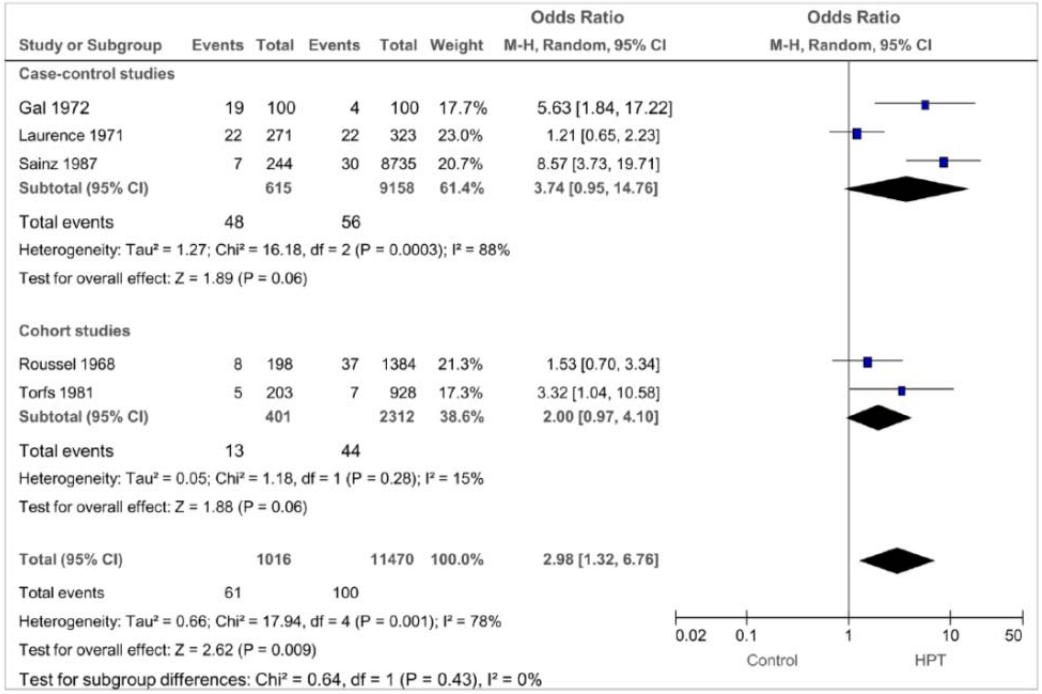


The CHMP noted that of the studies included, one study was published as a letter (Levy eta l, 1973), containing few details, which therefore cannot be taken into account. Further, for 6 studies (Ferencz et al, 1980, Janerich et al, 1977, Levy et al, 1973, Nora et al, 1978-2/3, Hadjigeorgiou et al, 1982, Torfs et al, 1981) it was questionable whether the results could be taken into account, as discussed above. In a post-hoc sensitivity analysis, removing one study that collected no confounding variables (Levy et al, 1973, 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 OR = 1.95 (95% CI 0.44 to 8.69; P = 0.38; I^2 = 32%). The meta-regression was not significant (P = 0.94).

Exposure to HPTs and nervous system malformations

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 et al, 1971; Sainz et al, 1987) and two cohort studies (Roussel, 1968; Torfs et al, 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; I^2 = 78%). In a post-hoc sensitivity analysis, removing the two studies that collected no confounding variables (Laurence et al, 1971; Roussel, 1968) 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, I^2 = 0%).

