REPORT OF THE COMMISSION ON HUMAN MEDICINES EXPERT WORKING GROUP ON HORMONE PREGNANCY TESTS

October 2017

REPORT OF THE COMMISSION ON HUMAN MEDICINES EXPERT WORKING GROUP ON HORMONE PREGNANCY TESTS

CONTENTS

Foreword

Hormone Pregnancy Tests (HPTs) were medicines available in the 1950s to 1970s that contained sex steroid hormones, most commonly an estrogen and a progestogen, and were used to diagnose pregnancy or treat a disorder of menstruation called secondary amenorrhoea. Whilst it seems anathema now to give or take a drug to diagnose pregnancy, HPTs were widely used in the past in the UK before the development of the simple over-the-counter pregnancy tests we currently have today. Hormone Pregnancy Tests have been in the spotlight for many years and have been the subject of long term debate because of a possible association between their use and harm in a developing pregnancy. It is largely due to the determination of members of the 'Association for Children Damaged by HPTs', who were given an HPT for diagnosing pregnancy and whose lives have been impacted by an adverse pregnancy outcome, that this scientific review has been undertaken.

To ensure this review was able to fulfil its main objective – to investigate all available evidence on the possible association between exposure in pregnancy to HPTs and adverse outcomes in pregnancy – a full panel of experts (Expert Working Group) in teratology, embryology, clinical genetics, neonatology and epidemiology, amongst others, was convened on behalf of the Commission on Human Medicines. The science behind a potential association between HPTs and adverse pregnancy outcomes was considered in depth and rigorously evaluated by the group members. At all times, the scientific review process was designed to protect it from any bias relating to the pharmaceutical industry, politics and media pressure.

The social, medical and regulatory environment at the time HPTs were available was very different from that which we take for granted today. Although the EWG had a clear scientific remit, we recognised how important it was to familiarise ourselves with some of the broader, non-scientific issues from that time to provide the context and perspective for the review. We therefore wanted to make sure we had complete access to the substantial amount of information gathered for the review – scientific and non-scientific – while being scrupulous about examining the scientific evidence objectively.

One point on which the members were unanimous from the outset was that this should be a forward-looking review with a focus on making sure that the EWG's recommendations are relevant today and in the future in helping to improve the safe use of medicines in pregnancy.

We are reassured that substantial and important changes have taken place from the regulatory perspective since HPTs were available in the UK; nevertheless, we consider that valuable steps could be taken to strengthen further the systems currently in place and that safety monitoring should be a shared industry and public responsibility. We recognise that some of these recommendations will need to be funded but we strongly believe they will have the potential to improve how we detect, evaluate and communicate safety concerns relating to use of medicines in early pregnancy.

Finally, the EWG felt strongly that all the evidence that had been gathered, together with the assessments of those data and the minutes of the EWG's meetings should be published at the end of the process to allow full public scrutiny and ensure complete transparency of the scientific process.

In conclusion, we hope this report will be recognised as the most comprehensive and up-to-date review of the available data thus far and that its legacy will be in helping safeguard pregnant women and their babies in the future.

Acknowledgements

As Chair of the EWG, I feel privileged to have worked with so many exceptional people who have bought exemplary professionalism, scientific rigour, commitment and compassion to this review. The EWG members took every opportunity to explore the evidence on HPTs and adverse pregnancy outcomes, and their desire to undertake the best possible review led to fresh approaches to analysing the data. Consequently, the EWG members are confident in their final conclusions and assured that they have done all they could with the available data.

The importance of the EWG's findings to all those whose lives or those of their families have been affected by adverse pregnancy outcomes after taking HPTs made the role of the lay members and observers on the EWG of critical importance. The patient voice was present throughout and hearing at first hand the personal experiences of the families involved was deeply moving. I thank them all for having the courage to tell us their stories in difficult circumstances.

Finally, on behalf of the EWG, I should like to thank the secretariat which MHRA provided to support our work which was of the highest standard and responsive in taking any extra steps requested by the group to ensure the validity and balance of the scientific process.

Dr Ailsa Gebbie, MB ChB FRCOG FRCP(Edin) FFSRH - Chair

October 2017

Lay Summary

Introduction

This section provides a lay summary of the report of the Commission on Human Medicines Expert Working Group on Hormone Pregnancy Tests (HPTs). It states the questions that the Group set out to answer and explains how this was achieved, what evidence was reviewed and, in the light of this, what conclusions were drawn.

Each section in this summary signposts the relevant chapter in the main report where interested readers can go to find out more. Scientific terms are explained within the glossary at the end of the report on page 111.

Background (Chapter 1)

In the UK, Hormone Pregnancy Tests (HPTs) first became available for diagnosing pregnancy in the 1950s. Hormone Pregnancy Tests contained synthetic versions of two hormones, alone or in combination, which are naturally found in the body. These hormones – progesterone and estrogen – are involved in the sexual development of women and in normal pregnancy and are therefore referred to as sex hormones. The most widely used HPT in the UK was called Primodos. Primodos contained a synthetic version of progesterone called norethisterone acetate[1] (10mg per tablet) and a synthetic version of estrogen called ethinylestradiol (0.02mg per tablet) and one tablet was taken on each of two consecutive days. This led to a withdrawal bleed a few days later in those who were not pregnant. Primodos was also used to treat disorders of menstruation.

Between the 1950s and 1978, when Primodos was withdrawn from the market in the UK, a number of studies were published which investigated a possible link between women being given an HPT to diagnose pregnancy and the occurrence of a range of congenital anomalies in the offspring. Although there was never any reliable evidence that HPTs were unsafe, concern about this issue, coupled with the development of better pregnancy tests meant that a series of precautionary actions were taken to restrict the use of HPTs to treating disorders of menstruation and to prevent their use in women who were pregnant. However, evidence suggested that these restrictions were not always being adhered to, and because the alternative non-hormonal pregnancy tests were becoming more widely available, the products were withdrawn from the market by the manufacturers. Whether these precautionary actions were sufficiently timely became a subject of controversy.

A campaign group, the 'Association for Children Damaged by Hormone Pregnancy Tests' (the Association), took the manufacturer of Primodos, Schering (now Bayer), to court in the 1970s. The case was discontinued at the request of the plaintiffs, with the judge stating that "the evidence would have to be very strong for a new trial"[2]. The body of information subsequently accrued by the Association and other campaigners since then, led to a Parliamentary debate in 2014 during which the then Minister for Life Sciences, George Freeman MP, stated that he would instruct that all relevant documents held by the Department of Health be released. In addition, he determined that an independent review of the papers and all the available evidence was justified.

[1] Norethisterone acetate in Primodos is broken down by the body to norethisterone. For simplicity, this lay summary refers to 'norethisterone' hereafter.

[2] New Scientist 8 July 1982, [extract of Judgment]

The purpose of a review would be to ascertain whether the totality of the available data, on balance, support a causal association between use of an HPT by the mother and adverse pregnancy outcomes (as was the assertion of the Backbench Business Committee made during that debate). Alternatively, whether the anomalies could have been due to chance alone or due to other factors (that is, the evidence does not support a causal association).

An Expert Working Group (referred to as the EWG) of the Commission on Human Medicines was established in October 2015 to conduct the review with the benefit of up-to-date scientific expertise.

The EWG was subject to a strict conflict of interest policy and comprised experts from a broad range of specialisms together with lay representation. The terms of reference of the EWG, were as follows:

- To consider all available evidence on the possible association between exposure in pregnancy to hormone pregnancy tests (HPTs) and adverse outcomes in pregnancy (in particular congenital anomalies, miscarriage and stillbirth) including consideration of any potential mechanism of action.
- To consider whether the EWG's findings have any implications for currently licensed medicines in the UK or elsewhere.
- To draw any lessons for how drug safety issues in pregnancy are identified, assessed and communicated in the present regulatory system and how the effectiveness of risk management is monitored.
- To make recommendations.

This report summarises the scientific evidence that was considered by the EWG, its conclusions on the evidence and its recommendations.

Scope of the review (Chapter 1)

Different HPT products contained different sex hormones, all of which have different actions in the body. The EWG focussed its review on the two components of Primodos, norethisterone and ethinylestradiol, separately and in combination.

To set the scene, the EWG considered the historical and scientific context from the time that HPTs were used, before going on to review how norethisterone and ethinylestradiol act in the body and what evidence there is from studies conducted in animals and in women for a possible association between their use to diagnose pregnancy and the development of congenital anomalies in the infant, or miscarriage of the pregnancy. Because the medical, scientific and regulatory landscape has changed so much since HPTs were on the market the EWG went on to consider how medicines are licensed today and what processes are in place to monitor, detect, evaluate and act on any safety concerns that are observed in pregnancy, how these are communicated to prescribers, women and patients and what checks are in place to make sure important new advice is acted on.

To ensure the review was comprehensive, published and unpublished evidence was gathered from a number of different sources including: pharmaceutical companies whose predecessors used to market HPTs; medicines regulators in other countries; the UK National Archives; archives in Berlin (the Landesarchiv Berlin); anyone who considered they had any relevant information following a public call for information; and from the published literature.

The EWG also heard evidence from several scientific experts, and members of The Association were invited to relate their experiences to the EWG.

Historical perspective (Chapter 2)

Unlike today, when home pregnancy kits can be bought in any chemist or supermarket and are used by many millions of women worldwide, in the 1950s and '60s testing for pregnancy was not common, and was mainly intended for women who were thought to be more at risk of having a difficult pregnancy. The alternative to HPTs was a physical examination by the doctor or a relatively lengthy and expensive laboratory test involving a specific type of toad. At the time of their introduction, HPTs were therefore considered to have several advantages over the alternatives and many HPT products rapidly became available. In the UK, it is estimated that well over a million women received an HPT between 1966 and 1978.

When HPTs first became available, pharmaceutical companies were not legally required to ensure that marketed medicines met appropriate standards of safety and efficacy and it was not until the thalidomide tragedy in 1961 that a framework was put in place to support the regulation of medicines. In the UK, this ultimately led to the introduction of the Medicines Act 1968, which came into force in 1971.

Sex hormones (Chapter 2)

Estrogens and progesterone are naturally occurring sex hormones that prepare the uterus (womb) for pregnancy and act to sustain pregnancy after implantation of a fertilised egg. High levels of progesterone dominate throughout pregnancy and have many diverse functions, with estrogens also contributing an important function. It has long been considered that low levels of progesterone can lead to miscarriage and so many women have been given high levels of synthetic progesterone (progestogens) for prolonged periods of their pregnancy in an effort to stop them from losing their baby.

No currently licensed medicines in the UK contain the same combination and dosage of progestogens and estrogens as were present in HPTs. However, varying combinations and dosages of similar progestogens and estrogens are used daily by many millions of women for contraception (the "Pill") and for treatment of gynaecological conditions.

Congenital anomalies (Chapter 2)

Information from Europe and the USA suggests that 24 to 40 babies out of every 1 000 are born with a major congenital anomaly (2.4% to 4%). Congenital heart defects are the most common anomaly, followed by limb defects, anomalies of the urinary system and nervous system defects. Today, the cause of the majority of all congenital anomalies remains unknown. Known causes include genetic causes (inherited or occurring without prior family history) and certain medications and medical conditions in the mother. It is likely that many congenital anomalies are caused by many environmental and genetic factors acting together.

The third to eighth weeks of human development (equivalent to the fifth to tenth weeks of pregnancy) are critically important for normal development of the baby as this is when most of its major organs and body systems are formed or become apparent. This time is called the period of organogenesis and is thought to be when the developing baby – the fetus – is

most sensitive to genetic or environmental factors and when most structural congenital anomalies are induced.

Evidence for a possible association between use of HPTs in pregnancy and congenital anomalies (Chapters 4 and 5)

Before considering the evidence for a possible association between using Primodos and having a baby with a congenital anomaly, the EWG set out the key conditions that would need to be met for this to have been possible:

1. Primodos must be administered during the critical period of fetal development
2. It must be able to cross the placental barrier between the mother and the fetus
3. The fetus must have estrogen and progesterone receptors that are capable of binding to the hormonal components of Primodos
4. These receptors must be present during the critical period of fetal development and be able to bind to, and be activated by, the drug
5. The drug should be at a sufficiently high concentration to cause a biological effect.

As well as norethisterone and ethinylestradiol in Primodos, natural progesterone and estrogen hormones produced by the mother during pregnancy act through the same receptors and may also reach the developing fetus. Little is currently known about the effects of these maternal hormones on the fetus, especially during early pregnancy, but norethisterone and ethinylestradiol from Primodos would be expected to mimic them.

To assess the five points above, evidence from laboratory studies, studies in animals and studies in humans was evaluated. However, limited evidence was found and only a very small amount of the data came from studies in pregnant women. As a result, a number of assumptions had to be made based on knowledge of how levels of norethisterone and ethinylestradiol change in the blood of women given doses broadly similar to those in Primodos.

Based on the available evidence, the EWG concluded that small amounts of norethisterone and ethinylestradiol could have reached the fetus as the result of taking Primodos tablets, for two days during the first trimester of pregnancy but that it was unlikely to have had an effect on the developing fetus, via a direct pharmacological action. Any action of these hormones would require the expression of functional receptors and would undoubtedly be affected by the relatively high concentrations of the very similar natural maternal estrogen and progesterone in early pregnancy.

- Evidence for a direct teratogenic effect

The EWG then examined the available evidence from studies in animals to determine whether norethisterone or ethinylestradiol, or both, can act as a teratogen (disturb the development of the embryo or fetus) and cause a malformation.

Reproductive toxicity studies in animals look specifically at whether a drug can have an adverse effect on the development of the young. A number of published studies (a total of 38) and unpublished studies (a total of 44) were therefore evaluated to see if there was any evidence for a teratogenic effect with norethisterone or ethinylestradiol. Preliminary findings from researchers at Aberdeen University were also presented to the EWG but at the time of writing remain unpublished.

Thirty-seven studies evaluated the effect of norethisterone and ethinylestradiol in combination: six in mice; 11 in rats; 12 in rabbits; 1 in guinea pigs; and seven in non-human primates (monkeys and baboons). Doses ranged from those roughly equivalent to the dose of Primodos in humans up to doses that were approximately 9 000 times the dose in humans. This wide dose range reflects partly the different sensitivities of the different animals to the effects of the sex hormones (the reason why studies need to be done in several different species) and partly the need to use high enough drug concentrations to cause an effect which has teratogenic potential. The numbers of offspring in all studies identified were relatively small and so the tests were not very sensitive for detecting small increases in rare events or showing whether any observed malformations that had occurred were spontaneous or due to the drug. Comparison of any potentially drug-related effects with a large pool of the laboratory's historical control data in a group of animals not given HPTs is therefore important. In many of the studies reviewed here, historical control data were not available and so assessing whether any random events that were observed were related to the drug was difficult.

Consistent findings in mice, rats, guinea pigs and rabbits were shown in these studies. Malformations of the genital tract or genital organs and the abnormal development of male sexual characteristics in a female (known as virilisation) were reported in some rats, mice and non-human primates that were exposed to norethisterone and ethinylestradiol during the period of sexual differentiation late in organogenesis. These effects reflect the known pharmacological action of these compounds and so the review focused on anomalies of non-reproductive tissues, for which there has been scientific uncertainty over evidence for a causal association.

Most studies found very little evidence for an increased risk of malformations in other (non-genital) organs apart from the occurrence of random events, such as cleft palate, absence of one or both eyes (anophthalmia) and a disorder in which the brain is outside of the skull (exencephaly) in a very small proportion of the young. In one study in mice there was evidence that giving a combination of the two hormones throughout the whole period of development at doses approximately 30 times higher than those found in HPTs was associated with an increase in malformations of the chest and mid-body. A similar increase in such malformations was not seen in rats, rabbits or non-human primates. While the effect seen in the mouse was considered to be related to the drug, the effect would therefore seem to be species specific, with the mouse being the sensitive species.

The EWG agreed on the different levels of evidence that would be needed to show that use of HPTs could cause malformations in the offspring i.e. a causal association (see Section 3.3.4 of the main report).

On this basis, the EWG concluded that the totality of the available data from studies in rats, rabbits, and non-human primates did not support a causal association between administering norethisterone and ethinylestradiol at the doses and durations found in Primodos and the development of malformations in non-sexual tissues of the offspring.

- Evidence for an indirect effect through disturbance of the pregnancy

Evidence for an indirect effect on the pregnancy caused by disruption or interruption of the intrauterine blood supply (so-called 'vascular disruption') was considered as another possible mechanism for congenital anomalies. However, no evidence in support of a possible disruptive effect of the components of Primodos on placental blood vessels was identified.

- Anomalies reported in babies whose mothers had been given an HPT

The EWG heard from 13 members of the 'Association for Children Damaged by Hormone Pregnancy Tests' who had, or whose child had had, one of a range of different anomalies. The Association members confirmed that the HPT had been taken within the critical period for fetal development, and that in many cases a test was recommended by the doctor rather than requested, that pills were given to first time mothers who did not consider themselves to be in any high-risk category, and that the doctor in several cases had taken what appeared to be free samples from a desk drawer[3], rather than making out a prescription.

In addition, the EWG examined reports of suspected adverse drug reactions (ADRs). Reports were received from a number of sources including the Yellow Card Scheme, The Association, individuals responding to the public call for information, pharmaceutical companies, and other regulatory agencies around the world. The final dataset comprised 235 reports which were categorised according to two separate medical classification schemes (European Surveillance of Congenital Anomalies, EUROCAT, and the World Health Organisation, WHO) to enable them to be directly compared. Five cases were reliably identified as having a genetic cause.

Congenital anomalies often affect more than one organ or limb. The EUROCAT method of classifying the cases allows patterns of anomalies to be examined and this identified 67 cases of possible multiple anomalies in babies whose mothers had been given an HPT. Upon manual review of the reports the expert geneticists on the EWG were unable to identify a combination of anomalies, repeated throughout the cases, which could represent an obvious syndrome or syndromes of drug-induced anomalies in the HPT cases.

To see if any anomalies were reported more frequently with HPTs than might be expected to occur in the general population and that might therefore have been suggestive of a drug-related phenomenon the types of ADRs reported with HPTs were compared to those reported to the EUROCAT congenital anomaly database. In addition, to identify any unusual patterns in the types of anomalies reported with HPTs, a comparison was made between the HPT cases and reports of congenital anomalies with all other medicines on MHRA's Yellow Card database.

The spontaneous reporting of a suspected ADR is typically subject to some general, well-recognised limitations, including an unknown and variable level of under-reporting. In addition to these general limitations, the EWG considered that a number of specific challenges applied to the HPT reports. First, some remaining duplication of ADR reports that could not be confirmed due to report anonymization. This was considered to be highly likely but, together with the small number of reports overall, would magnify the impact of residual duplicate cases. Second, the reports were of two different types: those reported spontaneously at the time they occurred and those reported in response to the call for information at the start of the review. Results from comparisons between the two types of report therefore needed very careful interpretation. Third, adverse events are generally reported shortly after they occur and so cases describing obvious physical anomalies may be expected to be reported more frequently than anomalies with no obvious external features that would tend to be diagnosed later. Any comparison of ADRs reported spontaneously with population congenital anomaly databases that collect information on all

[3] Today, there are strict requirements for the supply of free samples of medicines to prescribers, as set out in section 6.12 of the MHRA Blue Guide

anomalies irrespective of cause (eg. EUROCAT) therefore also need to be interpreted with great care.

With these limitations in mind, the EWG made a number of observations. The number of reports of congenital anomalies in babies of mothers given HPTs was relatively small in comparison to their extensive use. One in six of the babies with anomalies reported as exposed to HPTs had more than one congenital anomaly but no consistent pattern could be identified; for some of these it might now be possible to identify an underlying genetic basis with currently available genetic tests. Some differences were observed between the pattern of congenital anomalies reported with HPTs and the pattern reported to either the EUROCAT congenital anomaly database or MHRA's Yellow Card database with all other medicines, with some anomalies being over-represented and some under-represented, but for the reasons described above it was difficult to draw any firm conclusions.

Overall, the EWG concluded that the available adverse event reporting data had many limitations but did not support a causal association between use of HPTs, including Primodos, during pregnancy and congenital anomalies. Anomalies reported in association with HPTs were largely those that are clearly visible at birth and which occur relatively frequently in the general population.

- Evidence from studies on use of HPTs by pregnant women

Studies and articles on a possible association between the use of norethisterone or ethinylestradiol, or both, to diagnose pregnancy and the development of congenital anomalies was identified from searching the published literature. Hand-searching of references cited in individual papers was also performed to find any other relevant articles that might have not been captured. Reports of single or multiple cases, letters to journals, research studies, pooled analyses, review articles, editorials and other commentaries were all included. No date or language restrictions were applied to any of the searches, which used 12 different terms for HPTs combined with each of 69 congenital anomaly outcomes (using the EUROCAT description of congenital anomaly subgroups).

The search identified 4 390 potentially relevant publications of which 4 227 were excluded according to pre-defined exclusion criteria. Most papers were excluded because they had no data on congenital anomalies, they referred to animals or other pre-clinical studies, or they were studying a completely different treatment or irrelevant study population. Full review of the remaining 163 publications, and 12 others identified from hand-searching, resulted in the exclusion of a further 78 publications to leave 97 that were further evaluated.

A causal association between exposure to HPTs during pregnancy and development of a congenital anomaly in the fetus can never be proven or ruled out with absolute certainty through studies that observe individuals in their everyday life and measure outcomes, so-called observational studies. Nevertheless, it is easier to draw conclusions about effects that are assessed in studies that are well designed to minimise factors that may influence the results (referred to as bias and confounding) and that show consistency with other studies than it is to draw conclusions about effects that are assessed in studies that have major limitations in their design, or that show inconsistency with other studies, or both. Limitations such as these can falsely exaggerate or obscure an association. The EWG therefore agreed what would constitute good evidence, limited evidence, insufficient evidence or inadequate evidence within the available data for a causal association between the use of HPTs during pregnancy and congenital anomalies (see Section 3.3.2 of main report). In addition, the EWG defined the strength of any observed association between HPTs and congenital anomalies, from an extremely strong association to no association (see Section 3.3.2 of main report).

To evaluate all 97 publications a quality scoring system was developed and this was used to assess each individual study. When developing the scoring system, it was not possible to apply the same scientific rigour as we would to studies conducted today as this would have resulted in the exclusion of most of the studies. The key limitations that had been identified from a preliminary review of the data were therefore used to develop the quality criteria.

The quality scoring system was comprised of seven criteria that examined different aspects of the study:

1. whether the women selected to act as the comparator or control groups were appropriate and had the same baseline risk of giving birth to a baby with a congenital anomaly as those given HPTs
2. how reliably women's use of an HPT was determined (for example, from memory or by referring to medical records)
3. how specifically exposure had been recorded (for example, Primodos vs norethisterone and ethinylestradiol vs HPT vs sex hormones vs other)
4. how accurately timing of exposure had been recorded and whether it had been given during the critical period of development of the fetus
5. how well factors that could interfere with the study results had been adjusted for in analyses
6. whether the study was large enough to detect an association if one existed
7. whether studies that had investigated many outcomes had been adjusted appropriately to ensure a positive finding was not due by chance (referred to as multiplicity).

A traffic light scale of green/amber/red was used to indicate whether, for each of the seven quality criteria, it was considered to be good, moderate or poor quality, respectively. In addition to the quality assessment of each individual study design individual study findings had to be carefully considered alongside other factors, such as the size of the observed effect and whether it was consistent with the findings of other studies, to be able to make an informed judgement about the strength of any possible association between exposure to norethisterone and ethinylestradiol and each anomaly type.