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

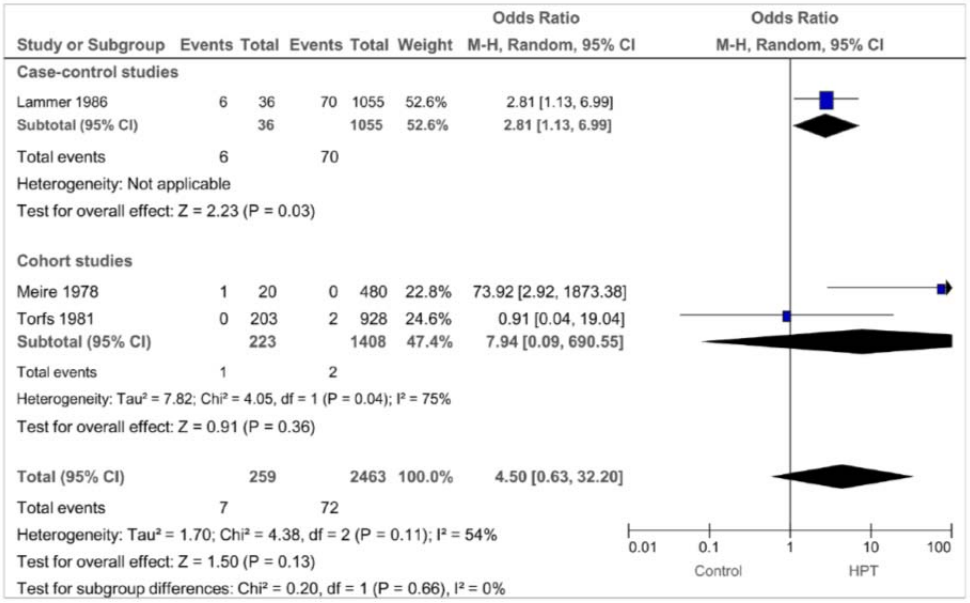


The CHMP noted that of the 5 studies included, one study (Laurence et al, 1971) was unpublished and two were published as a letter (Gal, 1972; Roussel, 1968) containing little detailed information, which therefore could not be taken into account. For the remaining two studies (Sainz et al, 1987, Torf et al, 1981), questions were raised on whether the results could be taken into account.

Exposure to HPTs and gastrointestinal malformations

Gastrointestinal malformations and exposure to oral HPTs were reported in three studies: a case-control study (Lammer & Cordero, 1986) and two cohort studies (Meire & Vuylsteek, 1978 and Torfs et al, 1981), providing data on 2,722 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; I^2 = 54%).

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

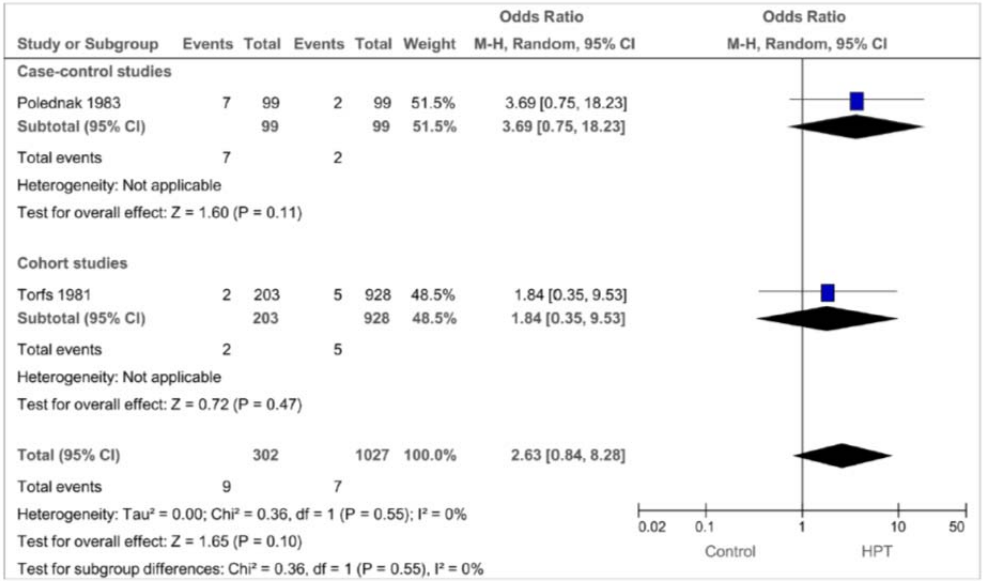


The CHMP noted that of the three studies included, one study (Meire & Vuylsteek, 1978) was published as a letter, containing limited information, and therefore could not be taken into account, and for two studies (Torfs et al, 1981, Lammer & Cordero, 1986) the results were questionable whether should be taken into account, as discussed above.

Exposure to HPTs and urogenital malformations

One case-control study (Polednak & Janerich, 1983) and one cohort study (Torfs et al, 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; I^2 = 0%) (Figure 6).