In general, the studies were judged to have important limitations in their design and to be of poor quality with respect to at least one (and up to five) of the seven criteria. This made it difficult to draw any robust conclusions: that is, the evidence from many of these studies was insufficiently strong to demonstrate with certainty either that there was a causal association between HPTs and congenital anomalies or conversely that there was no possibility of a causal association. However, after a very careful assessment of each study the following observations were made:

- There was limited evidence for a weak association between the use of HPTs and congenital heart defects, limb reduction defects, and oesophageal atresia, but it was felt this could be due to chance or confounding factors.
- The evidence reviewed did not support an association between the use of HPTs and neural tube defects, orofacial clefts (hare lip or cleft palate), digestive system and abdominal wall defects, skeletal defects (other than limb reduction defects) or overall congenital anomalies but the quality of the evidence is limited.
- From the evidence available it was not possible to draw any conclusions about a possible association between the use of HPTs and urinary system or genital defects, nervous system defects (other than neural tube defects) or VACTERL (Vertebral defects, Anal atresia, Cardiovascular malformations, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies and Limb defects).

Overall, the EWG concluded 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, during pregnancy and any single anomaly, or any pattern of anomalies. The weak associations that were observed could have occurred by chance or confounding.

– Overall conclusion on a possible association between HPTs and congenital anomalies

Taken together, the EWG considered that the complete body of available evidence from pharmacology, non-clinical, epidemiological and adverse event reporting data was very limited and did not, on balance, support an association between use of HPTs such as Primodos by the mother during early pregnancy and congenital anomalies in the child.

Evidence for a possible association between use of HPTs in pregnancy and miscarriage (Chapter 6)

A possible association between HPTs and miscarriage has been proposed for many years. To explore this in more detail the effects of the components of Primodos, separately or together, in animal studies were assessed. Many of the studies described above, in relation to congenital anomalies, also reported on loss of the developing embryo (embryo-lethality).

Death of the developing embryo with high doses of estrogens has been consistently observed in animal studies and is now considered to be a well-established effect. A similar effect has been observed in studies with norethisterone (or related progestogens). As may be expected, the combination of norethisterone and ethinylestradiol also showed consistent embryo-lethality in different animal species. This effect was dose-dependent and varied according to when and for how long during pregnancy it was given. The mechanism for this effect in animals is not established but may relate to disruption of the relationship between the mother's hormones that are required to maintain pregnancy and the developing embryo or fetus.

In humans, the effect of norethisterone and ethinylestradiol (at doses equivalent to two Primodos tablets taken together) on early human pregnancy was investigated in two small clinical studies in women seeking legal termination of pregnancy in Finland. Though not large enough to detect any small differences in rates of miscarriage between the women who received hormones compared with those who received placebo, no adverse effects were observed on the developing pregnancy in terms of bleeding or a fall in maternal levels of progesterone.

In addition, twenty-one published studies in women who were given ethinylestradiol or norethisterone or both were evaluated. Most of these studies were conducted in the 1950s and 1960s to prevent threatened abortion (a total of 14) and many were considered to have limitations in their design or analysis. Only two studies investigated the use of Primodos for diagnosing pregnancy. The available epidemiological data were considered to be limited and to provide no evidence for an abortifacient effect of NETA and/or EE when given to pregnant women.

– Conclusion on a possible association between HPTs and miscarriage

Taken together the EWG considered that while administration of ethinylestradiol and norethisterone, mostly at very high doses or for prolonged periods, can result in embryo-

lethal effects in animals, there was no evidence that administration of these hormones at the licensed doses used in Primodos during early pregnancy was associated with an increased risk of miscarriage.

Safeguarding future generations (Chapter 7)

Women take medicines in pregnancy for a number of reasons: they may have an existing condition that requires ongoing treatment (eg. epilepsy, diabetes, HIV, depression, chronic inflammatory diseases); they may develop a condition during pregnancy that requires treatment (eg. gestational diabetes, infections); or they may be inadvertently exposed, particularly in the critical early stages when they may not be aware they are pregnant and so continue to take a medicine. It is therefore essential that any risk from a medicine to the fetus or to the pregnancy is identified at an early stage of its development, and preferably before it is used in humans.

It was therefore important to consider what developments have taken place since HPTs were on the market in terms of identifying, evaluating, managing and communicating safety concerns with medicines in pregnancy and whether the systems currently in place could be further strengthened to safeguard future generations.

- Before a medicine is licensed

Before a medicine can be used in humans a series of studies must be carried out by pharmaceutical companies to identify any undesirable properties that may have relevance to humans. There is a legal requirement to find out what action(s) the medicine has on the body and what actions the body has on the medicine. There are also comprehensive guidelines on what companies need to do to uncover any teratogenic potential with a medicine. These studies in animals generally provide the first source of information about any potential safety concerns with a medicine.

If no concerns are observed, companies are legally obliged to conduct randomised controlled trials in humans to establish the efficacy and safety of a new medicine. For many years, the experience with thalidomide meant that women were not included in these trials. In Europe, this continued until 1997 when guidance was changed to stipulate that women should be included in trials but those of childbearing potential must use highly effective contraception. If a medicine is intended specifically for use during pregnancy a randomised trial in pregnant women with follow-up of the pregnancy and of the baby during its development and after its birth is required. For medicines that are not specifically indicated in pregnancy but expected to be used by pregnant women, safety data are now collected after licensing, through an observational safety study.

Safety information collected before licensing is used to determine what is known about the safety of a medicine, to anticipate what is not yet known, to decide what concerns may need to be studied further, and to determine whether any action is required to reduce the risk of any of the identified concerns. These points are documented within a ‘risk management plan’, which is a legal obligation and must be approved by the regulator before a medicine can be licensed.

- After a medicine is licensed

- Detecting potential safety signals with medicines used in pregnancy

Before the introduction of the Yellow Card Scheme (YCS) in the UK in 1964, there was no requirement to report or collect cases of suspected adverse drug reactions (ADRs) to licensed medicines. Today pharmaceutical companies are legally required to operate a system for recording suspected ADRs in association with their medicines and to report these to national regulatory authorities within strict timeframes. The MHRA currently receives about 40 000 ADR reports per year and the Yellow Card database now holds about 850 000 reports. Despite a great many improvements to the YCS, spontaneous reporting of ADRs is not the best way to detect drug safety signals in pregnancy. This is because it is not well suited to identify conditions that occur relatively frequently in the population, such as the more common congenital anomalies, or any conditions that take a long time to show up, such as developmental disorders (eg. language, learning or autistic spectrum disorders that only become apparent in early childhood).

Pregnancy registers can be set up to monitor the use of a specific drug substance in pregnancy, or to follow pregnant women with a specific medical condition. However, the accuracy and completeness of the information collected depends on whether there is access to the medical records of the mother and newborn and in many cases this information is incomplete and poorly or inconsistently recorded. While it may be possible to detect signals through pregnancy registers, recruitment is voluntary and so the number of women may be too small to be able to detect anything other than a substantial harm.

Teratology Information Services (TIS) were set up around the world, including in the UK, following the thalidomide tragedy. These centres collect pregnancy outcome data from women who have been exposed to drugs and chemicals in pregnancy, and work together as a collaborative network with the aim of identifying, at an early stage, major teratogens and providing advice to women. Working with geneticists can help to identify whether an adverse outcome of pregnancy has a genetic or other cause.

Assessment of the use of medicines in pregnancy through electronic healthcare databases which can link GP data with other sources of data offers promise for the future. However, despite recording information on large numbers of women, these databases do not reliably capture maternal health during pregnancy, pregnancy outcome or collect information on medicines purchased over the counter, they are not always able to link data on mothers with data on their babies, and do not always collect accurate data about congenital anomalies. This and other methodological considerations, currently limits their use for detecting drug safety signals in pregnancy.

Information on congenital anomalies identified before or after birth is routinely collected in England through the National Congenital Anomaly and Rare Disease Registration Service (NCARDRS); however, information on what medicines may have been used during pregnancy is not collected. In Wales, the Congenital Anomaly Register and Information Service (CARIS) collects information on congenital anomalies and on medicines used in the first trimester.

No single system currently records usage of prescription medicines and over-the-counter medicines, the outcomes of all pregnancies, including longer-term outcomes, and can automatically link pregnant women with their children.

- Evaluating potential safety concerns (see also page ix)

The signal for a possible association between HPTs and spina bifida (a neural tube defect) initially came from a study conducted in 1967. Historical records suggest that a number of steps were taken by the Department of Health and their advisory bodies (the Committee on Safety of Drugs, CSD, later to become the Committee on Safety of Medicines, CSM) to further evaluate this potential signal. This included discussing the findings with the study authors, requesting manufacturers of HPTs and academics working in the field to provide them with all relevant information, exploring the possibility of collaborating on studies to explore the finding, starting their own study to examine a possible association between HPTs and anomalies, and regularly examining new data as it emerged. Nevertheless, the CSD/CSM were criticised for taking too long to withdraw Primodos from the market and for having links with the pharmaceutical industry. In the different European countries and globally, decisions to withdraw HPT products were taken in a staggered and uncoordinated way. This was partly because the evidence linking the use of HPTs to congenital anomalies was inconclusive; partly because communication channels that existed between regulators in different countries at that time were limited; and partly because many similar HPT preparations were available, with different names but the same ingredients.

The key principles of signal evaluation have not changed fundamentally since then, and all the actions described above with respect to data gathering, seeking expert advice and taking action to minimise risk would still be undertaken today. However, to ensure that important decisions made on drug safety issues today are timely, consistent across member states and legally binding on pharmaceutical companies, the current legislation[4] focuses heavily on the coordination of any regulatory action taken for safety reasons throughout the EU and sets out timelines for action that are consistent with the importance of the concern.

In 2005, a review of the UK advisory bodies resulted in formation of the Commission on Human Medicines (CHM) and a revised Code of Practice. This provides detailed guidance on holding, declaring and managing relevant interests, to ensure that all advice given is impartial. Under the revised Code the chairman and members of the CHM are not permitted to hold any current personal interests in the pharmaceutical industry. Any other conflicts declared and the actions taken to manage them must be recorded and made public in the CHM's annual report.

- Minimising harm to women and their unborn child

For reasons including scientific uncertainty over the data that was emerging with use of HPTs in early pregnancy and the increasing availability of better pregnancy tests, the following precautionary actions were taken with Primodos in the UK:

- In 1970 Schering removed the indication 'diagnosis of pregnancy' from the Primodos datasheet, stopped promoting Primodos for pregnancy testing, and stopped providing free samples to healthcare professionals[5].
- In 1975 the Primodos product information was updated to include a warning about the possible risk of congenital anomalies and a contraindication in pregnancy.

[4] Directive 2010/84/EU of the European Parliament and of the Council of 15 December 2010 amending, as regards pharmacovigilance, Directive 2001/83/EC on the Community code relating to medicinal products for human use.

[5] Today, there are strict requirements for the supply of free samples of medicines to prescribers, as set out in section 6.12 of the MHRA Blue Guide

- In 1977, it became apparent that in addition to being used to treat secondary amenorrhoea, Primodos was still being used to diagnose pregnancy. CSM therefore issued a reminder that HPTs should not be used in pregnancy.

In 1978 Schering withdrew Primodos from the UK market.

Similar approaches for reducing potential harm to the fetus are used by regulators and pharmaceutical companies today and a guideline, developed in 2008, makes sure that a consistent approach is taken when deciding on the most appropriate risk minimisation tools to use, based on the available evidence (see table 21 of main report). For medicines that pose a substantial threat to pregnancy a range of additional measures can be recommended or imposed including the development of materials to educate prescribers and women on the risk, requirement for proof of a current negative pregnancy test before a medicine can be prescribed, or implementation of a Pregnancy Prevention Plan, which aims to ensure that women are not pregnant when starting therapy and do not become pregnant during the course of treatment or soon after stopping.

▪ Communicating safety concerns with medicines

In 1970 when Schering removed the indication of Primodos as a hormone pregnancy test from its data sheet, the change was not proactively communicated to prescribers, whose behaviour apparently remained relatively unchanged. With the publication of CSM's two warnings on HPTs, prescribing was reduced by around 60% in 1975 and by a further 30% in 1977. The observed fall in prescribing suggests that this may have been due to CSM communications but does not necessarily prove it as this period coincided with wider availability of modern pregnancy tests.

Today, the legislation makes it clear that pharmaceutical companies should communicate important safety information relating to their medicine to prescribers in a consistent and co-ordinated way. This is primarily achieved through circulation of a Dear Healthcare Professional Communication (DHPC). In the UK, MHRA has a number of ways to communicate drug safety messages, including through a web-based message cascading system to NHS, through the press or digital and social media, or by collaborating with Professional Societies to disseminate the message to those who most need to receive it. For more routine drug safety messages, since 2007 MHRA has published an online monthly drug safety bulletin, Drug Safety Update (DSU), for healthcare professionals including GPs. Before this a drug safety bulletin called 'Current Problems in Pharmacovigilance' was circulated by MHRA to healthcare professionals in the UK.

Full information on the safe and effective use of all licensed medicines is also provided to prescribers (through the Summary of Product Characteristics, SmPC) and to patients, through the Patient Information Leaflets (PILs) that accompany every medicine. Both documents are updated every time important new information is identified.

▪ Measuring how effectively risk to patients has been minimised

Even after the CSM communicated the advice on HPTs to healthcare professionals in 1975 and 1977, the new restrictions were not adhered to by all prescribers with the result that Primodos continued to be used by some to diagnose pregnancy for as long as eight years after the indication for prescription was restricted to treating women with menstrual disorders. Today, pharmaceutical companies may be required to evaluate whether action taken to minimise an important risk has had the desired effect, and if not why not. If the results suggest that the action has not achieved its goal, the advice is either repeated using

different channels, or further potentially more restrictive measures may be considered followed by further evaluation of their effectiveness. When all other regulatory options have been shown to be insufficient in addressing the risk, consideration is given to whether there is a need to withdraw the medicine from the market. The MHRA also regularly screens the action it has taken to see where there may be a need to conduct additional research.

- Conclusion on safeguarding future generations

Despite major improvements over the years in the systems in place to support the safe and effective use of medicines the EWG considered that there were several areas that could be further strengthened to safeguard future generations and that these should form the focus of their recommendations.

Summary conclusions of the Expert Working Group on the available evidence (Chapter 8)

The EWG set out to address three key issues in relation to the evidence reviewed. Its overall conclusions on each of these areas, based on a comprehensive evaluation of the available data as presented above, are as follows:

1. **To consider all available evidence on the possible association between exposure in pregnancy to HPTs and adverse outcomes in pregnancy (in particular congenital anomalies, miscarriage and stillbirth) including consideration of any potential mechanism of action**

   The EWG's overall finding is that the available scientific evidence, taking all aspects into consideration, does not support a causal association between the use of HPTs, such as Primodos, during early pregnancy and adverse outcomes, either with regard to miscarriage, stillbirth or congenital anomalies. All the available relevant evidence on a possible association has been extensively and thoroughly reviewed with the benefit of up-to-date knowledge by experts from the relevant specialisms.

2. **On whether the Expert Working Group's findings have any implications for currently licensed medicines**

   The findings of the review for HPTs, including Primodos, on a possible association between exposure in pregnancy to HPTs and adverse outcomes in pregnancy do not have implications for any currently licensed medicines. They are in fact considered to be reassuring for women who may inadvertently become pregnant whilst taking these hormones for contraception or gynaecological indications.

3. **To draw any lessons for how drug safety issues in pregnancy are identified, assessed, and communicated in the present regulatory system and how the effectiveness of risk management is monitored**

   There have been substantial and far-reaching advances in all areas of the development, regulation, study and use of medicines in pregnancy since HPTs were available in the UK, whereas there was a lack of transparency in the past. Nevertheless, ways to strengthen further how safety concerns in pregnancy are detected, managed, evaluated and communicated should be taken forward.

4. To make recommendations

The EWG considered that a number of steps could be taken to safeguard future generations through strengthening the systems in place for detecting, evaluating, managing and communicating risk with exposure to medicines in early pregnancy.

These include:

- undertaking an annual review of all reported congenital anomalies with independent scientific advice of CHM, published in its annual report
- facilitating research by optimising the collection of, access to and use of data on medicines in pregnancy
- safeguarding future generations through improved training and guidance of healthcare professionals
- working to improve the impact of safety messages on the risks of medicines in pregnancy.

In addition, families of the Association for Children Damaged by HPTs, whose lives have been impacted by adverse pregnancy outcomes and who were given HPTs to diagnose pregnancy should be offered a full up-to-date genetic clinical evaluation.

The recommendations are provided in full in Section 8.2.

Abbreviations

Abbreviation	Definition
ADR	Adverse drug reaction
BINOCAR	British and Irish Network of Congenital Anomaly Registers
CARIS	Congenital Anomaly Register and Information Service for Wales
CHM	Commission on Human Medicines
Cmax	Maximum total serum concentration
CPRD	Clinical Practice Research Datalink
CSD	Committee on Safety of Drugs
CSM	Committee on Safety of Medicines
DHSS	Department of Health and Social Security
DNA	Deoxyribonucleic acid
DSU	Drug Safety Update
E2	Estradiol
EAG	Expert Advisory Group
EE	Ethinylestradiol
ENCePP	European Network of Centres for Pharmacovigilance and Pharmacoepidemiology
ER	Estrogen receptor
EU	European Union
EUROCAT	European surveillance of congenital anomalies
EWG	Expert Working Group
FOI	Freedom of information
Fu	Unbound fraction
GMC	General Medical Council
GVP	Good Vigilance Practice
HPT	Hormone pregnancy test
ICD	International Statistical Classification of Diseases
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
LMP	Last menstrual period
MDI	Medical Data Index
MedDRA	Medical Dictionary for Regulatory Activities

MHRA	Medicines and Healthcare products Regulatory Agency – the regulatory divisions
µg	Microgram (a millionth of a gram)
MDI	Medical Data Index
mg	Milligram (one thousandth of a gram)
ml	Millilitre
mRNA	Messenger RNA
NCARDRS	National Congenital Anomaly and Rare Disease Registration Service
NCAS	National Congenital Anomaly System
NET	Norethisterone
NETA	Norethisterone acetate
ng	Nanogram (a billionth of a gram, or one thousandth of a microgram)
nM	Nanomolar
NO[A]EL	No Observed [Adverse] Effect Level
OPCS	Office of Population Censuses and Surveys
PASS	Post-authorisation safety study
pg	Picogram (a trillionth of a gram, or a thousandth of a nanogram)
PIL	Patient Information Leaflet
PPP	Pregnancy Prevention Plan/Programme
PR	Progesterone receptor
PRAC	Pharmacovigilance Risk Assessment Committee
RCT	Randomised controlled trial
RMP	Risk Management Plan
RNA	Ribonucleic acid
SAIL	Secure Anonymised Information Linkage
SCAR	Subcommittee on Adverse Reactions
SCAR	Scottish Congenital Anomalies Register
SHBG	Sex hormone binding globulin
SmPC	Summary of Product Characteristics
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
THIN	The Health Improvement Network
UKTIS	UK Teratology Information Service
VACTERL	Vertebral defects, Anal atresia, Cardiovascular malformations, Tracheoesophageal fistula,

	Esophageal atresia, Renal anomalies and Limb defects.
WHO	World Health Organisation
YCS	Yellow Card Scheme

1 INTRODUCTION

The Commission on Human Medicines (CHM) established an Expert Working Group (EWG) on HPTs in October 2015 to investigate a concern that the use of these products was associated with adverse outcomes in pregnancy, in particular congenital anomalies, miscarriage and stillbirth.

This report provides historical and scientific context to the use of HPTs before going on to summarise the data considered for each of the key areas of concern that were examined by the EWG and the findings in relation to those areas. It ends by considering what has changed, in the way drug safety concerns in pregnancy are identified, evaluated and communicated, since HPTs were on the market and looks forward to what further could be considered to safeguard future generations through strengthening existing systems and processes relevant to the safe use of medicines in pregnancy. The report provides a summary of all key data evaluated by the EWG and is intended as a stand-alone document.

This chapter outlines the background to the review and the establishment of the EWG, and provides information about the membership of the EWG, its remit and the scope of its work. Further information may be found in the full EWG papers in Annexes 1-5.

1.1 Background to the review

1.1.1 Concern with HPTs

Hormone Pregnancy Tests (HPTs) for diagnosing pregnancy first became available in the UK in the 1950s. Hormone Pregnancy Tests generally contained progestogens combined with an estrogen and were given orally as tablets or by an injection. This led to a withdrawal bleed a few days later in those who were not pregnant. The first publication which suggested that HPTs could cause malformations was by Edwards in 1958. Then in October 1967, the first observational study to suggest a link between use of HPTs in pregnancy and congenital anomalies in the exposed child was published in a letter to the journal Nature (Gal, 1967). This study showed an increased risk of spina bifida in the babies of mothers who had taken HPTs to diagnose their pregnancy. Against a background of heightened awareness of the possible teratogenic effect of medicines taken in pregnancy through recent experience with thalidomide (a medicine used to prevent morning sickness) and, primarily, phocomelia in the offspring, Gal's study stimulated major interest in the issue.

Over the next decade, a great many studies investigated a possible association between HPTs and a range of congenital anomalies including neural tube defects, heart defects and limb reduction defects (section 5.3.4.4).

1.1.2 The need for a scientific review

In the 1970s, a campaign group called the 'Association for Children Damaged by Hormone Pregnancy Tests' raised concerns about the possible association between the use of HPTs in pregnancy and congenital anomalies. This resulted in parliamentary interest in the issue and the campaign group subsequently brought legal action against Schering (the manufacturer for Primodos, a widely-used HPT in the UK), since taken over by Bayer. Although the case was ultimately discontinued, the more recent discovery of documents atthe National Archives in London prompted renewed campaigning by the group. This culminated in a debate on HPTs in the House of Commons by the Backbench Business Committee on 23 October 2014 in which the following question was put to the House:

"That this House notes that children were born with serious deformities due to hormone pregnancy test drugs[6] taken by expectant mothers between 1953 and 1975; also notes with concern that as the surviving victims enter their forties and fifties many of them face a host of new problems as their bodies continue to suffer; further notes that no official warnings were issued about these drugs until eight years after the first reports indicated possible dangers; further notes that some doctors continued to prescribe the drugs for pregnant women after official warnings from the Committee on Safety of Medicines; calls on the Secretary of State for Health to fully disclose all documents relating to the use of Hormone Pregnancy Tests held by the Department from the period between 1953 and 1978; and also calls on the Secretary of State to set up an independent panel to examine these documents."

During the debate, the then Minister for Life Sciences, George Freeman MP, stated that he would instruct that all relevant documents held by the Department of Health be released. Furthermore, an independent review of the papers and all the evidence would be conducted to determine whether the assertion of the Backbench Business Committee that there was a causal association between HPTs and congenital anomalies was justified, based on the available data.

In 2015, the Commission on Human Medicines agreed to establish an Expert Working group to review the available data on a possible association between HPTs and adverse outcomes in pregnancy, and to make recommendations.

1.2 The Expert Working Group on HPTs

1.2.1 Remit of the Expert Working Group on HPTs

The terms of reference of the EWG, were finalised and adopted by the EWG at its second meeting on 4 December 2015, as follows:

- To consider all available evidence on the possible association between exposure in pregnancy to hormone pregnancy tests (HPTs) and adverse outcomes in pregnancy (in particular congenital anomalies, miscarriage and stillbirth) including consideration of any potential mechanism of action
- To consider whether the EWG's findings have any implications for currently licensed medicines in the UK or elsewhere
- To draw any lessons for how drug safety issues in pregnancy are identified, assessed and communicated in the present regulatory system and how the effectiveness of risk management is monitored
- To make recommendations.

It was not within the remit of the EWG to make formal conclusions or recommendations on the historical system or regulatory failures. However, the EWG was provided with all the available information on the regulatory history so it could decide which evidence was most relevant to consider, and to comment or make recommendations, as appropriate.

The names of the members of the EWG are listed in section 1.3 and include those with specialist expertise in teratology, toxicology, medicinal chemistry, pharmacometrics, embryology, prenatal genetics, clinical genetics, neonatology, gynaecology, obstetrics, reproductive endocrinology, pharmacoepidemiology and statistics as well as lay representation.

[6] Note: A causal association between HPTs and fetal anomalies was presumed by the Backbench Business Committee but remained to be established.