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

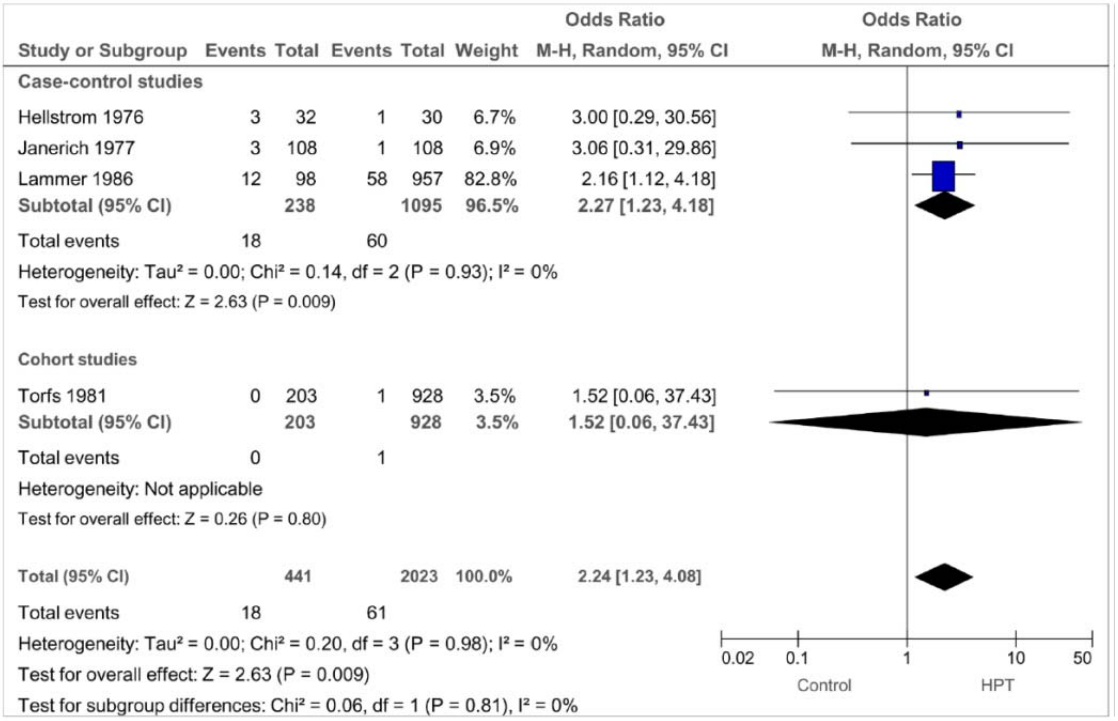


For the two studies included, the CHMP questioned the results and whether they could be taken into account, as discussed above.

Exposure to HPTs and musculoskeletal malformations

A relation between the exposure to oral HPTs and musculoskeletal malformations was reported in three studies: three case-control studies (Hellstrom et al, 1976; Janerich et al, 1977; Lammer & Cordero, 1986) and one cohort study (Torfs et al, 1981) (Figure 7), based on 2,464 women, with 79 exposed to HPTs. The pooled OR = 2.24 (95% CI 1.23 to 4.08; P = 0.009; I^2 = 0%). Removal of the zero study events (Torfs et al, 1981) did not affect this result.