The EWG met to consider the evidence seven times over the period October 2015 to April 2017; full members met again in July 2017 to work on the report.

1.2.2 Scope of the review

A range of HPT products first became available in the UK in the 1950s and contained natural or synthetic sex steroid hormones, usually a progestogen in combination with an estrogen. A short course of tablets (usually between one to four days) was taken by women who suspected they were pregnant, which led to a withdrawal bleed a few days later in those who were not pregnant.

The progestogenic hormones display diverse biological activities, as demonstrated in Table 1. As an example of this diversity, in addition to its direct progestogenic effects natural progesterone can have a range of indirect actions including anti-gonadotropic, anti-estrogenic, glucocorticoid, partial anti-androgenic and anti-mineralocorticoid effects.

Table 1. Biological activities of progestogens (from Schindler, 2003, based on studies in animals and humans.)

Progestin	Progesto-genic	Anti-gonado-tropic	Anti-estrogenic	Estro-genic	Andro-genic	Anti-andro-genic	Gluco-corticoid	Anti-mineralo-corticoid
Progesterone	+	+	+	−	−	±	+	+
Dydrogesterone	+	−	+	−	−	±	−	±
Medrogestone	+	+	+	−	−	±	−	−
17α-Hydroxy-derivatives								
Chlormadinone acetate	+	+	+	−	−	+	+	−
Cyproterone acetate	+	+	+	−	−	++	+	−
Megestrol acetate	+	+	+	−	±	+	+	−
Medroxy-progesterone-acetate	+	+	+	−	±	−	+	−
19-Nor-progesterone-derivatives								
Nomegestrol acetate	+	+	+	−	−	±	−	−
Promegestone	+	+	+	−	−	−	−	−
Trimegestone	+	+	+	−	−	±	−	±
Spirolactone-derivatives								
Drospirenone	+	+	+	−	−	+	−	+
19-Nortestosterone derivatives								
Norethisterone	+	+	+	+	+	−	−	−
Lynestrenol	+	+	+	+	+	−	−	−
Norethinodrel	±	+	±	+	±	−	−	−
Levonorgestrel	+	+	+	−	+	−	−	−
Norgestimate	+	+	+	−	+	−	−	−
3-Keto-desogestrel	+	+	+	−	+	−	−	−
Gestoden	+	+	+	−	+	−	+	+
Dienogest	+	+	±	±	−	+	−	−

(+) effective; (±) weakly effective; (-) not effective

Because of this complexity, the EWG considered it necessary to narrow the focus of the review, where possible, to the components of Primodos as this was the most widely-used HPT in the UK. Primodos contained a combination of norethisterone acetate (NETA) and ethinylestradiol (EE) and was available in the UK from 1958 until 1978.

1.2.3 Programme of work

The EWG agreed the following programme of work to address the adopted terms of reference:

- Pharmacology and pharmacokinetics – to outline the basic chemistry and metabolism of the active ingredients of Primodos, including a description of the pharmacokinetics and pharmacodynamics of these ingredients.

- Toxicology and studies of teratogenicity – to examine the non-clinical studies performed on the components of Primodos, with a particular focus on data from animal studies of single-dose and repeat-dose toxicity, reproductive toxicity and genotoxicity.
- Possible mechanisms of action for adverse effects of the components of Primodos on the developing pregnancy.
- Reports of suspected adverse drug reactions (ADR), including: reports from the UK Yellow Card Scheme (YCS) and other ADR reporting systems worldwide, information on individual cases, testimonials, and published cases and case series.
- Epidemiological evidence, including data from prospective and retrospective studies of the outcomes of pregnancy and exposure to the components of Primodos.

The EWG was provided with assessment reports prepared by MHRA which summarised the evidence, and had free access to all evidence gathered by MHRA for the review. Final versions of the assessment reports are provided in the annexes to this report and, together with all the supporting evidence, will be published at https://www.gov.uk, once they have been reviewed in line with duties under data protection legislation, and common law duty of confidence.

1.2.4 Key data sources

Due to the length of time that has elapsed since HPTs were on the UK market, very little historical documentation on these products was retained by the Government. To conduct a comprehensive review of all evidence relating to whether their use may have been associated with congenital anomalies, information was sought through a public call for evidence, in which MHRA invited interested individuals and organisations to provide any information they considered relevant to a possible association between the use of HPTs and adverse effects on the pregnancy or subsequent congenital anomalies in the child. Individuals responding to the call for evidence were encouraged to use the Yellow Card website to send any personal or confidential information relating to their own experiences with HPTs. To draw further attention to the review, the call for evidence was accompanied by a press release at launch and in June 2015 an article in MHRA's 'Drug Safety Update' bulletin.

Data and information on HPTs were also obtained from the following key sources:

– **Pharmaceutical companies whose predecessors marketed an HPT**

Every pharmaceutical company whose predecessors marketed an HPT were contacted by MHRA and asked to submit for the review all relevant information or documentation that it held, including all published or unpublished data from non-clinical, clinical, mechanistic or pharmacoepidemiology studies, all related internal and external communications and sales or usage data for the period spanning 1950 to 1980.

The Human Medicines Regulations 2012 require pharmaceutical companies to provide data to MHRA on request for medicines that have a current marketing authorisation; however, there is no legal provision which compels companies to submit data for medicines that have no current MA. Co-operation of these companies in providing information was therefore voluntary.

The majority of companies whose predecessors marketed an HPT other than Primodos stated they no longer held any relevant data. However, some data was provided.

– **Bayer**

Bayer is the successor company to Schering, the licence holder for Primodos. Bayer provided MHRA with a schedule of archived documents, retained by Schering's legal

representatives from the 1970s court case. The schedule documented reports of various types of evidence including: unpublished pre-clinical data, published scientific papers, medical histories, correspondence, reviews by Schering specialists, and expert comments and opinions.

The terms under which MHRA was able to receive some of these documents were such that it would not have been possible to disclose them to an audience wider than the EWG because of restrictions related to data protection, legal privilege and confidentiality. Because of this legal restriction, the data that was requested from Bayer was restricted to scientific data that could not have been obtained from any other source. Much of the other data that was not requested from Bayer could be obtained from other sources where these legal restrictions did not apply.

- **Information provided by other institutions**

The Royal Colleges of Physicians, Obstetrics and Gynaecology, Paediatrics and Child Health, and General Practitioners were asked to search their on-site and archived records for any relevant published or unpublished information or data for inclusion in the review. Most of the Royal Colleges confirmed they did not hold any relevant information; the Royal College of General Practitioners (RCGP) submitted information relating to a survey study on congenital anomalies, subsequently published in 1964 (Slater, 1964).

- **Information provided by other countries**

The MHRA asked the other EU Member States to provide details of:

1. Any HPTs available in their country in the 1950s, 1960s and 1970s.

2. Any communications or regulatory action taken by the National Competent Authority or the pharmaceutical companies with respect to HPTs and congenital anomalies (including, where possible, any relevant evidence and stating the key evidence responsible for triggering any action).

3. Details of any current or ongoing reviews of, or interest in, HPTs in each country.

Seventeen Member States responded, the majority of whom had no information to provide. Hormone Pregnancy Tests had been available in some Member States and limited information on their withdrawal (and the reasons for this) were provided. Inaccuracies within the responses (highlighted by the fact that the information provided was sometimes at odds with information within the archive documents) likely reflects the absence of records from this time currently held within the regulatory agencies in these countries.

Limited information about regulatory action taken in other countries was also obtained for Australia, India, Japan, New Zealand and a number of other Asia-Pacific countries. The Food and Drug Administration (FDA) in the US provided a copy of a thesis from 1976 entitled “A Study of the Teratogenic Potential of Oral Hormonal Pregnancy Tests in the White Rat” by Mark R. Kazmierski.

- **UK National Archives**

A search of the national archives was conducted by a professional researcher, with the aim of obtaining a complete set of historical documents relevant to this issue. Searches were performed for any documents which referred to ‘hormone pregnancy tests’, ‘hormonal pregnancy tests’ or to any of the 12 branded products known to be used as an HPT.

In total 151 files were ordered and reviewed of which 108 files contained no relevant information. Of those that contained relevant information, 32 files were copied partially and 6 were copied fully. Five files that appeared to contain relevant information from the

period January 1962 to December 1967 were marked 'Missing at transfer' and could not be recovered.

- **Academics**

Expertise on the social and medical environment in which HPTs were used was sought to provide a broader context and Dr Jesse Olszynko-Gryn from the University of Cambridge agreed to provide a copy of his thesis "Pregnancy testing in Britain, c.1900-67: laboratories, animals and demand from doctors, patients and consumers" to MHRA. Dr Olszynko-Gryn was also invited to give a presentation to the EWG (see below).

- **People who have been given HPTs and who have experienced adverse pregnancy outcomes**

Several individuals whose lives have been impacted by an adverse outcome of pregnancy after taking HPTs in pregnancy provided documentation, including Yellow Card reports, for the review.

- **Experts**

A number of experts, including some EWG participants, were invited to give evidence to the EWG on specific aspects of the review, as follows:

- Professor David Healy: "Spontaneous reporting systems and their strengths and limitations, particularly with respect to detecting/identifying congenital anomalies and adverse effects on the pregnancy".
- Dr Diana Wellesley: "A current update on congenital anomalies"; and "Congenital anomaly registration in the UK"
- Professor Helen Dolk: "Pharmacovigilance for medication safety in pregnancy: data sources and designs"
- Dr Rachael Williams: "CPRD and pregnancy research"
- Sarah Stevens: "The National Congenital Anomaly and Rare Disease Registration Service"
- Dr Ulla Wandell Liminga and Professor Corinne de Vries: "Good Pharmacovigilance Practice Guidance. Special populations III: Pregnancy and breastfeeding"
- Dr Neil Vargesson: "Review of new pre-clinical research". Dr Vargesson's work was unpublished (and therefore not peer-reviewed) and so a summary of the findings only was presented to the EWG.
- Dr Jesse Olszynko-Gryn: "Water under the bridge? A historical argument for regulatory failure in the case of HPTs"

- **Published literature**

Searches of scientific publications were made to identify all literature that might be relevant to the review. References cited are listed at the end of this report; complete lists of published literature that supported the review are provided within each of the relevant EWG papers in the annexes.

A summary of all the information provided through the call for evidence is provided in Annex 5.

1.2.5 Terminology

There are many different spellings and abbreviations in the literature relating to the substances and technical terms considered in these papers. Some reflect spelling and terminology differences between countries and others reflect changes in scientific language over time. To avoid confusion, consistent terminology has been used throughout the report, even where the original data or research papers used different terminology. This does not necessarily apply to the annexes to the report.

Synthetic hormones

The following were considered to best reflect current UK spellings and abbreviations for the most common of these:

- Estrogen(s) (rather than 'oestrogen'), for the naturally occurring estrogenic sex steroids
- Estradiol (rather than 'oestradiol'), for 17β-estradiol (E2) the most potent of the naturally occurring estrogens
- Progestogen (rather than 'progestins' or 'progestagen'), for synthetic versions of progesterone
- Ethinylestradiol or EE (rather than 'ethinyloestradiol' or 'ethinyl-estradiol'), for the estrogenic component of Primodos
- Norethisterone acetate or NETA (rather than the American 'norethindrone'), for the progestogenic component of Primodos.

NETA is considered to be a prodrug that is immediately converted to norethisterone (NET) by the body after administration. The report therefore refers to clinical use of NETA as a component of Primodos tablets, and to NET once the tablet has been ingested and absorbed and is in the body. Of note, some of the *in vitro* studies examined both NETA and NET.

Primodos was reformulated twice, the second time in 1963. This report uses the doses of NETA and EE that were in this final version of the product for all its calculations as this represents the version of greatest use ie. NETA (10 mg per tablet) and EE (0.02 mg per tablet), one tablet to be taken on each of two consecutive days.

Congenital anomalies

Congenital anomalies have variously been referred to as 'birth defects', 'malformations', 'congenital malformations', 'developmental defects', 'congenital abnormalities' or 'congenital anomalies' – all have the same meaning. This report uses the term 'congenital anomaly(ies)' when referring to humans and 'malformations' when referring to animals.

Many different classification systems have been used to describe and organise congenital anomalies into groups, including classification by cause, morphologic alteration, regional anatomy and body system. For this review the International Statistical Classification of Diseases (ICD) systems, in which congenital anomalies are given codes which are grouped together according to World Health Organisation (WHO) definitions for specific diagnoses, was used to classify ADR reports. The only exception is in Chapter 5, where adverse event reporting data are also categorised according to the EUROCAT (European surveillance of congenital anomalies; EUROCAT Guide 1.4, version 20.12.16) method.

Periods of fetal development and gestation (pregnancy)

Clinicians refer to the duration of gestation, or pregnancy, starting from the time of the woman's last menstrual period (LMP); embryologists describe the weeks of fetal development from the time of conception (fertilisation of the oocyte), ie. from about day 14 or 15 of the woman's last menstrual cycle.

There is a difference of about two weeks between these two definitions which must be taken into account when considering exposure to any potential teratogen. Thus, the embryo of a woman told she is six weeks pregnant will have undergone about four weeks of development. In this report, duration of pregnancy is referred to in terms of 'gestation weeks or days' and embryonic/fetal growth is referred to in terms of 'development weeks or days'.

Throughout the report the term embryo has been used where it is clear that the developing fetus is less than 8 weeks old; where this is not clear the term fetus has been used.

Scientific units

Prefixes have been used to denote quantities of less than one, to minimise the number of zeros. For example, 1milligram (mg) is a thousandth of a gram (g) and this can also be expressed as 0.001g (or $1x10^{-3}$g).

Thus:

- milli (m) = 1/1000 (10^{-3})
- micro (μ) = 1/1000,000 (10^{-6})
- nano (n) = 1/1000,000,000 (10^{-9})
- pico (p) = 1/1000,000,000,000 (10^{-12})

When expressing the amounts of two components in a single product that have very different concentrations, such as Primodos, the report refers to both in the same unitage to facilitate comparisons ie. NETA (10 mg) and EE (0.02 mg).

1.3 Membership of the Expert Working Group

1.3.1 Conflicts of interest

The CSD/CSM has been criticised for having links with the pharmaceutical industry. When these committees were established the chair was not permitted to have current personal financial interests in the pharmaceutical industry and interests of members had to be declared. In 2005, a review of the UK advisory bodies resulted in formation of the CHM and a revised Code of Practice. This provides detailed guidance on holding, declaring and managing relevant interests, to ensure that all advice given is impartial. Under the revised Code the chairman and members of the CHM are not permitted to hold any current personal interests in the pharmaceutical industry. Any other conflicts declared and the actions taken to manage them must be recorded and made public in the annual report (Annex 7).

To maintain the independence of the EWG, a policy was therefore developed that defined the conflicts of interest and level of involvement for four categories of participant: 1) full members, 2) invited experts, 3) visiting experts and 4) observers (Annex 6). Under the terms of this policy full members could not have: any interest in pharmaceutical companies whose predecessors marketed HPTs; any interest in the outcome of the review; and should not have publicly expressed a strong opinion (either favourable or unfavourable) about the possibility of a risk of congenital anomalies with use of HPTs or about any of the pharmaceutical companies whose predecessors produced them. The principles set out in the 'Code of Practice for Chairmen and Members of the Commission on Human Medicines, Certain Committees and Expert Advisory Groups' provided the basis for the policy developed for this EWG (Annex 7).

All those invited to participate in the EWG were asked to declare any interests held before the EWG first met – to establish the appropriate participant category – and at every meeting thereafter in case of any changes (see Annex 8 for declarations of interests).

1.3.2 Membership

Membership of the EWG was based on the required areas of expertise. At their second meeting the EWG recommended that experts in pharmacometrics and medicinal chemistry would be valuable additions. During the course of the review two of the initially invited experts stepped down due to a conflict of interest in one case and time pressures in the other; both had attended one meeting of the EWG. The final EWG membership was as follows:

Chair

Dr Ailsa Gebbie. MB ChB FRCOG FRCP(Edin) FFSRH
Consultant Gynaecologist and Co-Director of the Clinical Effectiveness Unit of the Faculty of Sexual and Reproductive Health, Chalmers Centre, Edinburgh
Chair of Medicines for Women's Health Expert Advisory Group to the Commission on Human Medicines

Members

Professor Pat Doyle BSc MSc PhD
Professor of Epidemiology, London School of Hygiene & Tropical Medicine

Mrs Joyce Epstein
Former Director of the Foundation for the Study of Infant Deaths. Member of NICE accreditation committee and NSPCC research ethics committee. **(Lay Representative)**

Professor Stephen Evans Ba MSc CStat FRCP (Edin) FISPE Hon. FRCP (Lon)
Professor of Pharmacoepidemiology, London School of Hygiene & Tropical Medicine
independent Expert member of the Pharmacovigilance and Risk Assessment Committee at the European Medicines Agency

Professor Joyce Harper
Professor in Human Genetics and Embryology, University College London

Professor Dr Axel Heep FRCPCH
Consultant Neonatologist, Southmead Hospital, North Bristol NHS Trust; Honorary Senior Clinical Lecturer, School of Clinical Sciences, University of Bristol; Professor, Technical University Munich, Germany

Professor Stephen Hillier OBE DSc FRCPath FRCOG
Emeritus Professor of Reproductive Endocrinology, University of Edinburgh

Professor Alison Macfarlane BA Dip Stat CStat FFPH
Professor of Perinatal Health, School of Health Sciences, City University London

Ms Sara Payne BA CPE LPC
Solicitor **(Lay Representative)**
Member of Paediatric Medicines Expert Advisory Group to the Commission on Human Medicines

Mrs Farrah Pradhan
Invited Reviews Coordinator at the Royal College of Obstetricians and Gynaecologists **(Lay Representative)**
Member of Herbal Medicines Advisory Committee; Member of Patient and Public Engagement Expert Advisory Group to the Commission on Human Medicines

Professor Shirley Price MSc PhD FBTS ERT FHEA FSB

Academic Quality and Professor of Toxicology and Head of Academic Appeals, University of Surrey
Member of the Commission on Human Medicines

Professor Siobhan Quenby MBBS BSc MD FRCOG
Professor of Obstetrics, Warwick University
Medicines for Women's Health Expert Advisory Group to the Commission on Human Medicines

Dr Richard Quinton MB Bchir
Consultant Endocrinologist, Endocrine Unit Royal Victoria Infirmary

Dr Connie Smith MB BS MFSRH
Retired Consultant in Sexual and Reproductive Health Care, Central London Community Healthcare

Professor Michael D Threadgill PGCE MA PhD DSc FRSC Cchem
Professor in Medicinal Chemistry, Department of Pharmacy and Pharmacology, University of Bath
Member of Chemistry, Pharmacy and Standards Expert Advisory Group to the Commission on Human Medicines

Dr Diana Wellesley BM, FRCP
Head of Prenatal Genetics, Consultant and Honorary Senior Lecturer in Clinical Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital Southampton

Invited Experts

Professor Leon Aarons BSc MSc PhD ICI
Professor of Pharmacometrics, Manchester Pharmacy School, The University of Manchester

Mr Nick Dobrik
Thalidomide Campaigner

Professor Helen Dolk DrPH
Professor of Epidemiology & Health Services Research, Ulster University

Professor Kay Marshall
Head of the Manchester Pharmacy School, University of Manchester

Dr Irene Petersen
Reader in Epidemiology and Statistics, University College London

Professor Faith Williams
Emeritus Professor of Toxicology, Medical Toxicology Centre and Institute of Cellular Medicine, Newcastle University

Dr Laura M Yates MBChB DRCOG MRCPCH PhD
Consultant in Clinical Genetics, Institute of Genetic Medicine; Head of Teratology, UK Teratology Information Service, Newcastle upon Tyne Hospitals NHS Foundation Trust; Honorary Senior Lecturer in Clinical Genetics, Newcastle University

Observers

Mrs Marie Lyon
Chair of the 'Association for Children Damaged by Hormone Pregnancy Tests'

PD Dr Elke Röhrdanz
EUROTOX registered Toxicologist; Head of the Unit Reproductive and Genetic Toxicology, Bundesinstitut für Arzneimittel und Medizinprodukte (Federal Institute for Drugs and Medical Devices), Germany

Individuals who presented evidence to the EWG

Representatives from the 'Association for Children Damaged by HPTs'

Thirteen members of the 'Association for Children Damaged by HPTs' presented their personal experiences to the EWG (at the meeting of 4th December 2015)

Mrs Marie Lyon, Chair of The Association gave a presentation to the EWG, entitled "Documents from the Landesarchiv Berlin".

Visiting experts who also gave presentations to the EWG

Professor David Healy, retired Professor of psychiatry and psychopharmacology (for the meeting of 4th December 2015)

Dr Rachael Williams, Research Programme Manager, CPRD (for the meeting of 11th August 2016)

Dr Sarah Stevens, Public Health Consultant for Public Health England (for the meeting of 11th August 2016)

Dr Ulla Wandell Liminga, Non-clinical assessor and PRAC delegate, Läkemedelsverket Swedish Medical Products Agency (for the meeting of 11th August 2016)

Professor Corinne de Vries, Head of Science and Innovation Support, Human Medicines Research and Development Support Division, European Medicines Agency (for the meeting of 11th August 2016)

Dr Neil Vargesson, Senior Lecturer, School of Medicine, Medical Sciences and Nutrition, Institute of Medical Sciences, University of Aberdeen (for the meeting of 18th October 2016)

Dr Jesse Olszynko-Gryn, Director of Studies, Department of History and Philosophy of Science, University of Cambridge (for the meeting of 24th April 2017)

2 HISTORICAL AND SCIENTIFIC CONTEXT

This chapter describes the social, medical and regulatory frameworks within which HPTs were used in the UK between the 1950s and 1978. Further information is provided in annexes 3, 9 and 10.

2.1 Historical perspective

2.1.1 Socio-medical environment

During the 1950s and 1960s, access to family planning advice and effective contraception was limited and abortion (other than in extreme medical circumstances) was illegal in the UK until 1967. Whilst pregnancy testing became more available from the 1920s onwards, it did not become mainstream or universal until much later, after the introduction of home pregnancy tests in the early 1980s.

Women did not usually attend antenatal clinics or consult a doctor about their pregnancy before the second or third trimester. This situation continued into later years, despite the increasing availability of the laboratory tests and then HPTs in the 1960s and 1970s. The unreliability of self-diagnosis of pregnancy through physical manifestations together with the invasiveness of bimanual examination by a doctor appear to have been at least partly responsible for driving the development of more accurate tests for pregnancy diagnosis.

A series of laboratory tests during the first half of the 20th century gave rise to the more reliable Xenopus laevis toad pregnancy test (1939–1960s) which involved injecting a female Xenopus laevis toad with a woman's urine (if the Xenopus laevis toad produced eggs within the next 24 hours the test was positive). Most women at this time were not offered a laboratory pregnancy test by their doctor and the 'toad test' was only used for cases where urgent diagnosis of pregnancy was required, or some other special circumstance.

In 1940, the 'treatment of delayed menstruation with prostigin' (a synthetic form of natural progesterone) as a 'therapeutic test for early pregnancy' was reported in the Journal of the American Medical Association (Soskin, 1940). By 1950 Schering had marketed a compound named Duogynon in West Germany as a pregnancy test and to treat secondary amenorrhoea. Duogynon was initially an injectable drug that contained natural progesterone and estradiol benzoate, the two hormones required for a successful pregnancy. Duogynon was subsequently marketed as 'dragees' (or tablets) that contained NETA and EE (synthetic versions of estrogen and progesterone) and tablets were taken over two days. If menstruation occurred in the days following, the pregnancy test was considered negative.

If implantation of a fertilised ovum does not occur (ie. a woman is not pregnant) a drop in the level of natural (endogenous) progesterone prompts shedding of the endometrium, causing a menstrual bleed; a similar bleed occurs during the pill-free week in non-pregnant women when synthetic progestogens are taken as part of oral contraceptives. This shedding process is called a 'withdrawal bleed'. Duogynon worked on the principle that a pregnant woman would be unlikely to have a 'withdrawal bleed' after taking the second Duogynon tablet because her own high levels of estrogen and progesterone would support the continued growth and development of the endometrium and would offset the impact of a fall in levels of the hormones in Duogynon.