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

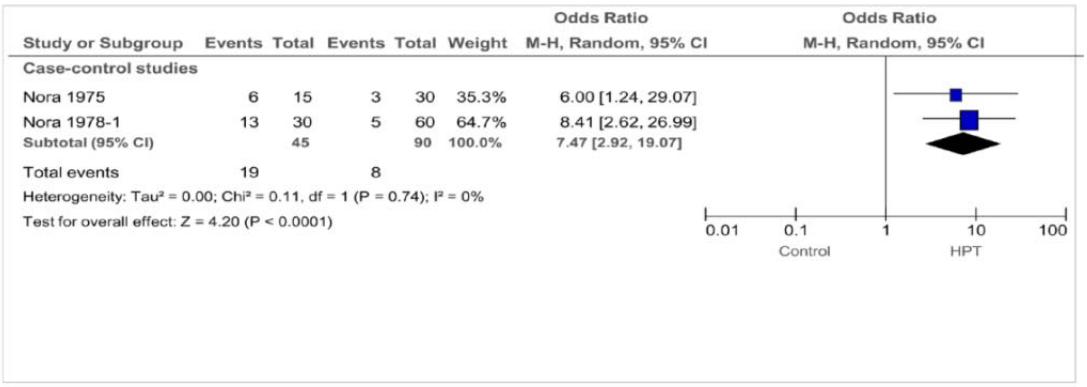


The CHMP noted also that of the four studies included, one study (Hellstrom et al, 1976) was published as a letter with very little detailed information, and therefore cannot be taken into account. For three studies (Janerich et al, 1977, Torfs et al, 1981, Lammer & Cordero, 1986) it was questionable whether the results could be taken into account.

Exposure to HPTs and VACTERL

The association of Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies, and Limb defects (VACTERL) syndrome with HPT exposure was reported in two case-control studies (Nora et al, 1978 (case-1) and Nora & 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; I^2 = 0%) (Figure 8).

Figure 8. Association of exposure to oral HPTs in pregnancy with VACTERL syndrome in the offspring.



The CHMP noted that for both studies included, it was questionable whether the results could be taken into account.

2.3. Discussion of the meta-analysis

Concerning the meta-analysis of Heneghan et al (2019), the CHMP was requested to provide an opinion on the suitability and robustness of the methodology, including the selection and application of the data quality score and to discuss if the findings would lead to any clinical implications. In addition, the CHMP was asked in case there are public health concerns with norethisterone and ethinylestradiol, whether these would need to be investigated further at Union level.

Regarding the suitability and robustness of the methodology, including the selection and application of the data quality score, the CHMP noted that most of the studies included in the meta-analysis are case control or cohort studies performed without the current scientific requirement level. The sample sizes used in the studies are possibly underpowered to the objectives of the studies; in some studies, the objectives could be better defined, and in general the methodologies to overcome the bias were not sufficient to reassure on the robustness of the results. The meta-analysis tried to compensate for the insufficiencies of the majority of the studies and seemed logical and possibly the only way to obtain some knowledge from these old studies.

The methodology of the meta-analysis seems adequate. As referred by the authors of this meta-analysis, the lack of adjustment for confounders, recall bias, problems in the ascertainment of the malformations and exposures, and the differences in severity of malformations with different risk estimates could introduce important bias to the results. However, this does not prevent that the meta-analysis has suitability to the objective of the work.

The Newcastle-Ottawa Scale (NOS), which was used for assessing the quality of non-randomised studies in meta-analyses, was not validated. Although NOS can show low agreement (Hartling et al, 2013) between authors, it can be useful for understanding the quality of the studies. The possibility of classifying the degree of bias by a quantitative scale can also be an advantage over the known STROBE statement (STrengthening the Reporting of OBservational studies in Epidemiology; www.strobe-statement.org) methodology, which is more descriptive and does not have quantification measures.

The meta-analysis did not present adjusted results but only crude data. This could limit the robustness of the results. However, most studies try to control for selection bias and confounding factors. Admittedly adjusted results could be somewhat different. However, the absence of adjusted results does not invalidate the published results.

Although, in general, the results give rise to an OR < 3 (cut-off point argued by Schaefer (2018) to consider teratogenicity), the CHMP noted that in all congenital malformations (OR 1.40), congenital heart malformations (OR 1.89), nervous system malformations (OR 2.98), and musculoskeletal malformations (OR 2.24), the confidence intervals are reassuring in that the OR is statistically significant, and these results cannot be completely discarded, even considering the heterogeneity and the small number of cases included in some studies. VACTERL syndrome (OR 7.47) is a peculiar condition, where the few cases could explain a higher upper limit of confidence intervals, but the lower limit is clearly above 2. Of the 12 studies with low risk of bias, six studies have an association with congenital malformations and the use of HPTs.

Overall, and although the CHMP agreed with some points and doubts in the Schaefer's position (2018) (namely the requirements of the biological plausibility for a teratogenic effect), the CHMP considered that the results of Heneghan and colleagues (2018) do not exclude the possibility of an association between malformations and hormonal pregnancy tests. Comparatively with the EWG "Report of the commission on human medicines expert working group on hormone pregnancy tests", these results did not add new information.

Regarding the question on any clinical implications, the CHMP noted that in this meta-analysis only data from HPTs was included. Cases with exposure to oral contraceptives or hormonal use with therapeutic intention were excluded from the meta-analysis. Therefore, the results can only have more direct implication in relation to the use of HPT in pregnancy. Although the oestrogen constituent in hormonal pregnancy test is in general ethinylestradiol, the progestogen component in some studies is mainly (but not restricted to) norethisterone. Nevertheless, ethinylestradiol and norethisterone are also constituents of oral contraceptives.

Many gaps of knowledge exist in the assessment of a claimed teratogenic effect of the use of HPTs in pregnancy. These gaps are even higher when an extrapolation is tempted to oral contraceptives. Not only the data in these studies with oral contraceptives is scarce, but also, the time and dose with HPTs are substantially different from those present in oral contraceptives.

In HPTs (Primodos/ Duogynon) the dose of ethinylestradiol was 0.02 mg and the dose of norethisterone was 10 mg. In oral contraceptives, the dose per tablet of the two substances is much lower, i.e. either 0.02 mg or 0.03 mg for ethinylestradiol, and 1.0 mg or 1.5 mg for norethisterone. A full dosage of 0.04 mg/20 mg of ethinylestradiol / norethisterone in HPTs (2 days) must be put against the usual dose of ethinylestradiol / norethisterone in a menstrual cycle when used as oral contraceptive (ethinylestradiol: 0.42 mg or 0.63 mg / norethisterone 21 mg or 31.5 mg) (in the total of 21 days of treatment). The EWG "Report of the commission on human medicines expert working group on hormone pregnancy tests" has made estimates about blood levels of norethisterone and of ethinylestradiol. After a single dose of Primodos, the maximum free concentration of maternal blood levels of norethisterone were estimated to be of about 2 ± 0.8 ng/ml, and the maximum free concentrations of maternal blood levels of ethinylestradiol of about 1.9 ± 0.6 pg/ml. It is also estimated the maximal total human placental serum concentrations and the maximal total human fetal concentrations of norethisterone, respectively, of about 15 ± 5 ng/ml and of about 10 ± 4 ng/ml. For ethinylestradiol, the estimate for the total human foetal plasma concentrations is of about 4 ± 1 pg/ml, and for free human foetal plasma concentrations of about 0.12 ± 0.04 pg/ml. It is of note, however, that for ethinylestradiol the EWG Report estimates about 20% accumulation (half-life 18 ±4.7 h) in maternal blood levels when dosed for 21 days. The EWG Report also suggested that although a dose of 5 mg of norethisterone (half-life 5-12h) is expected to have a limited accumulation with daily dosing it seems that co-administration with ethinylestradiol could lead to additional accumulation of norethisterone in women taking a daily dose. Nevertheless, as the combination of the two hormones, and especially with such high dose of norethisterone is not used for oral contraception the clinical relevance of these data is questioned.

In the studies included in the meta-analysis, a look to the cases reported with exposure to oral contraceptives should be considered. In at least 6 studies, a separation of oral contraceptives cases was found. In particular, the study by Janerich and colleagues (1974), raised the hypothesis of a positive association between oral contraceptives and limb defects in male infants; the author corroborate this hypothesis with an increase of 33% in limb defects (1968-1973 vs 1963-1967) in New York area, which in the author´s opinion could be due to an increase in the use of oral contraceptives.

Nevertheless, more recent studies show different results. The meta-analysis of Bracken (1990) of prospective studies did not find an association between oral contraceptives and birth defects (relative risk (RR): 0.99, 95%CI, 0.83-1.19, for all malformations; RR 1.06, 0.72-1.56 for congenital heart defects; RR 1.04-0.30-3.55 for limb defects). Although, Li and colleagues (1995) suggested an association between oral contraceptive use after conception and the risk of congenital urinary tract anomalies, others authors, (Waller et al, 2010) did not find a link between oral contraceptives in pregnancy and major structural birth defects in offspring. Similarly, the cohort study (Charlton et al, 2016) with 880,694 live births from Danish registries between 1997 and 2011 did not found an increase in the prevalence of major birth defects among women with recent oral contraceptive exposure before pregnancy (prevalence odds ratio 0.98 (95% CI 0.93-1.03)) or use after pregnancy onset (OR 0.95 (0.84-1.08)), compared with the reference group. No increase in prevalence of any birth defect by subgroup (e.g. limb defects) was observed either. In summary, overall data available on oral contraceptive use and birth defects did not show any increased risks.