In Britain Duogynon was first marketed by Schering in 1958 as 'Primodos' and by this time, a number of other pharmaceutical companies had launched similar 'clinical', 'hormonal' or 'withdrawal bleeding' pregnancy tests (see table 3). Despite questions about their reliability,

it was recognised that HPTs offered a more accessible, quicker and cheaper method of diagnosing pregnancy than laboratory testing.

2.1.2 Early medicines regulation

In the 1950s and early 1960s there were no legal requirements on pharmaceutical companies to ensure that marketed medicines met appropriate standards of safety and effectiveness; any studies performed on the drug were at the discretion of the pharmaceutical company. Similarly, with no centralised system for reporting suspected adverse reactions, any hazard would have been difficult to detect unless it was substantial, or the condition occurred naturally only very rarely or was unusual and deemed worthy of publication in a medical journal.

The association of the morning sickness medicine, thalidomide, with unusual congenital anomalies (primarily phocomelia) led to its withdrawal from the UK market in 1961 and resulted in a number of key developments relating to the regulation of medicines. One of the first of these occurred in March 1963 when a Joint Subcommittee of the English and Scottish Standing Medical Advisory Committee made the following recommendations:

- The responsibility for testing new drugs in clinical trials before they are used should remain with each individual manufacturer; and
- A Committee on Safety of Drugs (CSD) should be established, with four subcommittees to review evidence and advise on: i) toxicity of new drugs ii) clinical trials iii) therapeutic efficacy and iv) adverse reactions.

In June 1963, the UK's CSD was established. At that time, the CSD had no legal powers and operated an entirely voluntary scheme, working with the pharmaceutical industry, which meant companies undertook to seek CSD advice before marketing a new medicine. Between 1964 and 1967 new drug submissions started to be considered by the CSD, leading to the establishment of a number of important principles of modern medicines regulation including that:

- Medicines should not be released for testing in women of child bearing age until the appropriate tests in animals had been performed; and
- A minimum of teratogenicity testing in two animal species was required.

The Medicines Act 1968 (chapter 67) consolidated into a single Act the most necessary and desirable features of all previous rules, regulations and Acts in the UK. It covered all aspects regarding the control of medicines for human and veterinary use, from approval of the marketing authorisation to its withdrawal. The Act came into force in 1971 and was subsequently amended frequently to ensure it remained in line with European Community legislation. It has now been largely repealed and replaced by the Human Medicines Regulations 2012 (S.I. 2012/1916, as amended).

The Subcommittee on Adverse Reactions (SCAR) was set up by CSD to promote the collection and investigation of information relating to suspected adverse drug reactions (ADRs). In 1970 the CSD became the Committee on Safety of Medicines (CSM) whose primary function was to give advice to UK ministers relating to the quality, efficacy and safety of any substance to which any provision of the Medicines Act 1968 (c.67) applied.

In the UK HPTs were therefore prescribed by doctors for use by women for over a decade before the full extent of the thalidomide tragedy was known and in an environment where there was little statutory control over the manufacture, marketing, promotion and supervision of supply and terms of use of medicines.

2.2 Scientific context

This section provides an introduction to the hormonal steroids, the roles of the natural (endogenous) and synthetic sex hormones, the process of normal human embryonic development and the development of congenital anomalies.

2.2.1 Hormonal steroids

The hormonal steroids, including the sex hormones, are part of a class of low molecular weight compounds with a common ring structure derived from cholesterol. They have high lipid (fat) solubility and so are readily absorbed across biological membranes. Naturally occurring steroid hormones are broken down in the body by a number of steroid metabolising enzymes. This activity results in the synthesis of a large number of metabolites which can exert different biological effects. This review focuses on estrogens and progestogens, whose functions relate mainly to the development of secondary sexual characteristics, fertility and reproduction.

2.2.2 Role of sex hormones in the mother and fetus

Naturally occurring estrogens (estrone, estradiol, and estriol) and progesterone are essential for the sexual and reproductive development of women. Working together, estrogens and progesterone act via specific receptors to stimulate growth and differentiation in cells of the key target tissues (eg endometrium and mammary).

Estrogens and progesterone at high concentrations prepare the uterus for pregnancy and act to maintain human pregnancy after implantation of a fertilised ovum or ova. If fertilisation occurs, the corpus luteum (a temporary endocrine gland of the ovaries) begins to produce high levels of progesterone. At this early stage, progesterone has many diverse functions that are vital to the establishment and support of early pregnancy. While high levels of progesterone dominate throughout pregnancy, estrogens are also very important, with many of the functions of progesterone requiring prior stimulation of the tissue by estrogen. By the end of the first trimester, the role of sex steroid hormone production, particularly progesterone production, is taken over by the placenta and production levels increase a further ten-fold.

For many years, it was thought that low levels of progesterone were a cause of miscarriage and some women who were diagnosed as having low progesterone levels, were prescribed synthetic progesterone to prevent miscarriage. Evidence for the effectiveness of such treatment is weak but studies are still being carried out to investigate this (eg. Coomarasamy, 2015). Whilst the roles of progesterone and estrogens in supporting implantation and early placental development have been defined, their role in fetal organogenesis and development is less clear.

2.2.3 Congenital anomalies and their natural occurrence

Congenital anomalies are defined by the WHO as “structural or functional anomalies (for example metabolic disorders) that occur during intrauterine life and can be identified prenatally, at birth or sometimes may be detected later in infancy, such as hearing defects”[7].

Information from Europe and the USA suggests a total prevalence of major congenital anomalies of between 24 and 40 per 1 000 births (2.4% to 4%). In both territories, congenital heart defects appear to be the most common post-natal anomaly, followed by limb defects, anomalies of the urinary system and nervous system defects.

[7] http://www.who.int/mediacentre/factsheets/fs370/en/

The cause of at least half (60%) of all post-natal congenital anomalies remains unknown with the other half having genetic or environmental causes or both, as shown in Table 2. Many genetic conditions occur in individuals with no prior family history.

Table 2. Causes of post-natal human congenital anomalies[1]

Cause of congenital anomalies	Proportion of all congenital anomalies
Unknown aetiology	60%
Multifactorial (genetic and environmental)	20%
Environmental agents of which: Recognised teratogen Maternal illness Infection at birth	7–10% 2% 3% 2%
Genetic mutations	8%
Chromosomal abnormalities	6% (prenatally 30%)

[1] As presented by Dr D Wellesley at the EWG meeting on 25 April 2016 (annex 11), adapted from Emery's Elements of Medical Genetics, 10th Edition (Mueller and Young, 1998)

Single gene defects and most chromosomal defects occur prior to conception and, in many cases of congenital anomaly, one or both of these possible causes should be ruled out before alternative aetiologies are considered. As genetic research continues to progress, it is likely that more congenital anomalies will be identified as having a genetic cause.

2.3 Hormone Pregnancy Tests in the UK

2.3.1 Synthetic sex hormones

Synthetic sex hormones that mimic naturally occurring estrogens and progesterone, the latter termed progestogens, are used in a range of gynaecology indications, most notably contraception and the relief of menopausal symptoms. High dose progestogens, including NETA, are also widely used to prevent recurrent or threatened miscarriage), treat endometriosis, stop or delay menstruation and in assisted conception (*in vitro* fertilisation, or IVF).

Identifying the unique actions of the many different synthetic sex hormones is challenging because each can have more than one type of effect, depending on the molecular structure, the dose, the route of administration and whether it is co-administered with other sex hormones. There is also extensive interplay between the effects of steroid hormones. For example, prior exposure to estrogen is required for progestogenic effects to occur in most, if not all, tissues due to the estrogen-induced expression of progesterone receptors (PR).

Synthetic hormones can mimic the action of endogenous hormones by directly activating the receptor (agonist action), they can oppose the action of the natural hormone by binding to the receptor but not activating it (antagonist action) or they can have a mixed effect by binding to the receptor and activating it less well than the endogenous hormone (mixed agonist-antagonist action).

2.3.2 HPT products

A number of products containing a synthetic estrogen or progestogen or both were marketed in the UK from 1950 for use as HPTs and a range of other gynaecological conditions (Table 3).

Table 3. Hormone Pregnancy Test products available in the UK.

Estrogen/progestogen combination	Products	Dates available in UK
Ethinylestradiol and ethisterone	Amenorone	1950 to 1977
	Amenorone Forte	Until 1977**
	Disecron	Pre-1952* to 1969
	Menstrogen	1951 to 1975
	Orasecron	1950 to 1975
	Paralut tablets	NA
	Paralut Forte tablets	NA
	Pregornot	1973
Ethinylestradiol and norethisterone acetate	Norlestrin	1961 to 1975
	Norlutin-A	1961 to 1975
	Primodos oral	**1958 to 1978**
Ethinylestradiol and dimethisterone	Secrodyl	1961 to 1975
Estradiol benzoate and progesterone	Paralut injection	NA
	Paralut Forte injection	NA
	Primodos injectable	NA
Norethisterone	Norone	1965 to 1969

NA information not available
* exact date not known
**withdrawal date only known

A more detailed breakdown of the composition of these products, their licensed indications, when they were first available in the UK and when they were withdrawn from the market is provided in Annex 3. In many countries, a product equivalent to Primodos tablets was available and was called Duogynon or Cumorit. Injectable forms of Duogynon were also available but these contained natural forms of the sex hormones (estradiol benzoate and progesterone) rather than the synthetic forms found in Primodos (EE and NETA). HPTs were available in many countries worldwide but following decline in their use in the 1970s most were withdrawn by the manufacturers. In the UK, Primodos was the last HPT to be withdrawn in 1978.

2.3.2.1 Primodos

The synthetic steroids that were found in Primodos have been in clinical use for more than 50 years (see section 2.3.1). EE is the estrogenic component in a great many combined hormonal contraceptives and NETA is the progestogenic component in many combined hormonal contraceptives and hormone replacement therapies. EE binds to the estrogen receptors (ER) in the body and, while NETA binds to a range of receptors, it largely binds to PR.

NETA is derived from 19-nortestosterone and is structurally similar to both natural progesterone and testosterone. In the body, NET primarily mimics the actions of progesterone through its binding to the PR but is also a strong activator of the androgen receptor. It can also inhibit the androgen receptor and may have estrogenic and anti-estrogenic activity, depending on the conditions. The action that NET exhibits is therefore complex and will depend on its dose, route of administration, duration of use, the presence or absence of other hormones and the presence or absence of different hormone receptors.

EE is a semi-synthetic estrogen derived from natural estrogen (estradiol) to produce a compound that is much more resistant to inactivation by the liver. EE exerts potent estrogenic effects through its action at the ER and has similar or slightly stronger estrogen agonist activity than the naturally occurring estrogens.

Primodos was reformulated twice and therefore available in three different dosing schedules:

- From 1958 to 1960: each tablet contained NETA 10 mg and EE 50 µg (0.05 mg); one tablet to be taken each on four consecutive days.
- From 1960 to 1963: each tablet contained NETA 5 mg and EE 10 µg (0.01 mg); one tablet to be taken on each of two consecutive days.
- From 1963 to 1978: each tablet contained NETA 10 mg and EE 20 µg (0.02 mg); one tablet to be taken on each of two consecutive days.

No currently licensed medicines in the UK contain progestogens and estrogens in combination at the same doses as were present in HPTs. However, these hormones remain very common components of a wide range of medicines that are effective in the treatment of a broad range of gynaecological disorders, menopausal symptoms, cancer and in contraception.

The final version of Primodos contained 10 mg NETA per tablet. Historically, early NETA-containing combined contraceptive pills included doses of NETA of between 1 mg and 4 mg. The concentrations of NETA currently found in combined contraceptives typically range from 0.5 mg to 1.5 mg with one tablet generally taken every day. Today NETA is given at doses of 15 mg per day for 10 days to delay menstruation and at doses of up to 60 mg per day for disseminated breast cancer.

The final version of Primodos contained 20 µg EE per tablet. Historically, the first daily combined oral contraceptives used in the 1960s contained up to 100 µg EE per tablet, with lower dose pills containing 35 µg to 50 µg EE per tablet being introduced during the 1980s. The concentrations of EE currently found in combined contraceptives typically range from 15 µg to 35 µg with one tablet generally taken every day (with or without a seven day break every month).

2.3.3 Usage of HPTs in the UK

No reliable data on exposure to HPTs in the UK were identified during the review and so information was derived from two historical sources: a report by Dr Isobel Gal to the Minister for Health in 1978 (Annex 12) and a report by the Medical Director of Schering (RA Wiseman, report undated but circa 1981, Annex 13).

The estimates of exposure by Dr Gal were based on data of GP prescriptions for HPT products from Medical Data Index (MDI); MDI figures do not include sales to hospitals or clinics under the Area Health Authority and are therefore an underestimate. The source of the information on the proportion of use of HPTs that was for pregnancy testing (rather than secondary amenorrhoea, the other indication for HPTs) is not documented (Table 4). These data demonstrate the extent to which HPTs were used over the years and how their popularity for diagnosing pregnancy fluctuated.

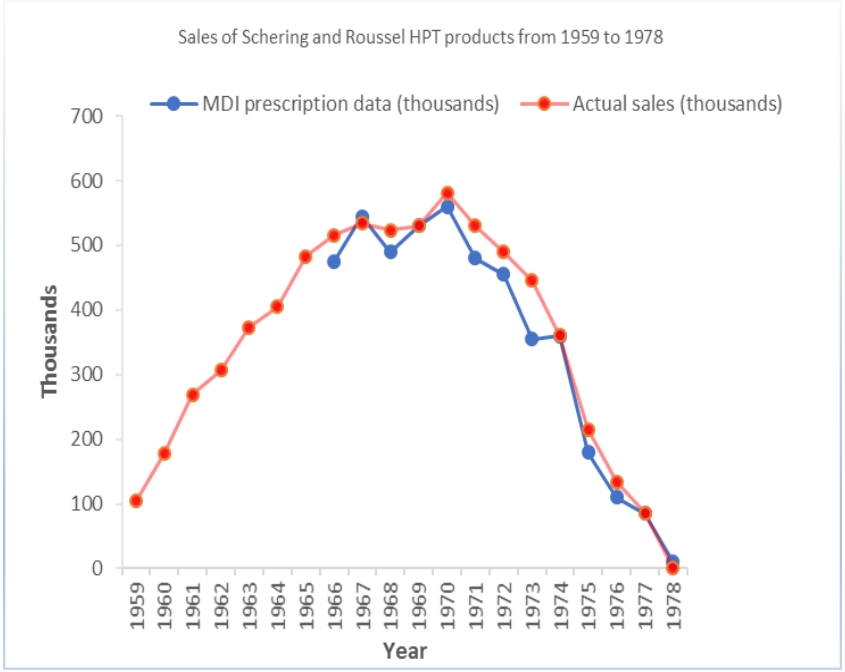
Table 4. Proportion of all pregnancies diagnosed using HPTs from 1968 to 1977 (based on Dr Gal's report to Roland Moyle in 1978)

Year	Total HPT use	Percentage of total used as pregnancy test*	Total HPT use as pregnancy test**	Live births in England and Wales***	Estimated percentage of mothers of all live births given HPTs
1968	1 212 000	15.1	183 012	819 272	22.3
1970	1 314 000	17.0	223 380	784 486	28.5
1971	1 202 000	19.4	233 188	783 155	29.8
1972	1 236 000	16.4	202 704	725 440	27.9
1973	1 036 000	11.5	119 140	675 953	17.6
1974	1 008 000	11.8	118 944	639 885	18.6
1975	371 000	12.8	47 488	603 445	7.9
1976	280 000	7.0	19 600	584 270	3.3
1977	306 000	2.3	7 038	569 259	1.2
Total	**7 965 000**	**-**	**1 154 494**	**-**	**-**

* Source for the proportions presented in the Gal report (1978, Annex 12) is unknown.
** Calculations based on the values in columns 2 and 3 from the Gal report (1978, Annex 12)
*** From Appendix 2 of the Schering report (Wiseman, Annex 13).

A report prepared by Schering's Medical Director similarly provided MDI prescription data for HPTs alongside sales figures for Schering Chemicals Limited (Primodos) and Roussel Laboratories Limited (Norone, Amenorone and Amenorone Forte), which together were stated to account for around 90% or more of total HPT sales (figure 1). The report also includes the number of samples of Primodos distributed by Schering, which are said to have fallen from 25 539 to 150 during the period 1966–1968.

Figure 1. Actual sales and MDI figures for usage of Schering and Roussel HPT products from 1959 to 1978 (Annex 13)



As with Dr Gal's figures, both actual sales and MDI prescription data show that the use of HPTs peaked around 1970 to 1971 after which their use fell steadily until their withdrawal from the market in 1978. The reason for the two to three-fold discrepancy in the estimates for total HPT prescriptions between Dr Gal's data and the Schering data (both derived from MDI) is not known. There is also a discrepancy between the estimates of Dr Gal and Schering for the proportion of HPTs that were used to diagnose pregnancy: a document from Schering Chemicals Ltd (dated 1967) suggested that in 1966 around 72.6% of HPTs use was for pregnancy diagnosis (75.3% for Primodos and 69.9% for Amenorone Forte)[8], this compares with Dr Gal's much lower estimate for that time of about 15%.

The background rate for congenital anomalies is at least 2% in the population. So, for the 223 380 women who were estimated to have been exposed to an HPT in 1970 (using Dr Gal's estimates for exposure to an HPT in pregnancy, table 4) the number who would be expected to have given birth to a child with a congenital anomaly, irrespective of HPT use, would be about 4 500.

For the 413 820 women who were estimated to have been exposed to an HPT in 1970 (using Schering's estimates of 570 000 for total exposure[9], and 72.6% for the proportion of use in pregnancy) the number who would be expected to have given birth to a child with a congenital anomaly, irrespective of HPT use, would be about 8 300.

If the same estimate is calculated using Dr Gal's figure of 1 314 000 women exposed to an HPT and Schering's figure of 72.6% for the proportion of use in pregnancy, a total of 953 964 women would have been exposed to an HPT in pregnancy in 1970. Of these, as many as 19 000 would be expected to have given birth to a child with a congenital anomaly, irrespective of HPT use.

2.4 Key UK regulatory action taken on HPTs

The sequence of regulatory actions leading up to the withdrawal of HPTs from the market in the UK can be deduced from historical records.

In 1967, the CSD considered the unpublished findings of a study by Dr Gal, which suggested there may be an association between HPTs and spina bifida in the offspring of women who had taken these products. No action was taken on the basis that the study was considered to suffer from a number of methodological limitations that cast doubt on the reliability of the findings. However, together with other emerging data, these findings, and concerns relating to other drugs used in pregnancy, prompted the CSD to start their own study: 'Maternal drug histories and congenital anomalies'.

In 1969, Schering stopped promoting Primodos for the diagnosis of pregnancy. Then in 1970, following a recommendation by the UK Standing Joint Committee on the Classification of Proprietary Preparations (otherwise known as the MacGregor Committee) that pregnancy tests should no longer be reimbursed by the health service, Schering removed the indication from the data sheet for Primodos (Annex 14[10]). Primodos remained available for the treatment of secondary amenorrhoea.

In 1975, CSM published the preliminary findings of their 'Maternal drug histories' study in a letter to the British Medical Journal (BMJ, Greenberg, 1975) and issued a warning to all prescribers advising them not to use hormonal tests for diagnosing pregnancy because of the possible hazard and the availability of other means of diagnosing pregnancy (Annex 15).

[8] Primodos and Amenorone Forte accounted for approximately 75% of the UK market share, personal communication from Bayer.

[9] Appendix 2 of Schering report (Wiseman, Annex 13)

[10] First ABPI Data Sheet Compendium, published in 1974

The data sheet for Primodos was changed in 1975 to introduce a contraindication in pregnancy (Annex 16).

In 1977, reports that Primodos was still being used as a pregnancy test, prompted CSM to issue a reminder to all UK doctors and pharmacists that these products should not be used for this purpose (Annex 17). The timing of the reminder coincided with publication of the final results of CSM's 'Maternal drug histories' study (Greenberg, 1977).

Records suggest that in the UK, some HPT products were discontinued as early as 1969, with the majority being discontinued before 1975. Primodos was the last available HPT in the UK and was removed from the market by Schering in January 1978.

Globally, different HPT products were similarly removed from the market over the course of several years. The motivations for their removal differed by country and by product and only in some countries was this based on possible safety concerns. In many cases the withdrawn HPT product was replaced by a different HPT product or by the same product with a different name. In addition, modern pregnancy tests began to be more widely available during the 1970s.

A more detailed chronology of events which describes the actions taken by: Schering AG in Germany; Schering Chemicals Ltd in the UK; the UK regulator; other regulatory agencies; the press; UK Parliament and the wider Government, including the Department of Health (formerly the Department of Health and Social Security, DHSS); and the 'Association for Children Damaged by HPTs', may be found in Annex 3.

3 INTRODUCTION TO PHARMACOVIGILANCE AND STRENGTH OF TYPES OF EVIDENCE

The EWG considered evidence from a wide variety of sources to evaluate a possible association between use of Primodos and adverse outcomes of pregnancy. Different types of evidence provide different perspectives and each source has its own strengths and weaknesses. This chapter provides a brief introduction to the process of pharmacovigilance and outlines the strengths and limitations of the different types of data considered during the EWG review, and the strength of different types of evidence for a causal association.

3.1 Pharmacovigilance

Any drug taken to treat a medical condition has potential risks and benefits. Today there are measures in place to ensure that, prior to a pharmaceutical company applying for a licence for a new medicine, randomised controlled trials (RCTs) are carried out to assess its efficacy and safety. By randomising patients to either the study drug or placebo or an active comparator, it is possible to assess whether any apparent beneficial effects are due to the drug, rather than to differences between the patients. However, the limited size and duration of most RCTs means they may detect only commonly occurring adverse effects (eg. that occur in at least 1 in every 10 people who take the drug), or that occur relatively close to start of treatment. RCTs typically do not include sufficient people to detect rare adverse effects, are too short to detect longer-term adverse effects, and mainly exclude pregnant women.

Consequently, a process is in place to monitor the safety of drugs in all patient populations to ensure that any less commonly occurring adverse effects, longer-term effects or effects in pregnancy are detected and any changes in the known safety profile can be identified and acted on. This monitoring is known as pharmacovigilance.

If a new safety concern is identified when the medicine is on the market, a number of regulatory options may be considered in order to prevent or minimise harm to patients, depending on the seriousness of the concern. Rarely, if the newly identified harm means that the risks with a medicine now outweigh the benefits, it may be necessary to remove the medicine from the market. More often, the risk of an adverse effect for individual patients may be avoided or reduced by one or more of the following measures: including warnings in the product information (the Summary of Product Characteristics, SmPC, and Patient Information Leaflet, PIL) or on the package label; reducing the dose; adding contraindications to use; or restricting the indications for use of a medicine. Any such action should be proportionate to the level and likelihood of harm to patients from the identified risk. Communication to health professionals and patients of important information on the nature of the risk is essential to support informed choices about treatment options.

A more detailed description of the process of pharmacovigilance and how it has evolved over time, based on experience, is presented in Chapter 7.

3.2 Sources of evidence – the importance and impact of study design

Pharmacovigilance involves continuous surveillance of all available evidence from a wide variety of sources, and assessing the impact of each new piece of information on the balance of benefits and risks for patients taking the drug. To reach a conclusion on the evidence, it is necessary to judge its robustness, or reliability, to provide an unbiased answer to the question that was asked.

The sections below cover the different sources of evidence with a discussion of the advantages and disadvantages of each.