In conclusion, the outcome of the review and meta-analysis performed by Heneghan and colleagues (2018) do not exclude an absence of teratogenic effect due to HPTs use in pregnancy. However, methodological flaws due to general poor quality of the studies involved cast doubts on the validity of the results of the meta-analysis.

The available data did not raise any new concerns of potential teratogenic effect associated with the use of oral contraceptives during pregnancy. It is also highlighted that the product information for oral contraceptives includes already a contraindication in pregnancy.

As this meta-analysis did not add new information to what is already known, the CHMP concluded that no regulatory actions were deemed necessary.

The CHMP also considered whether there are public health concerns norethisterone and ethinylestradiol which may be of Union interest and whether these need to be investigated further.

Based on the epidemiological data available, the EWG “Report of the commission on human medicines expert working group on hormone pregnancy tests” concludes that while the quality of the available epidemiological evidence was generally very limited, no strong associations were found between the use of HPTs, including Primodos/Duogynon, during pregnancy and any single anomaly, or any pattern of anomalies. The weak associations that were observed (congenital heart defects, limb reduction defects, and oesophageal atresia) could have occurred by chance or confounding.

The study of Henegahn and colleagues (2018) did not provide additional valid information than the one provided on the EWG “Report of the commission on human medicines expert working group” on hormone pregnancy tests. The CHMP considered that in respect to HPTs and the use in pregnancy the EWG Report conclusions remain valid. No further regulatory actions were deemed necessary by the CHMP.

In addition, the CHMP considered that this meta-analysis did not provide additional valid information than the ones provided on the previous published studies (Bracken, 1990; Waller et al, 2010; and Charlton et al, 2016) on the topic of potential association between birth defects and contraceptives use during the early phases of pregnancy. No further regulatory actions are deemed necessary in this aspect as well.

3. Overall Conclusions

Primodos contains 0.02 mg ethinylestradiol and 10 mg norethisterone. One tablet was to be taken for 2 days. Both components of Primodos, i.e. norethisterone and ethinylestradiol, are widely used.

The synthetic oestrogen ethinylestradiol is the oestrogen component in all combined hormonal contraceptives (CHCs) that are available on the EU market, except for two recently registered CHCs which contain estradiol valerate. The dose of ethinylestradiol in CHCs ranges between 0.015 – 0.035 mg. Contrary to the oestrogen component, the progestogen component varies in CHCs. There are several different synthetic progestogens including norethisterone. The norethisterone dose in CHCs ranges between 0.5 and 1 mg.

Norethisterone and several other progestogens are also used as monotherapy or in combination with an oestrogen for the treatment of menstrual bleeding irregularities and endometriosis. Norethisterone is used for menstrual irregularities in a dose range of 5 -15 mg and as monotherapy for endometriosis in a dose of 10-20 mg/day.

Although the exposure to Primodos is very limited and not comparable with the daily exposure to COCs containing norethisterone or norethisterone used as monotherapy, reference is made to information in published literature on exposure of COCs during early pregnancy.

COCs are the most frequently used contraceptive method in the Western world. Today, there are about 100 million women who use a COC around the world. Despite COCs being very effective when used according to the regimen, it has been estimated that about 9% of women nevertheless will get pregnant in their first year of use (Trussell, 2011) because of missed pills, possible interactions with concomitant medication, disease, or failure of the method. In such situations the woman might be unaware of the pregnancy and the foetus could be inadvertently exposed to exogenous hormones of the COC.