3.2.1 Randomised controlled trials

RCTs enrol individual patients and randomise them to either receive the treatment of interest or a comparator treatment or placebo. In general, RCTs are regarded as being the best study designs for establishing a causal association, because the randomisation process should minimise any bias that might exist and confound the outcome. A meta-analysis of randomised controlled trials is considered by many to be the most robust form of evidence because by increasing the number of study participants it increases the ability to measure an effect if one exists and examines consistency of effects across a number of trials.

Individual RCTs are generally limited by size due to feasibility or cost issues. In the absence of a meta-analysis, many RCTs will not be sufficiently large to detect rare outcomes. RCTs typically address the claimed benefits of treatments and focus on only one or two safety questions. As a result, likely harms need to be considered or predicted from the outset, and so unexpected harms may be less likely to be studied. Additionally, RCTs tend to exclude patients in vulnerable groups (such as pregnant women) and those with significant co-morbidities. The patient population in a RCT may therefore not be representative of the population taking a medicine in the real-world.

3.2.2 Observational studies

Observational studies are those in which individuals are observed under normal clinical conditions and certain outcomes are measured. This type of study allows for large numbers of patients to be included and tend to be much more representative of the patient population than RCTs. They are generally considered less robust than RCTs because they are typically subject to a number of biases which have to be carefully considered and controlled for in the analysis.

Observational studies can be either hypothesis-testing (those that seek to answer a specific question) or hypothesis-generating (those whose aim is to generate questions for further research) studies. The general strengths and limitations of the different types of observational studies are well documented (eg. guidance provided by the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, ENCePP, 2017).

3.2.2.1 Hypothesis-testing studies

All observational studies are not equal and those conducted prospectively – typically cohort studies – are generally considered to be more robust than case control studies, which are retrospective in nature and may need to rely on an individual's recall of events or treatments or on data that was collected for a different purpose (so key information may not be available). Case-control studies are generally better able to detect rare outcomes. A critical aspect of the robustness of an observational study is how well designed it is in terms of minimising bias and confounding.

3.2.2.2 Hypothesis-generating studies

Ecological studies examine the relationship between an exposure and outcomes at population rather than individual level eg. trends in the incidence of a disease in relation to prescribing trends of a medicine. This type of study is considered to provide less strong evidence than cohort or case-control studies because it infers associations at population level that may or may not exist at individual level. Timing of exposure in relation to the outcome is not considered.

A single-arm study (ie. one without a control group) can only be considered as hypothesis-generating as the results cannot be put into context as to whether observed incidence/prevalence of disease is higher or lower than expected.

Surveillance systems are designed to routinely monitor trends in the reporting of congenital anomalies in order to identify trends. Any trends may then be investigated by other means for possible associations with potentially causal factors. These can include exposures to drugs but also to other factors such as environmental hazards and infections. These systems are not usually designed to test for causation themselves.

3.2.3 Reports of suspected adverse reactions (spontaneous reporting data)

Case series and case reports of adverse reactions that are suspected to have occurred in association with the use of a medicine are generally considered one of the least robust sources of evidence and cannot be used to assess causal relationships between drugs and outcomes. A report of an adverse reaction does not necessarily mean that it is due to the drug in question as reporters are encouraged to report *suspicions* of an adverse reaction. However, spontaneous reports are valuable for hypothesis-generating activities. Also known as signal detection, the frequency with which a specific ADR is reported in association with a drug compared to the frequency with which it is reported with other drugs or other events in the same database can provide an indication of a drug safety issue.

Well recognised limitations of spontaneous suspected ADRs include that there is a variable and unknown degree of under-reporting for a number of possible reasons including: lack of time to report; lack of awareness of a reporting scheme; the possible association may go unrecognised by patients and prescribers, particularly if the effect occurs relatively frequently in the untreated population. Under-reporting therefore makes it impossible to calculate an incidence rate or absolute frequency for an ADR. Reporting may also be stimulated by the availability of a new drug or media attention about a possible safety concern.

3.2.4 Animal research

Studies in animals are important for understanding mechanisms by which drug substances exert their effects. They are also extremely valuable for rapidly screening out potentially harmful drug substances based on their effects in animal models and can be useful for detecting post-marketing safety concerns. Their inherent limitation is that their physiology differs from that of humans, which makes it difficult to judge how (well) their findings may relate to humans.

The aim of reproduction toxicity studies is to reveal any effect of one or more active substance(s) on mammalian reproduction. For this purpose, both the investigations and the interpretation of the results from animal studies should be related to all other pharmacological and toxicological data available to determine whether potential reproductive risks to humans are greater, lesser or equal to those posed by other toxicological manifestations. Further, repeated dose toxicity studies in animals can provide important information regarding potential effects on reproduction. To extrapolate the results to humans (ie. assess the relevance of the findings), data on likely human exposures, comparative kinetics and mechanisms of reproductive toxicity are essential.

3.2.5 *In vitro* or laboratory based research

In vitro studies carried out using animal or human cells provide simplified systems, which allow drug effects to be observed more directly. However, it is impossible to recreate a system that is as complex as the situation in the living body and the physiological control systems that are present in an organ or the whole body may be lacking. The effects that are observed in a test tube environment are therefore likely to differ from the effects in the human body.

3.2.6 Ideas, editorials, anecdotes, letters and opinions

Personal observation can highlight potential benefits and harms that have not previously been noted. As such, personal experience can provide valuable insight, both into new safety concerns and possible causes of harms. However, it can also be subject to limitations due to an over-reliance on striking anecdotal occurrences, a tendency to accept evidence that supports personal beliefs at the expense of evidence presented against that belief, or insufficient reliance on statistically significant strong evidence. Such data are considered to be hypothesis generating.

3.3 Strength of the evidence in this review

3.3.1 General considerations

The quality of each individual piece of data is of great importance; randomisation will always provide for the strongest causal inference, but may not have statistical power to detect rare adverse effects and this applies particularly to congenital anomalies. Hence, a well-designed observational study may provide stronger evidence to answer questions regarding rare events than the available randomised controlled trials (Glasziou, 2004).

A critical consideration during the EWG review was that, being carried out in the 1950s to 1970s, the design, conduct and quality of the studies were largely poorer than would be expected of those conducted today. This most likely reflects the lack of available databases and adequate routine data collection at that time coupled with advances in knowledge and understanding of study design and analysis. Similarly, when HPTs were first marketed, pharmaceutical companies were not legally required to ensure that the medicines they marketed met appropriate standards of safety, quality and efficacy. This had an important impact on the type of data available for the review, the quality of those data, and the interpretation of their findings.

A very careful evaluation of the available data was therefore necessary to determine whether, on balance, it supported a possible association between congenital anomalies in the children of mothers who had been given an HPT during early pregnancy as being causal, or whether its limitations were such that the association was more likely to have been due to chance or to other factors.

The EWG therefore considered it necessary to set out what would constitute different strengths of evidence to support a causal association between use of HPTs and an adverse outcome of pregnancy for the different types of available data.

3.3.2 Epidemiological data

If there was a causal association between exposure to HPTs in pregnancy and congenital anomalies in the offspring the EWG considered that:

- The best evidence would be consistent findings from randomised controlled trials (or a meta-analysis thereof) for a statistically significant increase in risk of an anomaly or anomalies in the offspring of those randomised to HPTs versus those randomised to no HPT. In these studies chance, bias and confounding could be ruled out with reasonable confidence.
- Good evidence would be consistent findings from very well conducted observational studies (or a meta-analysis thereof) for a statistically significant increase in risk of an anomaly or anomalies in the offspring of those who had been given HPTs compared with those who had not. In these studies, appropriate steps would have been taken to minimise chance, bias and confounding.

- Limited evidence would be mostly consistent but non-statistically significant findings from studies which are still at some risk of bias or that are too small to detect a statistically significant association should one exist, or consistent statistically significant findings from poor quality studies where the estimated magnitude of risk is such that it could not be explained by biases alone. A key bias is in comparing those who had been given a pregnancy test compared with those who had not.
- Insufficient evidence would be where a limited number of studies do not allow a conclusion to be reached about the likelihood of a causal association.

The EWG also considered the following criteria to denote the strength of any observed association between HPTs and congenital anomalies:

- An extremely strong association would typically refer to a large effect size (relative risk/odds ratio) that was unlikely to be due to chance or bias
- A weak association would typically refer to a small effect size (relative risk/odds ratio) that could be due to chance or bias
- No association would typically refer to risk estimates where half or more studies find no increased risk (i.e. risk estimates close to or below 1), and where studies are less likely to be susceptible to bias.

3.3.3 Adverse event reporting data

Bearing in mind that adverse event reporting data can neither absolutely confirm nor refute a causal association, particularly when considering outcomes in children following a single exposure in-utero, the only form of evidence that would highlight a possible association between HPTs and congenital anomaly(ies) would be: a consistent increase in the proportion of a particular anomaly or pattern of anomalies in association with HPT use, compared with the general population. Furthermore, the two sets of data would need to be as comparable as possible, and so the timeframe for reporting would need to be very similar, as large as possible, and all events reported with an HPT would ideally need to have been reported contemporaneously, based on suspicion of an association.

3.3.4 Animal data

With respect to evidence from animal studies being supportive of a causal association between administration of an HPT to women and development of an anomaly or anomalies in the offspring:

- The best evidence would be findings for an increase in malformations seen in treated animals relative to those found in controls in well-conducted studies with sufficient sample sizes, in multiple appropriate species with clinically relevant exposures of the test drug.
- Good evidence would be findings for an increase in malformations seen in treated animals relative to those found in controls in well-conducted studies, in more than one species with clinically relevant exposures of the test drug.
- Limited evidence would be findings for an increase in malformations seen in treated animals relative to those found in controls in relatively well-conducted but small studies, in a single species with clinically relevant exposures of the test drug.
- Inadequate evidence would be findings for a small increase in malformations relative to controls in relatively well-conducted but small studies, in a single species with clinically irrelevant exposures of the test drug.

3.3.5 Overall conclusion

When drawing overall conclusions, the EWG agreed that all the available evidence must be taken into consideration such that if limited evidence were to be identified from several different types of data this would add some weight to the findings.

More detailed information on the strengths and limitations of each of the different types of data reviewed by the EWG to address the question on a possible association between HPTs and adverse outcomes in pregnancy is included in Chapter 5 and Chapter 6.

4 PHARMACOLOGICAL CONSIDERATIONS RELATING TO EFFECTS OF HORMONE PREGNANCY TESTS ON THE DEVELOPING FETUS

The third to eighth weeks of human embryonic development (equivalent to the fifth to tenth weeks of pregnancy) are critically important for normal development as this is the time when most major organs and body systems are formed. This time is called the period of organogenesis, or embryogenesis, and is thought to be when the embryo is most sensitive to genetic or environmental factors. Most structural congenital anomalies are thought to be induced during the time of organogenesis.

Knowledge of the sensitive period for each human target organ development helps to assess whether it is possible that a medicine taken during pregnancy may have caused a specific anomaly. Other factors also play a role in determining if a medicine may have a teratogenic effect, for example, its dose (intensity of exposure), how effectively it crosses the placenta (actual fetal exposure), how long it is taken for and how long it lasts in the body (duration of exposure).

In reaching a conclusion on whether Primodos could have had an effect on the developing embryo the EWG considered that certain pharmacological conditions would need to be fulfilled:

i. It must be administered during the critical period of embryonic development
ii. It must be able to cross the placental barrier
iii. The embryo must express receptors that can bind the drug
iv. These receptors must be expressed at the right period of organogenesis
v. The receptors must be functional (ie. able to bind and be activated)
vi. The drug should be present at concentrations sufficiently high to elicit a pharmacological response.

This chapter discusses what is known from limited studies conducted *in vitro*, in animals and in humans about the way in which norethisterone (NET) and ethinylestradiol (EE) are metabolised and, based on a series of extrapolations, considers whether the above criteria are fulfilled. Some numbers calculated in the earlier sections of the chapter are used in later calculations and are presented in bold font to make it easier to see how numbers have been derived.

Much of the information discussed in this chapter assumes a basic understanding of pharmacology and pharmacokinetics, and therefore may not be suitable for all readers. Further information is provided in Annexes 18 and 19 that may be helpful.

4.1 Timing of administration of HPTs

To cause a congenital anomaly, any medicine administered during pregnancy would generally need to be given during the critical period of organogenesis, when all the organs and body systems are developing.

Population data on the exact timing of exposure to HPTs were not available. However, HPTs were expected to have been used at least two weeks after a woman's first missed period, ie. from about six weeks of gestation (or four weeks of fetal development). Hormone Pregnancy Tests would also have been used in a proportion of women after two missed periods (particularly those with irregular menstruation) but it would seem unlikely that they would

have been used after three missed periods. Although use outside this time frame would seem unlikely, some earlier or later use could not be excluded.

As a conservative estimate the likely window for HPT use was therefore considered to extend from the week of the woman's first missed period to the end of the first trimester; that is, 4 to 12 weeks of pregnancy (2 to 10 developmental weeks). Since this covers most of the critical period of fetal development the first criterion for a possible drug-related effect was considered met.

4.2 Crossing the placental barrier

During pregnancy, the placenta forms as an intermediate organ between the growing embryo and the maternal uterus and serves to provide the embryo with oxygen and nutrients and remove embryo-fetal waste products. After HPT use, NET and EE would have been present in the maternal circulation during the critical period of organogenesis; the question is whether or not NET and EE were able to cross the placental barrier to the embryo.

The fetal circulation is separated from the maternal circulation by the syncytial membrane of the placenta and endothelial cells of the fetal capillaries. Hence the concept of the 'placental barrier' developed, which was initially thought to act as a protective mechanism against infectious and other potentially harmful agents. Before the thalidomide tragedy in the 1960s, doctors believed the placenta was also a barrier to drugs. It is now known that maternal hormones and antibodies readily cross the placenta, as do most drugs and their metabolites, many viruses and other agents.

Since both NET and EE have relatively low molecular weights and are lipophilic they would be expected to cross into the placenta in the same way as the natural estrodiol and progesterone in pregnancy would do. However, the amount of NET and EE that reaches the embryo would be influenced by how they are metabolised and distributed within the mother's body, and how effectively they cross the placenta.

In addition, the placenta expresses proteins that act as transporters or "pumps", which can transport hormones and drug substances across biological membranes, both towards and away from the embryo. Although it is known that the placenta expresses a large number of transporters (Iqbal, 2012; Joshi, 2016), what is less clear is how transporter expression varies in the different stages of pregnancy and between individuals. In addition, the effect of concomitant maternal disease on transporter expression is unclear and very little work has been done in this area.

4.2.1 Maternal absorption of Primodos

There is significant variation between individuals in the levels of NET and EE in the blood after administration (Orme, 1989). This variation can be due to differences in body weight, absorption, metabolism, tissue receptor levels, diet, and drug interactions. Oral bioavailability particularly was highly variable, most likely due to differences in the extent of first pass effects in the gut wall or liver.

Mathematical models built to describe how the physiological changes in pregnancy affect the disposition and metabolism of drugs showed changes in various pharmacokinetic and physiological parameters in the third trimester (such as increased volume of distribution, increased cardiac output, and increases in CYP 3A4 enzymes responsible for hepatic and gut metabolism) (Xia, 2013). These models suggested that being pregnant would decrease the exposure and plasma concentrations of NET and EE in the mother compared to the same dose(s) taken by a non-pregnant woman.

As these models were developed mainly from second and third trimester pregnancies it was unclear to what extent the observed pregnancy-related changes would be apparent during the first trimester. If the changes were already fully established, this would decrease the effective dose of NET and EE from HPT exposure relative to non-pregnancy. However, if these changes were not significant during the first trimester, exposure levels for NET and EE from HPTs would be similar to those observed in non-pregnant women.

One published clinical trial of pregnant women awaiting termination of pregnancy (Pulkkinen, 1984) looked at the pharmacokinetic parameters of NET (see section 4.2.1.1). Maximal total plasma NET concentrations were stated to be similar to data from non-pregnant females, although no direct comparisons were made in this study. Comparable pharmacokinetic data for EE were not identified. It was therefore assumed that maternal blood levels of NET and EE in the first trimester were comparable to those in non-pregnant women.

4.2.1.1 Maternal blood levels of norethisterone

After oral administration, NETA is rapidly hydrolysed to NET in the intestinal tract and liver. Therefore, the pharmacokinetics and pharmacodynamics of NET following treatment with NETA or NET are similar.

NET is completely absorbed when given orally and underwent extensive bio-transformation, mainly in the liver such that only 60% of the administered dose reached the blood. Metabolites conjugated with glucuronic or sulfuric acid were present in plasma in at least similar amounts to unconjugated NET but are expected to be inactive. NET is also metabolised to EE; the precise amount produced is uncertain but may be similar to that from a 0.03 mg dose of EE. In the blood, about 97% of NET circulated bound to plasma proteins (about 36% to sex hormone-binding globulin, SHBG and 61% to albumin) and 3% was 'free', or unbound.

No pharmacokinetic studies with the same 10 mg dose of NET as in a Primodos tablet were identified. However, studies with administration of 5 mg NET reported maximum total serum concentrations (Cmax) of 28 ± 6.5 ng/ml (Goldzieher, 1994). The pharmacokinetic parameters of NET have been reported to be non-linear with Cmax increasing at doses above 5 mg due to saturation of binding to plasma proteins and absorption processes or both. Therefore, at doses higher than 5 mg, increased NET exposure would be expected.

Plasma half-life after a dose of 5 mg NET was about 5 to12 hours and so limited accumulation was expected with daily dosing. However, additional accumulation was observed with co-administration of EE which was attributed to increases in NET plasma protein binding due to EE-induced increases in SHBG and to saturation of NET metabolism or both. Such accumulation might be expected to affect NET plasma levels in women taking a daily dose; however, accumulation would be expected to be minimal in women exposed to only one tablet of Primodos on two consecutive days.

One published clinical trial of pregnant women awaiting termination of pregnancy (Pulkkinen, 1984) looked at the pharmacokinetic parameters of NET. In this study, 10 women were given NETA 20 mg and EE 40µg (equivalent to two Primodos tablets taken at the same time) at six to seven weeks of gestation. Limited plasma samples were taken during the first six hours, which prevented reliable establishment of Cmax, however maximal total plasma NET concentrations of **74 ± 27 ng/ml** and a $t_{1/2}$ of 12 to 48h, were observed. These were stated to be similar to data from non-pregnant females, although no direct comparisons were made in this study. Based on this study, the amount of NET estimated to be free and therefore available to bind receptors on target tissues would be about 2 ± 0.8 ng/ml (ie. Cmax x Fu, or 74 ± 27 x 0.03, where Fu is the unbound fraction and 0.03 corresponds to 3%).

Although 20 mg NET was equivalent to a double dose of Primodos, because the actual Cmax was not measured accurately in Pulkkinen's study, this was taken to represent a conservative estimate of NET levels after a single dose of Primodos, to yield an estimated:

- maximum free concentration of about **2 ± 0.8 ng/ml NET** in the mother after taking a single Primodos tablet.

4.2.1.2 Maternal blood levels of ethinylestradiol

Many of the identified data on metabolism of EE related to contraceptives rather than HPTs. However, because the dose of EE in Primodos was equivalent to the dose of EE in modern contraceptives, the findings were considered to be relevant.

EE given orally was rapidly and almost completely absorbed from the gastrointestinal tract and underwent extensive metabolism in the liver to form a number of metabolites. Metabolites conjugated with glucuronic or sulphuric acid were present in amounts higher than parent EE but are inactive. Due to the first-pass metabolism of EE in the gut and liver, only 50 to 60% of the administered dose reached the blood.

Plasma concentrations of EE peaked about two hours after administration, with a second peak evident ~10 to 14h after administration (due to deconjugation of metabolites and entero-hepatic recycling). This acted to prolong the activity of EE, but contributed to a relatively small proportion of the total drug exposure.

The time for plasma concentrations of EE to halve (the half-life, $t_{1/2}$) was 18 ± 4.7 hours at steady state. This contributed to a slight accumulation of EE with repeated daily dosing. Thus, EE concentrations were found to increase by 19% when dosed for 21 days, but did not increase when dosed for six days.

After administration of a single dose of tablets containing 0.02 mg EE (the same dose as in one tablet of Primodos) in combination with levonorgestrel (a different progestogen) to 22 non-pregnant women under fasting conditions, maximum total serum concentrations (Cmax) of EE of **62 ± 21 pg/ml** were reached at 1.5 ± 0.5 hours (Shi, 1987). After six days of daily dosing, Cmax was 77 ± 30 pg/ml and occurred at 1.3 ± 0.7 hours. Similar peak concentrations (ranging from 60 to 160 pg/ml) were observed following repeated dosing with a slightly higher concentration of EE (0.03 mg) in combination with various progestogens (Orme, 1989), suggesting that EE absorption was not affected by which progestogen was co-administered.

EE circulated in human blood almost completely (97%) bound to plasma albumin. Only **3%** was therefore unbound (Fu), or 'free', and available to bind receptors on target tissues. Assuming that the Cmax of 62 ± 21 pg/ml identified after an administered dose of EE 0.02mg (+ levonorgestrel) was representative of levels following a single dose of Primodos, the relevant free concentration in non-pregnant women would be about 1.9 ± 0.6 pg/ml (ie. Cmax x Fu, or 62 ± 21 x 0.03, where Fu is the unbound fraction and 0.03 corresponds to 3%).

The pharmacokinetic parameters of EE during early pregnancy do not appear to have been measured directly. However, from the concentrations estimated in non-pregnant women above, and assuming a conservative estimate of exposure in the first trimester of pregnancy, this suggests that taking a dose of EE 0.02 mg would give:

- maximum free concentrations of about **1.9 ± 0.6 pg/ml EE** in the mother.[11]

Since no accumulation in Cmax was found in the first six days of dosing EE and levonorgestrel, similar maximum free plasma concentrations would be expected after taking a second Primodos tablet.

[11] These estimates do not take into consideration any possible contribution to EE levels from the metabolism of NET (see section 4.2.1.1).

4.2.2 Transfer of norethisterone and ethinylestradiol across the placenta

No data on placental transfer of NET and EE in humans were identified. Possible human placental and fetal levels were therefore estimated by extrapolating from the limited data available from animal studies (Tauber, 1984; Humpel, 1982). To aid interpretation, Table 5 shows how the periods of major organogenesis and duration of pregnancy in animals compare with humans.

Table 5. Comparative development times during gestation in development days.

Gestational stage	Mouse	Rat	Rabbit	Guinea pig	Rhesus Macaque	Human
Embryonic stage: primitive streak to closure of the hard palate	7–14	8–17	6–19	13–33	16/18–45/47	13/15–56/58
Full term	20–21	21–22	31–32	65–68	164–168	~266

4.2.2.1 Placental and fetal concentrations of norethisterone

Limited data on the administration of NET to rats at day 18 of gestation (ie. close to full term, see table 5) suggested that:

- the placenta contained about **20%** of the peak maternal serum concentration, measured 30 minutes after administration
- the amniotic fluid contained 11% of the maternal serum concentration
- the fetus contained **14%** of the maternal serum concentration, where 'fetal' levels related to total concentration derived from pooled fetal tissues, rather than fetal serum levels.

It was not known whether differences between animal species in placental perfusion or different stages of pregnancy would affect the amount of NET or EE transferred from the maternal circulation to the fetus. However, assuming it was valid to extrapolate from the rat studies to humans, that is, NET transferred from the maternal circulation in similar fractions to the placenta and fetus, this suggested the following:

- maximal total human placental serum concentrations of about 15 ± 5 ng/ml NET (ie. 20% of 74 ± 27 ng/ml) and
- maximal total human fetal concentrations of about 10 ± 4 ng/ml NET (ie. 14% of 74 ± 27 ng/ml).

It was also not known whether NET was protein-bound in the fetus to the same extent as in the mother. In the unlikely scenario where no binding to plasma proteins occurs in the fetus, the free fetal concentration of NET would equal the total concentration ie. 10 ± 4 ng/ml. Alternatively, if binding occurs at the same level as in non-pregnant women (ie. 97% is bound) this would yield the following:

- free human placental concentrations of about **0.4 ± 0.16 ng/ml** NET (ie. 20% x 2 ± 0.8 ng/ml) and
- free human fetal concentrations of about **0.3 ± 0.11 ng/ml** NET (ie. 14% x 2 ± 0.8 ng/ml).