There are no specific studies per type of combined COC, but there are a number of studies which investigated inadvertent exposure by COCs in general during early pregnancy and the risk of major birth defects, including a meta-analysis (Bracken, 1990). The meta-analysis, based on 12 prospective studies, did not find an association. The largest and most recently published study is a very large Danish prospective cohort study in which it was investigated whether oral contraceptive use around the time of pregnancy onset is associated with an increased risk of major birth defects. All oral hormonal contraceptives were taken into consideration, i.e. combined oral contraceptives, progestogen-only contraceptives and emergency contraceptives. The data on oral contraceptive use and major birth defects were collected among 880,694 live births from Danish registries between 1997 and 2011. The main outcome measure was the number of major birth defects throughout one year follow-up (defined according to the European Surveillance of Congenital Anomalies classification). Based on the results, the authors concluded that “oral contraceptive exposure just before or during pregnancy does not appear to be associated with an increased risk of major birth defects” (Charlton et al, 2016). This study did not contain specific information on the active ingredients of the oral hormonal contraceptives taken into account, but as norethisterone containing COCs are on the market since the 70-ties it is likely included.

Based on above clinical evidence, the product information of COCs contains common information in section 4.6 of the summary of product characteristics that extensive epidemiological studies have revealed neither an increased risk of birth defects in children born to women who used COCs prior to pregnancy, nor a teratogenic effect when COCs were taken inadvertently during pregnancy.

Regarding the meta-analysis under review, the CHMP concluded that although the methodology of this meta-analysis is considered standard and appears to be acceptable, the performance of the study selection and data extraction are questioned.

The meta-analysis includes studies reported between 1968 and 1987. Two studies were unpublished and eight studies were published as a letter. The amount of information in these publications is too limited to properly determine the quality of the studies and the results presented.

For the other studies, it appears that the quality ascertainment applied lacks sufficient details. It is unclear what is considered a "relevant control group". Furthermore, assessing the quality of the studies can be subjective and details on the exact criteria in applying the NOS score were not provided. For example, in the determination whether or not the "study controls for the most important factors", it is unclear whether the reviewers determined the most important factors in advance, or relied on statements about importance in the original paper. Finally, the quality ascertainment using NOS scoring, appeared to have been overestimated while the study had to rely on a mix of patients (diseased or healthy) to complete the set. Several studies were scored as adequately controlling for the most important factor, while this was based on sex and date of birth only; exposure ascertainment by (late) interview, possibly influenced by recall bias, or studies where norethisterone/ethinylestradiol containing HPTs (Primodos/Diogynon) are mixed with other HPTs or hormonal preparations used to prevent miscarriage.

Therefore, the quality of most studies used is questioned and, as a result, the conclusions of the meta-analysis cannot be considered reliable. Due to the multiple limitations of the meta-analysis study, the results described in this manuscript cannot be used to further expand clinical knowledge. The results of this meta-analysis, thus, have no clinical implications. As a consequence, the conclusion that current clinical data available do not support a signal of teratogenicity of a combination of norethisterone/ethinylestradiol remains valid. The CHMP therefore did not recommend any further regulatory actions based on the above data.

4. References

Bracken MB. (1990) Oral contraception and congenital malformations in offspring: a review and meta-analysis of the prospective studies. Obstet Gynecol. Sep;76(3 Pt 2):552-7.

Charlton BM, Mølgaard-Nielsen D, Svanström H, statistician, Wohlfahrt J, Pasternak B, Melbye M. (2016) Maternal use of oral contraceptives and risk of birth defects in Denmark: prospective, nationwide cohort study. BMJ 2016;352:h6712.

Ferencz C, Matanoski GM, Wilson PD, Rubin JD, Neill CA, Gutberlet R (1980) Maternal hormone therapy and congenital heart disease. Teratology 21(2): 225-239.

Fleming DM (1978) Abnormal outcome of pregnancy after exposure to sex hormones. Personal communication. (Provided by Bayer).

Gal I (1972) Risks and benefits of the use of hormonal pregnancy test tablets. Nature 240: 241-242.

Greenberg, G, Inman WHW, Weatherall JAC, Adelstein AM, Haskey JC (1977) Maternal drug histories and congenital anomalies. BMJ 2: 853-856.

Goujard J, Rumeau-Rouquette C, Saurel-Cubizolles MJ (1979) Hormonal tests of pregnancy and congenital malformations. Journal de gynecologie obstetrique et biologie de la reproduction 8: 489-196.

Hadjigeorgiou E, Malamitsi-Puchner A, Lolis D, Lazarides P (1982) Cardiovascular birth defects and antenatal exposure to female sex hormones. Developmental pharmacology and therapeutics 5(1-2): 61-67.

Haller J (1974) Hormontherapie wahrend der graviditat.Deutsches Arzteblatt 14: 1013-1015.

Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, Dryden DM (2013). Testing the

Newcastle-Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol 66(9):982-993.

Hellstrom B, Lindsten J, Nilsson K (1976) Prenatal sex-hormone exposure and congenital limb reduction defects. Lancet 2:372-3.

Janerich DT, Dugan JM, Standfast SJ, Strite L (1977) Congenital heart disease and prenatal exposure to exogenous sex hormones. BMJ 1(6068):1058-60.

Kullander S, Kallen B (1976) A prospective study of drugs and pregnancy. Acta Onstet Gynecol Scand 55: 221-224.

Lammer EJ, Cordero JF (1986) Exogenous sex hormone exposure and the risk for major malformations. JAMA 255(22): 3128-3132.

Laurence M, Miller M, Evans K, Carter C (1971) Hormonal pregnancy tests and neural tube malformations. Nature 233: 495-496.

Levy EP, Cohen A, Fraser FC (1973) Hormone treatment during pregnancy and congenital heart defects. Lancet 1(7803): 611.

Li DK, Daling JR, Mueller BA, Hickok DE, Fantel AG, Weiss NS (1995) Oral contraceptive use after conception in relation to the risk of congenital urinary tract anomalies. Teratology 51(1):30-36.

Meire F, Vuylsteek K (1978) Apr 1;1(6116):856. Continued use of hormonal pregnancy tests. BMJ 1:856.

Michaelis J, Michaelis H, Gluck E, Koller S (1983) Prospective study of suspected associations between certain drugs administered during early pregnancy and congenital malformations. Teratology 27: 57-64.

Nora AH, Nora JJ (1975) A syndrome of multiple congenital anomalies associated with teratogenic exposure. Arch Envir Heatth. Vol 30.

Nora JJA, Nora H, Blu J, Ingram J, Fountain A, Peterson M, Lortscher RH, Kimberling WJ (1978) Exogenous progestogen and estrogen implicated in birth defects. JAMA 240(9): 837-843.

Polednak AP, Janerich DT (1983) Maternal characteristics and hypospadias: a case-control study. Teratology 28(1):67-73.

Rothman KJ, Fyler DC, Goldblatt A, Kreidberg MB (1979) Exogenous hormones and other drug exposures of children with congenital heart disease. Am J Epidemiol 109(4): 433-439

Roussel (1968) GP Survey – An investigation into the effects of oral pregnancy tests on the incidence of CNS malformations. Unpublished.

Rumeau-Rouquette et al (1978) Malformations congenitales risques perinatales. L'institut National de la Sante et de la recherce medicale (INSERM).

Sainz MP, Rodriguez Pinilla E, Martinez-Frias ML (1987) Progestogens and estrogens in high doses (hormone pregnancy tests): the risk of appearance of spina bifida with anencephaly. Medicina clinica (Barcelona) 89: 272-274.

Schaefer C, (2018) unpublished letter on the Heneghan et al 2018 (version 1) publication.

Torfs CP, Milkovich L, Van den Berg BJ (1981) The relationship between hormonal pregnancy tests and congenital anomalies: A prospective study. Am J Epidemiology 113(5): 563-574.

Trussell J. (2011) Contraceptive failure in the United States. Contraception. May;83(5):397-404.

Tümmler G, Rißmann A, Meister R, Schaefer C (2014) Congenital bladder exstrophy associated with Duogynon hormonal pregnancy tests-signal for teratogenicity or consumer report bias? Reprod Toxicol; 45:14-9.

EWG, expert working group, UK Government (2017)"Report of the Commission on Human Medicines' Expert Working Group on Hormone Pregnancy Tests" - https://www.gov.uk/government/publications/report-of-the-commission-on-human-medicines-expert-working-group-on-hormone-pregnancy-tests

Waller DK, Gallaway MS, Taylor LG, Ramadhani TA, Canfield MA, Scheuerle A, Hernández-Diaz S, Louik C, Correa A (2010) National Birth Defects Prevention Study. Use of oral contraceptives in pregnancy and major structural birth defects in offspring. Epidemiology. 21(2):232-239.