4.2.2.2 Placental and fetal concentrations of ethinylestradiol

Very few data on the placental transfer of EE were identified. In four rhesus monkeys (71% term) fetal plasma levels measured 20 minutes after an intravenous dose of EE were shown to be 15-fold lower than maternal plasma levels (Slikker, 1982). Assuming maternal concentrations of 62 ± 21 pg/ml following a single Primodos dose, and similar plasma protein binding levels as in humans (3% unbound) this suggested the following:

- total human fetal plasma concentrations of about 4 ± 1 pg/ml EE (ie. 62 ± 21 pg/ml / 15) and
- free human fetal plasma concentrations of about **0.12 ± 0.04 pg/ml** EE (ie. 4 ± 1pg/ml x 0.03)[12].

In light of the above, the second criterion for a possible drug-related effect – that NET and EE must be able to cross the placental barrier – was considered met.

4.3 Receptor expression during fetal development

The role of endogenous estrogens and progesterone in fetal development has not been fully determined; however, expression of genomic ER and PR was shown in both reproductive and non-reproductive tissues of the fetus. In general, messenger RNA (mRNA) for hormone receptors was found before the receptor proteins were detected.

In the mouse, mRNA for ER was detected mainly around gestational day 12 to14, although expression was earlier in heart, mesonephric tissues (remnants of the Wolffian duct), midgut and brain (Lemmen, 1999). Binding of radiolabelled estrogens to ER was detected later, from gestational day 16 onwards. Relatively consistent results were found in studies in rats and guinea pigs, in which radiolabelled estrogen was bound to receptors at gestational days 18 and 29 respectively. It is not clear if the detection of receptor mRNA in advance of the receptors reflected the natural sequence of events, mRNA synthesis followed by protein synthesis, or the sensitivity or timing of the assays.

ER were detected in fetal reproductive tissues of most species, usually after differentiation of the gonads (in humans this occurs around week 8 of development, week 10 of pregnancy). The expression of ER mRNA and estrogen binding was also found, transiently, for several fetal tissues not associated with reproduction.

PR were mainly detected from gestational days 20 in rats, 29 in rabbits and 48 in guinea pigs. Thus, PR expression occurred sometime after that of ER and was consistent with induction of expression of PR by the action of estrogens in most tissues.

Limited animal studies therefore suggest that ER and especially PR expression may have occurred relatively late in embryonic development and around the end of the period of organogenesis in most species (guinea pigs being the exception) in both reproductive and non-reproductive tissue, though expression in the latter tissues was found to be transient. Lack of detectable receptors for estrogen or progesterone in fetal tissues during the critical period in most animals may have reflected the limitations of the methods used or it may have suggested that estrogens and progesterone were not vital for formation of the main non-reproductive body structures.

No data from human embryos on the expression of receptors for estrogen or progesterone during the first trimester were identified, but data from the second trimester were consistent with the order of receptor expression observed with laboratory animals ie. estrogen followed by progesterone.

[12] These estimates do not take into consideration any possible contribution to EE levels from the metabolism of NET (see section 4.2.1.1).

There was therefore some evidence to support the expression of functional ER and PR in the developing human fetus; however, the timing of receptor expression, the tissues in which they are expressed and when they become fully functional (able to bind hormone and become activated) remains unknown.

4.4 Placental and fetal activity of ethinylestradiol and norethisterone

It has been established that: i) HPTs were given during the critical period of development of the fetus; ii) both NET and EE can pass through the placenta to the fetus; and iii) the fetus expresses ER and PR, though evidence for their expression in non-reproductive tissue during the critical period of organogenesis was lacking.

The next question was whether the estimated fetal concentrations for NET and EE were high enough to have bound their respective receptors. This again depended on the reliability of extrapolations from the animal data to humans, the extent to which NET and EE were protein-bound in the fetal circulation and how much competing fetal endogenous hormone may have been present. Whilst it was considered unlikely that no binding to plasma proteins occurred, the actual level of binding to proteins was not known for the placenta or the fetus.

4.4.1 Norethisterone

Assuming that the level of protein binding in the human placenta and fetus was the same as in maternal plasma (97%), the concentration of NET estimated for the fetus in section 4.2.2.1 that was free (0.3 ± 0.11 ng/ml, or 1 ± 0.4 nM[13]) and therefore available to bind to any receptors that were present was approximately equal to the binding affinity (strength of attraction) of NET for human PR and so would theoretically be expected to bind to 50% of any unoccupied PR in the fetus. However, the proportion of PR that would actually have bound NET *in vivo* would have depended on whether or not endogenous progesterone in placental and fetal tissues was present and able to compete for the receptor binding sites.

Similarly, at these concentrations, free NET would have been able to bind to 50% of the unoccupied fetal androgen receptors, but the total proportion of androgen receptors that would actually have been bound by NET would likewise have depended on the levels of other androgen receptor ligands (testosterone and testosterone derivatives) in the same tissues.

By contrast the estimated free concentrations of NET were about 1 000-fold lower than its binding affinity for human ER and any significant degree of interaction would therefore have been unlikely if endogenous placental or fetal estrogens were also present.

4.4.1 Ethinylestradiol

Assuming that the level of protein binding in the human placenta and fetus was the same as in maternal plasma (97%), the concentration of EE estimated for the fetus in section 4.2.2.2 that was free (0.12 ± 0.04 pg/ml, or 0.0004 ± 0.0001 nM[14]) would only have been expected to bind to fetal ER if there was no plasma protein binding of EE in the placenta or fetus and no competing endogenous estrogens. These estimates do not take into consideration any possible contribution to EE levels from the metabolism of NET. If NET metabolism results in peak maternal blood EE levels occurring at a similar time to those from a tablet containing 0.02 to 0.3 mg EE, this could double the free and total EE levels in the fetus. At these levels

[13] NET molecular weight = 298.19, so 74 ng/ml = 248 nM; so maternal free concentration 2.2 ng/ml = 7 ± 3 nM; assuming 97% binding, fetal free concentration 0.3 ± 0.1 ng/ml = 1 ± 0.4 nM

[14] EE molecular weight = 296.18, so 62pg/ml = 0.2 ± 0.07 nM; so maternal free concentration 1.9 pg/ml = 0.006 ± 0.002 nM; assuming 97% binding, fetal free concentration 0.12 ± 0.04 pg/ml = 0.0004 ± 0.0001 nM

EE would still only be expected to bind to fetal ER if there was no plasma protein binding of EE in the placenta or fetus and no competing endogenous estrogens.

4.5 Limitations of the evidence

An understanding of how NET and EE exert their effects in the body was important for exploring whether the use of Primodos could in theory have affected a pregnancy. However, the data identified were subject to some important limitations. When Primodos was developed, recognition of the importance of pharmacokinetics to understanding drug exposure was in its infancy and pharmacokinetic analysis was not generally used in the development of medicines. The available pharmacokinetic data on NET and EE and their metabolites were accordingly sparse, particularly in humans, and the parameters that would be considered relevant today were not measured in all studies that were identified. Most of the studies also predated current understanding of the diversity and interdependency of hormonal action and, as most were conducted *in vitro* or in animal models, estimations for humans in many cases required extrapolation. The validity of such extrapolations is unknown.

Only very few data were identified in pregnancy, human or otherwise, and still fewer in the first trimester. Similarly, only limited data from animal studies were available on the exposure of the fetus to NET or EE and in these, data was incomplete and hormonal concentrations were measured in units which made comparison with maternal blood concentrations difficult. In addition, several laboratory species with potentially different sensitivities to, and metabolism of, the hormones were used, limited experimental detail was provided in the publications and the immunoassays of the time were relatively insensitive, which made it difficult to judge the reliability of any negative findings.

With respect to the relevance of findings for Primodos tablets specifically, models used various non-oral routes of administration (which may affect circulating plasma concentrations and which metabolites are formed), studied different durations of exposure (which may influence the effects of receptor activation), or studied NET and EE in isolation rather than in combination.

4.6 Discussion of the pharmacological data

For HPTs such as Primodos to have had a direct effect on the developing fetus, the receptors through which NET and EE act would need to be present and functional in the fetus at the time of HPT use. In addition, sufficient NET or EE would need to reach and bind to the receptors and be able to activate them. Progesterone and estrogen hormones, produced by the mother (mainly from the placenta in early pregnancy) and which act through the same receptors, almost certainly also reach the fetus during pregnancy. Little is currently known about the effects of these maternal hormones on the fetus, especially during early pregnancy, but they would be expected to 'compete' with any NET and EE that was present from Primodos and to diminish their effects. Any effects of HPTs would therefore be in comparison to the effects of high concentrations of estrogens and progesterone that the fetus received naturally from the mother compared to the amounts of EE and NET received through taking Primodos, although it is almost impossible to quantify this.

No data were available to address these points directly and so a number of key observations and conclusions were inferred from knowledge of how levels of NET and EE in the blood of women change after taking broadly similar doses of NET and EE to those in Primodos, and from limited studies conducted in animals, including some degree of evidence for the natural hormones. Because there was so much uncertainty it was important that the worst-case scenario was considered.

4.7 Key observations

- Based on animal data, it is likely that a small amount of the NET and EE in Primodos tablets that were taken by a mother reached the fetus during its development in human pregnancy.
- In women, NET and EE are mainly bound to proteins in the blood, such as albumin. This reduces the amounts of NET and EE that are free to bind to and activate hormone receptors in other tissues. It is not yet known if the same levels of protein binding occur in fetal blood, especially in early development.
- Limited data from animal studies suggested that ER and PR appear in both reproductive and non-reproductive tissue in the fetus relatively late in development and around the end of the period of development of the major organs. This period would be equivalent to the end of the first trimester in human pregnancies. How soon the hormone receptors became functional after they were first formed, and how long they were present in tissues other than the reproductive organs, was unclear.
- Although the amount of NET from a single Primodos tablet that was estimated to reach the fetus was potentially high enough to bind to any fetal PR, it was not clear whether it could activate these receptors – this would depend on how much NET was bound to proteins in the fetal blood. It would also depend on how much 'competing' maternal progesterone reached the fetus. Both these values are unknown.
- The small amount of EE from a single Primodos tablet that was estimated to reach the fetus may have been high enough to bind to ER; however, it was only likely to have activated any fetal ER if there was less binding to albumin in the fetal blood than occurred in adults. Any potential effects would also have depended on how much 'competing' estrogen hormones from the mother reached the fetus. Both these values are unknown. Any effect would not be expected to differ from that of natural estrogen.

4.8 Overall conclusion on pharmacological considerations

From the evidence available, the EWG considered that small amounts of norethisterone and ethinylestradiol could have reached the fetus as the result of taking Primodos tablets, for two days during the first trimester of pregnancy but that it was unlikely to have had an effect on the developing fetus, via a direct pharmacological action. Any action of these hormones would require the expression of functional receptors and would undoubtedly be affected by the relatively high concentrations of the very similar natural maternal estrogen and progesterone in early pregnancy.

On the basis of these findings the EWG recommended that:

1. For medicines used commonly in pregnancy, particularly the first trimester, pharmacokinetics and pharmacodynamics studies in pregnant women should be performed, where possible, to understand better how pregnancy affects the levels of drug to which the mother and fetus are exposed and to develop evidence-based dosing and frequency of administration for use in pregnancy.
2. In support of the above opportunities should be provided for obstetricians to receive training in pharmacology.
3. A strategy to co-ordinate and promote research on drug transporter expression in the placenta, particularly in early pregnancy; how it differs between individuals; and how it is affected by maternal disease should be taken forward with appropriate experts in the field.

5 EXPOSURE TO HORMONE PREGNANCY TESTS IN PREGNANCY AND EVIDENCE RELATING TO POSSIBLE ASSOCIATION WITH CONGENITAL ANOMALIES

This chapter summarises the key evidence considered by the EWG on the possible association between the use of HPTs in pregnancy and the adverse outcome of congenital anomaly.

Three key types of evidence were considered by the EWG as follows:

1. Mechanistic evidence for development of congenital anomalies through:
    a. a direct teratogenic effect on the developing fetus of the active ingredients
    b. an indirect effect on the developing fetus caused by temporary disturbance of the developing pregnancy.
2. Reports of suspected adverse effects received from healthcare professionals and patients from a range of sources.
3. Published epidemiological evidence.

5.1 Mechanistic evidence

The evidence for either a direct teratogenic effect of HPTs on the developing fetus or an indirect effect, through perturbation of pregnancy caused by vascular disruption, was examined by the EWG (van Gelder, 2010). Further information is provided in Annexes 20 and 21.

5.1.1 General toxicity

Sex hormones have been extensively studied in repeated dose studies in rodents, dogs and non-human primates, largely to support their use as combined hormonal contraceptives and hormone replacement therapy. The most prominent effects observed were related to exaggerated pharmacodynamic effects, in other words expected but amplified effects in steroidal sex hormone response in primary and secondary reproductive tissues, including suppression of body weight gain, decrease of white blood cells, proliferation of gland tissue, with lactation of mammary gland. Atrophy of sexual organs in males and females has also been observed.

5.1.2 Genotoxicity

Given the timing of HPT administration during early organogenesis the concern was whether NET or EE or both had the potential to cause genetic damage to the somatic cells of the developing embryo, resulting in malformations. This could occur through chromosomal damage or effects on mitosis and cell division. Long-term exposures to clinical doses of natural estradiols and progesterone are associated with carcinogenic effects in humans and experimental animals. However, the mechanisms of tumour induction are related to the hormone receptor-mediated stimulation of growth and differentiation in cells of the target tissues (endometrium, mammary, testis, and prostate), rather than through a genotoxic pathway.

The genetic toxicity of synthetic estrogenic and progestogenic hormones has also been tested in a number of *in vitro* and *in vivo* test systems. These suggest that NETA, EE and other steroidal compounds do not directly interact with DNA but under some circumstances, and at doses more than 1 000 times higher than those used therapeutically in humans, may

produce nonspecific chromosomal damage (clastogenesis). Given the high exposures required to observe genotoxic effects it was considered unlikely that at the levels found in Primodos, NETA and EE would be genotoxic.

5.1.3 Reproductive toxicity

The available non-clinical studies (published literature and unpublished studies from the pharmaceutical industry) on the potential for reproductive toxicity with NETA and EE, separately and in combination, were reviewed to determine whether NETA and/or EE had the potential to cause adverse effects on the male or female reproductive system or on the developing conceptus. The majority of the studies reviewed were conducted in the 1960s and '70s by standards considered acceptable at the time; they were not conducted to current standards, including Good Laboratory Practice. Consequently, the reporting and conduct of the studies are deficient when compared to modern practice. For example, chemical characterisation and stability of the hormones tested in these studies was not controlled. Importantly, in common with the majority of studies from this period, data which confirm the exposure in pregnant animals of the sex hormone and its metabolites were not collected.

5.1.3.1 General considerations

When adverse events in reproductive toxicity studies are observed, the results have particular significance if:

- a relationship exists between the administered dose and the observed response
- effects are observed in more than one species
- a multiplicity of effects is observed

For example, the observation in more than one model or species of the same or related effects would give rise to more concern than a single observation. The following are important considerations when interpreting the findings from non-clinical animal studies:

- Choice of species

Studies assessing the potential for reproductive toxicology are most often conducted in rodents and rabbits but an understanding of the mechanism of any observed effects and how drug exposure in animals relates to humans, is required in order to assess the relevance of any potential risk to humans. Studies of sex hormones in non-human primates are generally considered to offer advantages over rodents and rabbits because of the similarity of their reproductive systems to humans; however, their use poses ethical and practical limitations, particularly as they usually produce single offspring at a time, having a default litter size of one.

Drug-associated teratogenicity is dependent on many factors including the species, since different animal species can be more, or less sensitive to the teratogen. These differences can relate to a number of factors including: differences in the pharmacokinetics of the drug, the stage of prenatal development when the exposure occurs, the dosage of drug administered and the mechanism of teratogenicity. These factors all need to be considered when interpreting the outcome of reproductive toxicity studies in animals.

- Dose

For many teratogens, there is a window of exposure between a dose which results in fetal death and a dose which has no effect. To investigate whether a new drug is a possible teratogen, dose selection is therefore critical and must cover an exposure that would be sufficient to cause a specific defect, if it were teratogenic, without bringing about significant embryo-fetal loss.

- Timing

The time of greatest sensitivity to teratogens is during organogenesis. Table below provides the major developmental milestones for a variety of species during organogenesis, presented in gestational days (starting from the formation of the primitive streak to the closure of the hard palate).

Table 6. Comparative development times during organogenesis in gestational days (adapted from DeSesso 2006).

Embryonic stage	Mouse	Rat	Rabbit	Guinea pig	Macaque	Baboon	Human
Primitive streak	7	8.5	6.5	13.5	16–18	16–18	13–15
First myocardial contractions	8	9.5	8.5	16.5	-–	–	21–24
Anterior limb bud	9.5	11	10–11	16.5	27–29	28	26–30
Hind Limb bud	10–11	11–12	10.5	17.5	28–29	28–30	28–32
Stomach appears	11.5	11.5	10.5	16.5	28–29	–	28–32
Histologic differentiation testis	13.5	12.5	–	–	38–39	–	46–48
Palate fusion	14	16–17	19–20	33	45–47	46–48	56–58
Full term	20–21	21–22	31–32	65–68	164–168	–	~266

- Sample size

Animal reproductive toxicity studies usually lack the statistical power to detect subtle increases in rare events. Currently, for all but the rarest events, evaluation of 16 to 20 litters for rodents and rabbits is considered necessary to provide a degree of consistency between studies. For this reason, reproductive toxicity studies include high doses of test substance so as to induce some maternal toxicity and maximise the possibility of detecting a response. Studies conducted many years ago, including those on NETA and EE were mostly based on the evaluation of fewer than 16 litters per group and so lack the power of studies conducted today.

One of the most challenging aspects in interpreting developmental toxicity studies is evaluating the relevance of sporadic findings and determining if these occurred spontaneously or if they are related to the drug exposure. Like humans, laboratory animals have a spontaneous malformation rate for which no cause is currently known. These may be indistinguishable from teratogen-induced defects when they occur with low frequency. The chance of spontaneous rare effects occurring in a study with one control group and three drug-treated groups is 3:1 in favour of the finding occurring in the drug-treated groups. As a result, comparison of any effects in the drug-treated groups with the laboratory's historical control data is important for determining whether a small increase or decrease (not statistically significant) in an endpoint constitutes a treatment-related effect. In many of the studies reviewed here, historical control data were either not available or not discussed in the study report or publication. Evaluating the significance of developmental toxicity findings also requires a comparison of results from other reproductive and developmental studies as well as studies in other species.

– Extrapolating risks to humans

If an adverse effect is recognised in animals, a key aspect of evaluating the potential risk to human pregnancy is determining a 'safety margin'. This is done by comparing the highest systemic concentration of test substance at which no observed developmental effects occurred in the animals (referred to as the 'No Observed [Adverse] Effect Level' or NO[A]EL) with the maximum systemic concentration attained in humans when using the highest dose permitted for that substance. For example, a medicine that reaches a maximum systemic concentration of 5 µg/ml in humans and which causes malformations in animals at 1000 µg/ml but not at 500 µg/ml has a NO(A)EL of 500 µg/ml and a safety margin of 100.

In the absence of data on drug levels in the blood in animal studies, dose equivalence must be estimated. A dose comparison based on an assumption that doses scale 1:1 between species when normalised to body surface area (mg/m^2) is currently considered to provide the best exposure estimate[15]. In many older studies, dose equivalence of the administered substance in mg/kg body weight to that used in humans was used.

Key considerations when assessing the relevance to humans of any observations in animal studies were therefore:

- Did increasing the dose of NET or EE increase the severity or incidence of the effect?
- Were toxicokinetic (drug blood levels in toxicological studies) comparisons made between experimental animal studies and estimated human exposure?
- Were there significant differences in the pattern, timing or magnitude of blood levels of drug between the experimental studies and humans?
- Was a NO(A)EL (or NOEL) demonstrated?
- Were similar effects observed in one or more species?
- Was the mechanism of action known or deducible?

5.1.3.2 Data considered

Non-clinical studies of the reproductive toxicity of sex hormones were identified from the published literature or from unpublished studies conducted by Schering AG. NETA and EE, used singly or in combination, were studied in mice, rats, guinea pigs, rabbits and non-human primates (Table). Eight studies with related estrogens, seven studies with related progestogens and 14 studies with related estrogen + progestogen combinations were also evaluated (Annex 20).

In addition, preliminary findings were presented from studies, looking into the effects of NETA and EE on embryonic development (including blood vessel formation, nerve outgrowth, and eye and ear development) using the zebrafish and chick embryo models, to the EWG on the 18th October 2016 (Dr Neil Vargesson). This work is currently unpublished. However, no developmental effects of NETA and EE on chick embryos had been found at two developmental timepoints tested, even at very high doses, but lethality had been found. Dose-dependent damage in zebrafish embryos which were viable had been found (eg. small eyes and ears; bent spines; yolk sac damage; loss of movement). Lethality at high concentrations had also been observed. Using in-vitro tissue culture assays (using human and mouse cell lines) direct effects of NETA and EE upon nerve outgrowth and blood vessel formation were seen. Further work on the effects of the NETA and EE on the zebrafish embryo model has been submitted for publication. This additional work had confirmed and

[15] FDA Guidance for Industry Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers (https://www.fda.gov/downloads/drugs/guidances/ucm078932.pdf)

extended the results presented for dose dependent developmental effects in the zebrafish embryo; that the NETA/EE mixture acts on the embryo in a time sensitive manner; that the drug accumulated in the embryo and that damage was observed rapidly.

This work is ongoing and is to be continued via studies in rodent species to establish if the effects reported in the zebrafish embryos with NETA and EE could also occur in mammalian embryos.

Table 7. Non-clinical studies of the reproductive toxicity of NETA and EE from the published literature or from unpublished studies conducted by Schering AG.

Data source	Number of studies
EE	
Published literature (from academia or the pharmaceutical industry)	
Mice	3
Rats	6
Rabbits	2
Non-human primates	1
Industry (unpublished)	
Rats	4
Rabbits	4
TOTAL	**20**
NETA	
Published literature (from academia or the pharmaceutical industry)	
Mice	3
Rats	8
Guinea pigs	1
Rabbits	1
Non-human primates	2
Industry (unpublished)	
Mice	2
Rats	7
Rabbits	1
TOTAL	**25**
NETA + EE	
Published literature	
Mice	2
Rats	2
Rabbits	1
Non-human primates	6[1]
Industry (unpublished)	
Mice	4
Rats	9
Rabbits	11
Guinea pigs	1
Non-human primates	1
TOTAL	**37**

[1] One publication (Hendrickx 1987) included three different studies, one of which was also available as an industry unpublished report.

5.1.3.3 Studies with norethisterone acetate alone

Twenty-five studies in mice, rats, rabbits, guinea pigs and non-human primates provided information on the effects of NETA exposure during organogenesis. A wide range of NETA doses were given for different durations, on different days of gestation and through different routes. Key aspects of the studies and the observed adverse effects are provided in Tables 6 and 7 of Annex 20.

The majority of studies with NETA (and related progestogens) showed genital malformations of the fetus with doses generally higher than those used in HPTs. This finding was observed in mice, rats, guinea pigs, rabbits and non-human primates. The induction of genital abnormalities is related to the known androgenic activity of NET and structurally related hormones. The review therefore focused on anomalies of non-reproductive tissues, for which there has been scientific uncertainty over evidence for a causal association.

In addition to the known effects of NET on genital tissue, the following adverse effects were noted:

- increase in embryo-fetal loss when given at high doses during pregnancy in rodents and rabbits (discussed in Chapter 6)
- increase in skeletal variations in rodents and rabbits
- equivocal increase in malformations in one rabbit study (Schering #2300, 1976 – two fetuses with umbilical hernia) at doses higher than those used in HPTs.

From studies that evaluated dosing with NETA throughout the period of major organogenesis, no good evidence for non-genital teratogenicity was reported in any species.

5.1.3.4 Studies with ethinylestradiol alone

Twenty-one studies in mice, rats, rabbits and non-human primates provided information on EE exposure during organogenesis. EE was administered over a wide range of doses and treatment durations, during different periods of gestation and by different routes. Key aspects of each study and details of the adverse effects found is provided in Tables 4 and 5 of Annex 20.

Adverse effects noted were as follows:

- increase in embryo-fetal loss when given at high doses during pregnancy in rodents and rabbits (discussed in Chapter 6)
- increase in skeletal variations (wavy ribs, retarded ossification) when given at higher doses (as calculated on a body surface area mg/m^2estimate) than those used in Primodos
- equivocal increase in malformations in one study in the rat (Schering study #983, 1973) which was not repeated in a larger study (Schering study #4136, 1980)
- effects on reproductive tissue (genital malformations) in rats in two studies, with continuous administration of EE at relatively low doses throughout pregnancy and lactation
- two literature reports of changes in maternal behaviour and emotional alterations in rats exposed throughout organogenesis.

Overall, these studies do not support an increase in malformations in non-reproductive tissue following EE exposure.

5.1.3.5 Studies with norethisterone acetate and ethinylestradiol combined

Thirty-seven studies in mice, rats, rabbits, guinea pigs and non-human primates provided information on effects of NETA and EE exposure during organogenesis. A wide range of doses were administered, for different durations, on different days of gestation, and through different routes.

Consistent findings were shown in studies with NETA and EE (or structurally related synthetic progestogens) across mice, rats, guinea pigs and rabbits. In accordance with their

hormonal actions, genital organ malformations were observed at high doses in some of these studies.

Most studies found very little evidence for an increased incidence of malformations outside of hormone-responsive reproductive tissue. Key aspects of the studies and the observed adverse effects, including malformations, are provided in Tables 17 and 18 of Annex 20; details of the more robustly designed studies are in Table to Table 11 below.

- Mice

In Schering study (#3579; 1978, and #3861; 1979) mice were exposed to NETA and EE (up to 50 + 0.1 mg/kg/d) throughout organogenesis. The high dose is about 30 times higher than the daily dose of NETA and EE in Primodos (calculated on a body surface area mg/m^2 basis). In this study, there was significant embryo-loss at the mid and high doses. At these doses, there was also an increase in fetuses with marked dilation of the renal pelvis, and at the top dose aI increase in slight retardation of skeletal ossification. A single incidence of cleft palate in the low and mid dose groups and a single incidence of exencephalocele in the mid and high dose groups was reported.

Additional thoracic sectioning followed by microscopic examination of the embryos reported a significant increase in visceral malformations, including the heart, lung, and thorax wall at the high dose (table 8).

Table 8. Findings from Schering study #3579/3861 in mice dosed with NETA and EE: thoracic anomalies.

	Frequency of thoracic anomalies			
Type of thoracic anomalies	Group 1 Control	Group 3 5.0 + 0.01 mg/kg/d	Group 4 15.0 + 0.03 mg/kg/d	Group 5 50.0 + 0.1 mg/kg/d
No. of fetuses examined	75	96	56	33
Transposition of the azygos vein	1	0	0	1
Reduction in the size of the azygos vein	0	1	0	0
Imperfect closure of the ventricular wall with intraventricular septal defect	0	0	0	1
Incomplete development of the left thoracic wall with herniation of the pleura	0	0	0	1
Absence of postcaval lobe of the lung	0	0	0	2
Rotation of the heart to the left associated with apparent reduced development of the left lung	0	0	0	1

A number of these changes are known to occur spontaneously and mice can be particularly prone to malformation clusters. Comparing 75 control fetuses with 185 test fetuses and having no historical control data with which to compare the findings, made it difficult to put the findings into context. Lower doses did not produce an increase in malformations in earlier experiments with mice. Nevertheless, the increase in malformations in this mouse study should be considered drug-related.

- Rats

In an equivalent study to the mouse study described above, rats were exposed to NETA (50 mg/kg/d) and EE (up to 0.1 mg/kg/d) during gestation days 6 to 15 (Schering study #3578, 1978 and study #4284, 1980, Table). At the high dose, there was a very high resorption

rate, with only two live fetuses recorded. In this study, no external, skeletal or visceral malformations were observed. Microscopic analysis of the thorax was conducted and, with the exception of subcutaneous oedema, there was no increase in thoracic anomalies (table 9).

Table 9. Thoracic anomalies in Schering rat study #3578[11]/4284[12] (NETA and EE).

Group	Group 1 control	Group 3 5.0 + 0.01 mg/kg/d	Group 4 15.0 + 0.03 mg/kg/d	Group 5 50.0 + 0.1 mg/kg/d
No. of fetuses examined	111	93	90	3
No. of fetuses with: subcutaneous oedema	1	0	3	2
Folding and thickening of dorsal skin	0	0	1	0
Late embryonic death: autolysed fetus possibly with a high interventricular defect	0	1	0	0
Distortion of the lobes of the lung	1	0	1	0
Incomplete development of the trachea and bronchus Additionally, a defect in the development of the atrium	2 1	0	0	0
Intramural haemorrhage in the aorta	0	0	1	
Transposition of azygous vein	0	0	1	0

In some studies, with higher doses of NETA and EE (Schering study #3578 and #4284 and Joshi, 1983) there was also evidence of an increase in developmental variations, such as delayed or non-ossification of the fetal skeleton and wavy ribs. Developmental variations also occurred in the control animals. An increase in skeletal variations was also observed in two Schering studies (study #2221 and #1631) in which rats were given the natural forms of estrogen and progesterone. Developmental variations such as these are not typically considered to be congenital malformations because these abnormalities are often seen in the presence of maternal or embryo-fetal toxicity and post-natally improvement is generally seen.

Standard embryo-fetal development studies by necessity are generally designed to dose a pregnant animal throughout the period of major organogenesis. The production of malformations, however, often requires a specific set of conditions, and dosing a pregnant animal on a single day or on two days of gestation can be an efficient means of eliciting malformations. In a series of studies conducted by Schering in rats and rabbits the hypothesis was tested that after treatment with a combination of NETA and EE spanning different periods of organogenesis, teratogenic effects might occur, which otherwise might have been masked by a more markedly embryo-lethal effect observed after treatment throughout organogenesis.

The doses used in the rat were relatively high (≥40-fold) compared to the Primodos dose in humans based on a body surface mg/m^2 comparison. The number and pattern of malformations observed provided no evidence of teratogenicity in the rat (Table 10). An increase in skeletal variations and an increase in small fetuses (runts) in the study group treated on days 14 to 15 of gestation was reported and considered an indication of embryo-fetal toxicity (see Annex 20, table 18). The variations reported were mostly consistent with those found in control animals.

Table 10. Distribution of external, skeletal and visceral malformations in rats dosed with NETA and EE; studies #4037 [1979]; #4042 [1979]; 4044 [1979]; #4045 [1979]; #4046 [1979]) (number of malformations/number of fetuses examined).

Gestation Day	Vehicle control	Low 50 + 0.01 mg/kg	Mid 150 + 0.03 mg/kg	High 500 + 1.0 mg/kg	Very high 1500 + 3.0 mg/kg
6–7	0/172	0/152	0/128	0/22	-–
8–9	1/225 (0.4%)[a]	1/177 (0.6%)[a]	0/109	0/14	-–
10–11	0/247	0/191	0/157	0/62	–
12–13	0/264	1/206 (0.5%)[b]	0/203	1/228 (0.4%)[c]	–
14–15	0/284	–	0/250	0/151	0/230

[a] Umbilical hernia (external)
[b] Malformation of the tail and atresia ani (external)
[c] Malformation of the tail (external)

- Rabbits

A parallel set of experiments to those conducted in the rat evaluated the effect of dosing on two consecutive days during organogenesis. The rabbit is more sensitive to the reproductive effects of the hormonal combination; the doses evaluated in this study were significantly lower than those used in the rat study with the low dose in same range as used in Primodos when calculated on a body surface mg/m^2 basis.

A single malformation occurred in each of two studies in rabbits dosed with NETA and EE from gestation days 7 to 16 throughout organogenesis. Single malformations were also observed in the control rabbits in each of two studies and in two controls in another study (Table 11).

In all except the group dosed on days 16 to 17 of gestation there was a small but dose-dependent increase in the number of skeletal variations, which can be an indication of embryo-fetal toxicity. The variations reported were mostly consistent with those found in control animals and were mainly associated with presence of the 13th rib, a common variation in this species of rabbit, and lack of ossification of one or more sternebrae.

The number and pattern of malformations observed in rabbits in these studies provided no evidence of teratogenicity.

- Non-human primates

Studies in non-human primates offer advantages over rodents and rabbits with regard to developmental and reproductive studies of sex steroids because of the similarity of reproduction to humans, eg. endocrinology of ovarian function and early pregnancy, placental morphology and physiology, timing of implantation and rates of embryonic development, and similar responses to known human teratogens. A number of embryo-fetal developmental studies have been conducted with combinations of NETA and EE during pregnancy in non-human primates.

The genital abnormalities observed with NETA and EE in rodent and rabbit studies were also seen in studies conducted in non-human primates, with virilisation of female cynomolgus monkey fetuses at high doses of NETA and EE (300 – 1 000 times the Primodos dose based on mg/kg/day).

In three studies, NETA and EE were given orally at doses ranging from 1 to 1 000 times the Primodos dose (on a mg/kg/day basis) to rhesus monkeys, cynomolgus monkeys or baboons over a 30-day period, from gestation day 20 to 50 (the period of organogenesis).

Table 11. Distribution of external skeletal and visceral malformations in rabbit studies dosed with NETA and EE; studies #3581 [1978]; #4036 [1978]; #4038 [1978]; #4039 [1978]; #4040 [1978]; #4041 [1978]; #4043 [1978] (number of malformations/number of fetuses examined).

Gestation Day	Vehicle control	Very low 0.15 + 0.0003 mg/kg	Low 0.50 + 0.001 mg/kg	Intermediate 1.5 + 0.003 mg/kg	Medium 5.0 0.01 mg/kg	High 15 + 0.03 mg/kg
6–7	0/103	–	0/90	–	0/65	0/13
8–9	0/64	–	0/52	–	0/54	0/1
10–11	0/67 1/34[a] (2.9%)	–	0/48	–	0/74	0/2
12–13	1/71[b] (1.4%)	1/89[c] (1.1%)	1/72[d] (1.4%)	–	0/30	–
14–15	0/64	0/74	0/79	0/59	–	–
16–17	1/67[b] (1.9%) 1/35[e] (2.9%)	0/55	0/56	0/54	–	–

[a] Fusion of the 7th and 8th rib on the left side;
[b] Umbilical hernia;
[c] Exencephaly;
[d] Severe malformation;
[e] Anophthalmia on the right side

Virilisation of female cynomolgus fetuses following NETA and EE treatment was manifested as cases of clitoral enlargement in the group dosed with 300 times the Primodos dose and cases of increased anogenital distance with reduced vaginal opening in the group dosed with 1 000 times the Primodos dose. A single case of scoliosis, rib and vertebral malformation was detected in a treated group across the studies which was described as within the normal background rate. No increase in non-genital malformations was detected across these studies (Hendrickx, 1987). The blood concentrations of NETA and EE were not determined in these studies and so it was not possible to compare the exposures directly to those in humans.

5.1.3.6 Key observations from reproductive and developmental toxicity studies

- *Virilisation*

Genital tract abnormalities / malformations including virilisation of female fetuses were reported in rodents and non-human primates exposed to these hormones during the period of sexual differentiation. The effects on male and female reproductive tissue reflect the known hormonal action of these compounds.

- *Genetic toxicity*

Genetic toxicity studies of NETA and EE, alone or in combination, indicated that these hormonal agents do not directly interact with DNA, as evidenced by the negative mutagenicity results, but under some circumstances they could produce nonspecific chromosomal damage.

- *Skeletal variations*

In rodents and rabbits there was evidence that NETA and EE in combination increased the frequency of skeletal variations. Such effects were not considered to be mechanistically linked to malformation because these abnormalities are often seen in the presence of maternal or embryo-fetal toxicity and post-natally improvement is generally seen.

- *Congenital malformations*

In a study in mice there was evidence that a combination of NETA and EE given throughout organogenesis at doses higher than those used in HPTs was associated with an increase in thoracic malformations. Since similar investigations in rats did not show a similar finding and there was no evidence from rabbits and non-human primates for an increase in thoracic malformations, the observed effects were considered specific to mice.

5.1.3.7 Overall conclusion on animal data

The totality of the available data from studies in rats, rabbits, and non-human primates did not support a causal association between administering norethisterone and ethinylestradiol at the doses and durations found in Primodos and the development of malformations in non-reproductive tissues of the offspring.

5.1.4 Vascular disruption with norethisterone and ethinylestradiol

Vaginal bleeding has been widely used as a surrogate for disruption of the endometrial lining in non-pregnant women; in pregnancy, it is viewed as a symptom of disruption of the placenta. In the non-pregnant uterus, the fall in NET and EE levels after taking the second Primodos tablet resulted in shedding of the endometrium and a withdrawal bleed. The possibility that the progestogen withdrawal following Primodos in the non-pregnant uterus could likewise have led to some degree of vascular disruption in the pregnant uterus, even if partial or temporary, was therefore explored by the EWG as a possible mechanism for the development of congenital anomalies.

During pregnancy, the placenta forms as an intermediate organ between the growing embryo and the maternal uterus. It develops in parallel to the embryo during early pregnancy and, after the first trimester, is how maternal blood supplies the growing fetus with oxygen and nutrients. During most of the first trimester the developing embryo is maintained by the yolk sac via the vitelline circulation (blood flow between the yolk sac and the embryo). As fetal development continues, embryo-fetal vasculature develops to link the blood supply from the placenta via the umbilical circulation to the fetus and the vitelline circulation regresses. The embryo-fetal vascular system thus develops in several phases to accommodate the changes in blood requirement as the major organs and body systems develop.

Since new blood vessel formation occurs in parallel with organ development, any premature or delayed replacement of embryonic blood vessels with fetal blood vessels can result in inappropriate blood supply to the tissue or organ and may result in abnormal growth of the target organ and hence congenital anomaly.

Disruption or interruption of intrauterine blood supply can occur due to i) effects on the mother's blood supply between the uterus and placenta, ii) disturbances of the supply between the placenta and the embryo or iii) effects on the embryo-fetal circulation. Further information is provided in Annex 21.

5.1.4.1 Causes of vascular disruption

Vascular disruption results from interference with or extrinsic breakdown of prenatal development of the vasculature that was previously normal. It has been proposed to arise from a number of different underlying causes that fall broadly into two categories: i) disturbances to blood flow within 'normal' vessels and ii) disturbances to blood vessel development. Possible mechanisms for vascular disruption include:

- mechanical pressure on the umbilical cord that slows the flow of blood

- physical trauma to the fetus by the umbilical cord wrapping around a limb and reducing its blood supply
- physical trauma to the fetus through a failed abortion or chorionic villus sampling (Firth, 1997)
- formation of a clot within the umbilical cord due to excess twisting (hypercoiling) or (external) trauma
- dispersal of thrombi (blood clots) through the umbilical cord, or between co-twins, and blockage of the fetal blood vessels
- maternal medical conditions such as diabetes or hypertension
- pharmacological effects (eg. cocaine use later in pregnancy, alcohol)
- premature rupture of membranes
- maternal to fetal transfusions of immunoglobulins or vasoactive substances from a damaged placenta causing vascular damage and obstruction in the fetal circulation
- mechanical damage allowing fetal to maternal transfusion, resulting in fetal hypovolaemia followed by hypo-perfusion.

Anomalies caused by vascular disruption depend on the duration, extent and timing of the disruption, the site of disruption and which vessels are affected.

5.1.4.2 HPT-induced vascular disruption

Limited data were identified on the potential for the components of Primodos to cause vascular disruption and NET has not been cited in the literature as a cause of vascular disruption. However, based on the pharmacological profile of NET and EE, a number of possible options for a disruptive effect were explored:

a) Disruption of the endometrial lining or utero-placental structure due to reduction of progesterone effects:

During early pregnancy progesterone release from the corpus luteum maintains the endometrium to allow implantation to occur and allows the uterus to remain in a 'relaxed' (non-contractile) state and the placenta to develop. The placenta itself takes over progesterone production around the end of the first trimester for the remainder of the pregnancy. Studies in animals, studies with human cells *in vitro*, and studies in pregnant women prior to an elective termination of pregnancy have shown:

- Progesterone 'withdrawal' by removal of the corpus luteum from pregnant women in one study increased uterine tone and responsiveness to oxytocin
- In early pregnancy NET administered to 24 women in various regimens at total doses of up to 1 000 mg orally per day (Nygren, 1975), induced a transient but significant decrease in plasma progesterone levels, although comparative data for untreated women were not available. The reductions in plasma progesterone levels were not associated with vaginal bleeding in any of the women. Histological examination of fetal tissue following elective termination of pregnancy did not differ between the treated and untreated groups.
- NET has mainly progestogenic action, but high doses have been shown to have anti-progestogenic action (Markiewicz, 1994).
- A study by Schering in 1963 in pregnant rats reported a dose-dependent effect of NET on vaginal bleeding, with no effect observed with a single dose of 3 mg NET (about 70 times higher than the equivalent dose of Primodos in humans), uterine

bleeding observed in half the cases treated with 10 mg NET (depending on time of dosing) and bleeding observed in all cases treated with 30 mg NET.

- Two studies in which 25 women were given NETA (40 mg) and EE (0.04 mg) or 11 women were given NETA (20 mg) and EE (0.04 mg) at weeks six to seven of gestation, found no effect on endogenous progesterone or estradiol levels, and did not identify any macroscopic or microscopic differences indicative of placental damage, (eg. thrombosis or in degenerative changes and leucocytosis in the termination products) compared with similar numbers of untreated women (Pulkkinen, 1984).
- No studies of, or citations for, a possible effect of NET and/or EE on uterine tone or contractions, or data suggesting an anti-progestogenic effect of NETA in pregnancy, when given in combination with EE, were identified.

b) Acute reductions in maternal blood flow through the uterine artery / placenta through formation of a blood clot

Estrogen-progestogen combinations increase the risk of thromboembolic events. If a blood clot were to form during early pregnancy in the general circulation it could in theory travel to the uterine artery, causing local reductions in blood flow to the placenta. However, perfusion of the intervillous space of the placenta by maternal blood is not thought to occur until the end of the first trimester and no studies of, or citations for, a possible effect of EE and/or NET on uterine arterial blood flow during pregnancy were identified.

c) Formation of a blood clot in the fetal circulation

EE and NET can pass through the placental barrier and so it was questioned whether this could increase the thrombotic risk in the fetus and provoke occlusion of embryo-fetal blood vessels. Coagulation proteins have been detected in the fetus from about 5 weeks of gestation but these are thought to act as regulators of tissue proliferation and differentiation rather than as clotting factors during embryogenesis (Manco-Johnson, 2005).

5.1.4.3 Discussion

Overall, very little data on the potential effects of NET and EE on the vasculature in pregnancy were identified. The underlying clinical and/or experimental evidence in support of vascular disruption as a cause of congenital anomalies in general is drawn largely from studies conducted following different 'insults' during pregnancy. Definitions for vascular disruption defects differed between studies whilst those proposed to be drug related often included different exposure times, durations, and/or doses.

The data were largely generated before the 1990s and did not necessarily reflect more recent understanding of pregnancy and/or embryonic development. Thus, alternative explanations now exist for the aetiology of some anomalies that were previously attributed to vascular disruption. Similarly, it is now considered that the onset of placental perfusion by maternal blood occurs towards the end of the first trimester of pregnancy and that the embryo develops in a relatively hypoxic environment until this time (Huppertz, 2007; Webster, 2007). It was therefore not clear that reduced uterine blood flow or hypoxia could induce vascular disruption following an insult during the first two months of pregnancy, although other mechanisms such as oxidative stress may play a role.

Against a background of very similar maternal hormones circulating at much higher levels during pregnancy, it was difficult to understand how use of two tablets of Primodos taken 12 to 24 hours apart could block a physiological effect. NET is an anti-inflammatory steroid that

interacts with the glucocorticoid receptor at certain doses and would be expected to have anti-disruptive rather than pro-disruptive effects during pregnancy.

5.1.5 Overall conclusion on vascular disruption

No evidence that norethisterone and/or ethinylestradiol could disturb a pregnancy through vascular disruption was identified.

5.2 Reports of suspected adverse drug reactions

Further information is provided in Annexes 22-24.

5.2.1 Information obtained

Reports of suspected ADRs to HPTs were collated from several sources, as follows:

a) MHRA database of spontaneous reports (Yellow Card reports)

MHRA's spontaneous reporting database was interrogated to identify cases of suspected ADRs that had been reported in association with the active ingredients in HPTs available in the UK, including NETA and EE. These hormones are also ingredients of other products including oral contraceptive pills so to identify reports associated with HPTs only, the following exclusion criteria were applied to the search:

- brand names for active substances which were different from the hormone pregnancy test brands;
- drug was started or stopped after 1978 (when HPTs were no longer available in the UK);
- duration of treatment longer than four 4 days (HPTs were taken for two to four days);
- the indication given for the suspect drug of 'contraception'.

Norethisterone as a single active ingredient was not included in the search, because initial searches showed it was difficult to distinguish reports for the norethisterone-only HPT (Norlutin®) from other products which contain norethisterone as a single active substance. The initial data-lock point for analyses of the MHRA's Yellow Card database was 9th November 2015; a subsequent search of the database with a datalock point of 31st January 2017 identified one additional case. Suspected ADR reports were manually reviewed to identify any reports of miscarriages, fetal death, congenital anomalies or any other pregnancy complications and to exclude any additional reports not relating to HPTs.

b) Information submitted through the public call for evidence

The following relevant information was received in response to the MHRA's public call for evidence:

- Yellow Card reports of suspected anomalies and adverse effects on pregnancy in association with HPTs
- Information on anomalies and adverse effects in pregnancy with use of HPTs in the UK and Germany (provided by the 'Association for Children Damaged by Hormone Pregnancy Tests' and by a Duogynon campaigner in Germany)
- Follow-up detailed information of cases from the 'Association for Children Damaged by Hormone Pregnancy Tests'

- Testimonials from people describing their experiences with HPTs, submitted either directly to MHRA or through Members of Parliament
- Details of all cases spontaneously reported to the UK Teratology Information Service (UKTIS).

c) Spontaneous reporting data from the World Health Organisation

The WHO was asked to provide details of all relevant worldwide spontaneous reports of congenital abnormalities following exposure to HPTs in its Vigibase repository. Anonymised reports were filtered by WHO for terms indicative of congenital anomalies using four Standard MedDRA Queries.

d) Regulatory authorities outside the UK

In December 2014 MHRA asked other Member States of the EU to provide details of any current or ongoing reviews of interest in HPTs; no spontaneous reporting data were provided. Information on regulatory action that had been taken in countries outside the EU, including in the US, Japan, India, Australia and New Zealand was also requested.

e) Pharmaceutical companies

All pharmaceutical companies whose predecessors marketed HPTs were asked to submit all relevant information or documentation held. In order to preserve the legal privilege for the cases notified to Schering during the course of the litigation process, Bayer provided anonymised case report data.

Table 12 describes the data that were initially obtained from the different sources and Annex 22 provides a full descriptive assessment of all cases received by MHRA.

The EWG also heard from 13 members of the 'Association for Children Damaged by Hormone Pregnancy Tests' who had, or whose child had had, one of a range of different anomalies. The members confirmed that the HPT had been taken within the critical period for fetal development, and that in many cases a test was recommended by the doctor rather than requested, that pills were given to first time mothers who did not consider themselves to be in any high-risk category, and that the doctor in several cases had taken what appeared to be free samples from a desk drawer, rather than making out a prescription.

5.2.2 Handling of duplicate reports

Spontaneous ADR reports received through the YCS were manually screened for duplicates (two or more reports concerning the same patient experiencing the same ADR at the same time), based on an algorithm that compared the similarity of certain fields in the database (including patient identifiers and suspect drugs) and flagged any reports that were potential duplicates.

The main area for potential overlap of cases was considered to be the Bayer litigant data and the cases provided by 'The Association of Children Damaged by HPTs'. However, it was not possible to identify duplicates on the basis of the limited information provided and further details of cases could not be provided. Cases provided by 'The Association of Children Damaged by HPTs' were more detailed than those provided by Bayer and so all cases from Bayer were excluded from further analyses to avoid introducing unnecessary duplication of information.

Table 12. Number of reports of suspected congenital anomaly reports and other adverse drug reactions in pregnancy with HPTs.

Source	Number ADR reports[1]	Comment
MHRA database of spontaneous reports (Yellow Card reports)	28[2]	Ethisterone and EE (Amenerone, Amenerone Forte): 17 reports NETA and EE (Primodos): 163 reports initially identified, of which 93 were excluded as being unlikely to relate to HPT (no reports of congenital anomalies, miscarriages, fetal deaths or pregnancy complications were amongst those excluded). In 59 of the remaining 70 reports it was not possible to confirm nor exclude with certainty whether the ADR related to an HPT; none included an outcome of congenital anomaly so were excluded from further evaluation. The 11 remaining reports specified or suggested that the product was an HPT; four of these were submitted between 2007 and 2014 and seven were reported during the public call for information. Few details were provided in many cases.
Patient testimonials	7[3,4]	Describe a mix of outcomes (including 1 report with no ADR). Key information lacking in some cases, notably on maternal medical history and possible confounding factors.
German campaigner	302	Information unclear but a total of 332 patient reports describing 302 adverse effects. Details of adverse events in German.
'Association for Children Damaged by HPTs'	133[4]	92 women took Primodos, 33 did not provide the name of the HPT and 8 took other HPTs; 30 of the 134 offspring were reported to have died (age of death from miscarriage up to 47 years of age); there were six reports describing siblings who were both exposed to HPTs *in utero* and several reports of developmental delay
Bayer	94[5]	Anonymised summary of cases grouped by Bayer into: i) structural defects eg. limb defects, cleft palate ii) functional/developmental defects eg. intellectual disability, behavioural disorders iii) unclassified defects eg. sensory problems, blindness, cerebral palsy and iv) unknown defects eg. no information, unspecified anomalies. 20 claimants had multiple defects involving several categories.
UKTIS	5	Information from five enquiries on use of Primodos in pregnancy. The similarity of the details provided in three of these cases suggested they represented enquiries from different health care providers about the same patient. As no personal identifiers were provided at the time of enquiry it was not possible to confirm this.
UK Royal Colleges/ professional bodies	0	No information available.
Regulatory authorities outside the UK	7[6]	Line listing of six cases reported 1973 to 1978 in Australia; details of one case reported to Centre for Adverse Reactions in New Zealand.
WHO ADR database (Vigibase)	205[7]	Case-level narrative information not provided. Six cases with ethisterone and/or EE; 205 cases with NETA and/or EE. Indication not provided in 152 cases.

[1] Reports for which an HPT was confirmed as suspect drug
[2] Two cases were reported to MHRA in 2016 and so were not included in the initial case review
[3] Eight in total but one was from the German campaigner and is included in 'Information from the German campaigner'
[4] One case did not provide an adverse event
[5] Bayer states they have information for a total of 196 claimants, but that no information on the outcome of pregnancy was provided for 102 of these claimants
[6] Six from the Therapeutic Goods Administration in Australia (one excluded as not adverse ADR of pregnancy) and one from the New Zealand Pharmacovigilance Centre
[7] 30 cases originated from the UK.

Other cases excluded were reports that did not record an adverse reaction; where an HPT could not be confirmed as the suspect drug (and the reaction was not an adverse outcome of pregnancy); that were reported in a foreign language; or that did not report an adverse outcome of pregnancy. The final dataset comprised 235 reports as follows:

- MHRA database of UK spontaneous reports (Yellow Card reports) (n=30[16]);
- Patient testimonials (n=7);
- Association for Children Damaged by HPTs (n=133);
- UK Teratology Information Service (n=5);
- World Health Organisation (n=53);
- New Zealand regulatory authority (n=1);
- Australian regulatory authority (n=6)

5.2.3 Re-categorisation of cases

To enable direct comparison of cases, the information in reports was re-categorised from their verbatim terms according to the WHO ICD10 medical classification system (10th revision of the International Statistical Classification of Diseases and Related Health Problems) and according to the EUROCAT method (EUROCAT Guide 1.4, version 20.12.16). ICD10 is a hierarchical system with broad chapters for each system organ class that are then sub-grouped into blocks and categories of related conditions down to specific event terms. ICD10 codes were then mapped to the rest of the ICD10 hierarchy. The EUROCAT guide provides a defined approach for the coding and classification of suspected congenital anomalies, and a standardised method for analysing and displaying the data.

BioMedical Computing Ltd re-categorised the anonymised HPT-exposed ADR data, based on a Microsoft Excel database provided by MHRA. EUROCAT categorisation displays only 'major anomalies' or 'major + minor' anomalies; minor anomalies alone are not captured unless associated with a major anomaly within the same case report. 173 cases were categorised as having a major congenital anomaly according to the EUROCAT method of which five reports were identified as reliably having a genetic cause:

- Cornelia de Lange syndrome
- DiGeorge syndrome
- Chromosomal abnormality unspecified (x2)
- Angelman syndrome.

5.2.3.1 Relationship with dose, timing of exposure, and gender

- Dose

An effect of the dose of a drug on the frequency of reporting of an adverse event (dose-relationship) can often be a useful parameter for assessing causality. In the case of HPT products, very few had more than one dose strength and few reports provided any information on dose.

- Timing of exposure

Timing of HPT use during pregnancy was not stated or was not known in two-thirds of the reports but would be expected to have been used most frequently from about six to seven

[16] Includes two additional cases reported in 2016

weeks of gestation. For the majority of reports where timing was reported, the type of anomaly reflected the critical period of exposure (Table 13).

Table 13. Reports of anomaly by timing of exposure.

	ICD10 Block	Critical period (weeks)[1]	Total number reports	Exposure during critical period[2]		
				Yes	No	Unknown
Q65-Q79	Congenital malformations and deformations of the musculoskeletal system	5 to 8.5	82	22	8	50
Q20-Q28	Congenital malformations of the circulatory system	4 to 6	39	10	6	23
F80-F89	Disorders of psychological development	Up to 16	27	16	0	11
Q00-Q07	Congenital malformations of the nervous system	2.5 to 6	26	10	5	11
Q35-Q37	Cleft lip and cleft palate	7 to 9.5	19	3	2	14
Q38-Q45	Other congenital malformations of the digestive system	5 to 10	18	5	1	12
G90-G99	Other disorders of the nervous system	2.5 to 6	16	3	5	8
G40-G47	Episodic and paroxysmal disorders	2.5 to 6	15	8	3	4
G80-G83	Cerebral palsy and other paralytic syndromes	2.5 to 6	14	7	2	5
R25-R29	Symptoms and signs involving the nervous and musculoskeletal systems	2.5 to 8.5	13	6	2	5
Q50-Q56	Congenital malformations of genital organs	7.5 to 10.5	12	0	4	8
Q10-Q18	Congenital malformations of eye, ear, face and neck	3.5 to 9	12	5	0	7
F70-F79	Mental retardation	Up to 16	12	5	0	7
H53-H54	Visual disturbances and blindness	5 to 8.5	12	6	1	5
Q60-Q64	Congenital malformations of the urinary system	Up to 7	10	4	0	6
K55-K64	Other diseases of intestines	5 to 10	10	3	1	6
R47-R49	Symptoms and signs involving speech and voice	Up to 16	10	6	0	4
H90-H95	Other disorders of ear	3.5 to 9	9	4	2	3
M40-M43	Deforming dorsopathies	5 to 8.5	10	5	3	2

[1] Approximate values (derived from multiple published sources)
[2] Where administration was reported to fall on the cusp of the critical period for organogenesis this was categorised as 'yes'.

Events relating to the musculoskeletal system were reported most commonly (36% of all reports) followed by events relating to the heart and circulatory system (17% of all reports).

- Gender

No substantial difference in the overall number of reports according to fetal gender was observed, with 36% of cases reported to be female, 29% male and in 34% the gender of the fetus was not reported. When specific anomalies were considered, higher proportions (an arbitrary 5% or more) of reports that included anomalies of the musculoskeletal system, the nervous system, the digestive system, the urinary system or cerebral palsy were observed in female offspring and higher proportions of reports that included anomalies of the circulatory system, genital organs or the palate were observed in male offspring. Formal statistical comparisons were not considered appropriate with spontaneous ADR data.

The reason(s) for these differences were unknown but gender disproportions are known to occur for many congenital anomalies. The observed differences may also be influenced by the small numbers of reports and possible remaining duplicate cases.

5.2.3.2 Identification of potential congenital anomaly syndromes

A limitation of analysing reports at the individual event level is that it splits cases for which more than one event may have been reported and does not allow clusters of events that may be commonly co-reported, and may therefore represent a potential syndrome, to be identified. In addition, analysing the data at a very specific level may separate medically

similar conditions that are occurring in the same individual (eg. multiple anomalies involving different limb events).

A relatively crude analysis was therefore conducted to see if there was any clustering of events that could be suggestive of a multiple congenital anomaly syndrome within the 20 most commonly reported ICD 10 blocks (which accounted for 91% of all HPT reports and 69% of events). Events within two different blocks were co-reported in only 3.5% of cases – within the blocks for i) musculoskeletal anomalies and anomalies of the circulatory system and ii) musculoskeletal anomalies and anomalies of the nervous system.

Using the EUROCAT 'multiple congenital anomaly' algorithm 67 cases of possible multiple anomalies were identified. After manual review by expert geneticists[17], 23 of these were excluded as being most likely not true multiples. A further 44 were considered to be possible multiple anomalies but this could not be confirmed as the reports lacked sufficient detail.

The proportion of possible multiple congenital anomalies within the HPT reports (18.7%) was higher than the proportion reported to EUROCAT (8.3%) (Calzolari, 2014) and were more likely to involve visible anomalies of the limbs, orofacial clefts, eyes, or ears, the face and neck. This may not be unexpected as there is a difference between the possible multiple anomalies identified by the algorithm and those identified by geneticists.

Upon manual review the expert geneticists were unable to identify a combination of anomalies, repeated throughout the cases, which could represent an obvious syndrome or syndromes of drug-induced anomalies in the cases reported in association with HPT exposure.

5.2.3.3 Reporting patterns of anomalies with HPTs

- Compared with population congenital anomaly databases

A number of comparisons were undertaken to identify any unusual patterns in reporting of anomalies with HPT use that might have been suggestive of a drug-related effect.

To see if any anomalies had been reported more frequently in association with HPTs than might be expected in the general population comparisons were made between the reported HPT cases and the EUROCAT (European) and BINOCAR (British and Irish) congenital anomaly databases. To identify any such patterns in reporting, the frequency with which each type of anomaly had been reported was expressed as a proportion of all anomaly reports. The proportions of each of the different types of anomalies reported in the HPT-exposed ADR dataset were then compared with the proportions of the same anomalies reported to the EUROCAT international anomaly database and the BINOCAR national anomaly database.

Taking thalidomide as an example, it would be expected that the event reported most frequently in association with its use would be phocomelia and that, as a proportion of all other thalidomide-related anomalies, this would be far higher than the proportion of phocomelia cases within the congenital anomaly databases (which reflects all anomalies that are occurring in the population regardless of the possible reason). If HPTs were causing a specific anomaly it should therefore have been possible to identify this through these comparative analyses. BINOCAR reports to and therefore forms a subset of the EUROCAT data. Not surprisingly, findings for BINOCAR were highly comparable with those for EUROCAT and so results for the comparisons with EUROCAT only are presented here.

Compared with the EUROCAT database, a higher proportion (at least two-fold higher) was observed for 18 anomalies with HPTs (Table 14); a lower proportion (at least two-fold lower) was observed for six anomalies. A further 36 anomalies present in the EUROCAT dataset had no reported instances in the HPT-exposed dataset. The choice of a two-fold difference

[17] Dr Diana Wellesley and Dr Laura Yates

was arbitrary. For completeness, all preferred terms and Higher Level terms that showed more than a two-fold difference are shown in table 14.

Table 14. Comparison of proportional reporting ratios for anomalies two-fold higher or more in the HPT-exposed dataset compared with the EUROCAT dataset (excluding genetic conditions).

	HPT-exposed data, proportions (%) of all anomalies[1] (n)	EUROCAT Proportions (%) of all anomalies (n)[2]	≥2-fold difference HPT vs EUROCAT[3] [95% CI]
Nervous system	22.6 (37)	10.5 (6479)	2.1 [1.6-2.9]
Neural Tube Defects	9.8 (16)	4.7 (6308)	2.1 [1.3-3.3]
Encephalocele	1.8 (3)	0.5 (700)	3.6 [1.2-10.9]
Spina Bifida	7.3 (12)	2.3 (3167)	3.1 [1.8-5.4]
Eye	5.5 (9)	1.6 (1046)	3.4 [1.8-6.5]
Microphthalmos	1.8 (3)	0.3 (208)	5.7 [1.8-17.6]
Anophthalmos	1.8 (3)	0.1 (113)	23.6 [7.0-79.7]
Congenital glaucoma	0.6 (1)	0.1 (176)	4.3 [0.6-31.4]
Transposition of great vessels	1.8 (3)	0.8 (178)	2.3 [0.7-7.0]
Tricuspid atresia and stenosis	1.2 (2)	0.3 (401)	4.1 [1.0-16.4]
Hypoplastic right heart §	0.6 (1)	0.3 (169)	2.3 [0.3-16.6]
Cleft lip with or without palate	8.5 (14)	3.8 (2465)	2.2 [1.4-3.7]
Ano-rectal atresia and stenosis	3.7 (6)	1.2 (754)	3.0 [1.4-6.6]
Bladder exstrophy and/or epispadia	0.6 (1)	0.3 (402)	1.9 [0.3-13.3]
Limb reduction defects	17.1 (28)	2.3 (3107)	7.4 [5.3-10.4]
Situs inversus	1.2 (2)	0.2 (51)	5.2 [1.3-21.6]
Vascular disruption anomalies §	7.9 (13)	3.4 (2212)	2.3 [1.4-3.9]
Lateral anomalies §	2.4 (4)	0.9 (561)	2.8 [1.1-7.4]

[1] The figures in the 'proportions' columns in the tables cannot be summed or expected to equal 100% since an individual anomaly can be captured in multiple rows of the table due to the hierarchical nature of the EUROCAT structure (eg. 'spina bifida' is captured as a stand-alone condition and also under 'neural tube defects')

[2] Changes in the proportions of certain anomalies reported to EUROCAT were observed over time and so the database was analysed over four time periods (1980 to 1989; 1990 to 1999; 2000 to 2009; and 2010 to 2014)

[3] HPT-exposed proportions divided by the lowest EUROCAT proportion in any of the four time periods (considered to provide the most conservative estimate) except for subgroups with incomplete or missing ICD codes, in which case the 2010 to 2014 dataset was used

The anomalies in the HPT group for which the proportion differed most, compared with the EUROCAT dataset were: anophthalmos/microphthalmos (n=3), limb reduction defects (n=28), situs inversus (n=2), and congenital glaucoma (n=1). The anomalies in the HPT group with the lowest proportion compared with the EUROCAT dataset were: anencephaly and similar (five-fold lower, n=1); atrial septal defect (three-fold lower, n=1); and ventricular septal defect (three-fold lower, n=5). In all cases the wide 95% confidence intervals reflected the small number of cases in the HPT-exposed group. Possible remaining duplicate cases in the HPT dataset and differences in the way the two types of data are collected mean that any comparisons between the HPT-exposed dataset and EUROCAT require cautious interpretation.

5.2.4 Comparison of reporting with the Yellow Card database

To identify unusual patterns in the reporting of anomalies with HPTs relative to other drugs the proportions of all major and minor anomalies reported in association with HPTs were compared with the proportions reported for all other medicines on MHRA's spontaneous reporting database (the Yellow Card database).

For this analysis, the data were restricted to UK only cases for both datasets because the large majority of HPT-exposed cases were from the UK. For the comparison, the time period of reporting for the MHRA spontaneous dataset was chosen to match as closely as possible the HPT-exposed dataset. Cases with genetic conditions were removed from both datasets using the Higher Level Group Term for 'Chromosomal abnormalities and abnormal gene carriers' (Table 15).

Results for the comparisons of congenital anomalies reported with use of HPTs versus all other medicines in MHRA's spontaneous reports dataset at the HLGT were unremarkable. A lack of detailed information in the MHRA spontaneous reports dataset prevented analysis at a more detailed level of the MedDRA hierarchy as many reports were coded to non-specific MedDRA terms rather than a specific anomaly. However, compared with other anomalies reported in association with HPTs, limb reduction defects were reported in relatively high numbers (n=14) and were therefore analysed in more detail.

The proportions of anomalies that were coded specifically as 'limb reduction defects' or that could have been limb reductions but were less specifically coded, were calculated for the HPT dataset and the MHRA spontaneous reporting dataset. The HPT reports were associated with a higher proportion of 'reports of 'limb reduction defects' than were all other drugs within the same time period (n=14, 7.3% versus n=21, 2%). However, compared with the cases reported in association with HPTs, which were extremely detailed, the historically-reported cases on the MHRA Yellow Card database had higher proportions of non-specific anomalies such as 'limb malformation' (8.2% versus 4.7%),'congenital musculoskeletal anomaly' (12.6% versus 0%) and 'multiple congenital abnormalities' (15.4% versus 0%) of which some may have included limb reduction defects (see Annex 24).

The large amount of non-specific data in the MHRA dataset, the possible remaining duplicate cases in the HPT dataset and the different modes of reporting of the two sets of reports mean that any comparisons between the HPT-exposed dataset and the MHRA spontaneous dataset for all drugs require cautious interpretation.

5.2.5 Discussion of adverse event data

The general limitations of spontaneous reporting data are outlined in section 3.2.3. A number of specific challenges were also identified when analysing the HPT-specific data. Firstly, reports of adverse events were received from a variety of different separate sources and this increased the likelihood of overlap between datasets, and multiple recordings of the same anomaly. Anonymisation of a substantial proportion of the reports meant that identification of possible duplicates was not possible in many cases. This, together with the small number of reports and even smaller numbers of individual anomalies – in many cases fewer than five – meant that the likely impact of any residual duplicate cases would be magnified.

Secondly, the reports received by MHRA in association with HPTs, including Primodos, or other medicines reporting a congenital anomaly, fell into two broad categories: those reported spontaneously at the time they occurred and those reported retrospectively in response to MHRA's call for information or following the widespread media and Parliamentary interest in HPTs, and Primodos in particular, and a possible association with congenital anomalies. The former cases were reported on the basis of a suspicion that the

Table 15. Proportional analysis for anomalies in the UK HPT-exposed dataset (1959 to 1979*) versus MHRA UK spontaneous data for all drugs (1963 to 1979), excluding genetic conditions.

MedDRA High Level Group Term[4]	UK HPT-exposed data, proportions of all anomalies (%)[1] 1959-1979 (n=193[3])	MHRA UK spontaneous data[2], proportions of all anomalies, (%) 1963-1979 (n=1067)
All Anomalies	100.0	100.0
Cardiac and vascular disorders congenital	16.1	15.4
Musculoskeletal and connective tissue disorders congenital	39.4	35.3
Congenital and hereditary disorders NEC	4.7	16.9
Gastrointestinal tract disorders congenital	11.9	19.5
Neurological disorders congenital	21.1	16.4
Reproductive tract and breast disorders congenital	4.7	2.6
Renal and urinary tract disorders congenital	5.2	4.7
Respiratory disorders congenital	0.5	1.7
Eye disorders congenital	2.1	2.5
Blood and lymphatic system disorders congenital	1.0	0.0
Skin and subcutaneous tissue disorders congenital	1.6	0.1
Ear and labyrinthine disorders congenital	3.6	2.3
Endocrine disorders congenital	0.5	0.5
Hepatobiliary disorders congenital	0.5	0.5

* Dates refer to onset of ADR

[1] The figures in the 'proportions' columns in the tables cannot be summed or expected to equal 100% since an individual anomaly can be captured in multiple rows of the table

[2] Data excludes any HPT-exposed spontaneous reports received by MHRA

[3] UK-only cases, minus genetic disorders

[4] It was not possible to recode the entire MHRA Yellow Card database using the EUROCAT or ICD10 coding systems and so these analyses were carried out with events coded using MedDRA (version 19.0)

adverse event was associated with exposure to the HPT or other medicine; the latter cases were reported after the safety concern had been recognised and generally tended to be far more detailed. Comparisons between the two types of report therefore needed careful interpretation. Because many of the cases did not specify whether the HPT that had been used was Primodos, this review included all cases that recorded use of any HPT.

Thirdly, adverse events are generally reported shortly after they occur and so cases describing obvious physical anomalies may be expected to be reported disproportionately more frequently compared with anomalies that have no obvious external features that would tend to be diagnosed later. Any comparison of these cases with data reported to congenital anomaly databases such as EUROCAT or BINOCAR would therefore also need to be interpreted with care. With the above limitations in mind, a number of key observations on the data were made.

5.2.6 Key observations

- The number of reports of congenital anomalies in the offspring of mothers given HPTs that were considered in the review was relatively small in comparison to their extensive usage and given the known rate of congenital anomalies (see section 2.3.3). Many of these reports were limited by lack of medical confirmation or insufficient case details, and the majority were reported retrospectively.
- One in six of these offspring had more than one congenital anomaly but no consistent pattern could be identified. For some of these it might now be possible to identify an underlying genetic basis.
- The total number of pregnant women who were given HPTs in the UK is not known with precision and so, frequencies of anomalies could not be calculated. The pattern of reporting of the different types of anomalies amongst the babies of these women was therefore compared with the patterns in two congenital anomaly databases and with the MHRA's spontaneous ADR database for all other medicines.
- Comparison of the pattern of congenital anomaly reports in the offspring of women who were given HPTs, with the pattern of anomalies reported to the EUROCAT (international) and BINOCAR (national) databases showed a higher proportion (≥2 fold) in about 30% of anomalies, a lower proportion (≤0.5 fold) in about 10%, and for 60% of anomalies there were no reported cases. The anomalies with the highest proportion in the HPT group compared with the EUROCAT dataset were: limb reduction defects; anophthalmos/microphthalmos; situs inversus; and congenital glaucoma. The anomalies with the lowest proportion were: anencephaly and similar; atrial septal defect and ventricular septal defect. However, the limitations of the data do not allow any firm conclusions to be drawn and the EWG concluded that these differences could have occurred as the result of chance.
- Comparison of the pattern of congenital anomaly reports for the offspring of women who were given HPTs in the UK with the pattern of anomalies reported to the MHRA's spontaneous ADR database for all other medicines, showed a higher proportion of reports specifically describing 'limb reduction defects' in the HPT-exposed dataset and a lower proportion of 'congenital musculoskeletal anomalies'. However, the limitations of the data do not allow any firm conclusions to be drawn and the EWG concluded that these differences could have occurred as the result of chance.

5.2.7 Overall conclusion on adverse event data

The available adverse event reporting data had many limitations but did not support a causal association between use during pregnancy of HPTs, including Primodos, and congenital anomalies. Anomalies reported in association with HPTs were largely those that are clearly visible at birth and which occur relatively frequently in the general population.

5.3 Epidemiological evidence

This section summarises the epidemiological evidence for a possible association between the use of sex hormones, and particularly those containing NETA and EE, in early pregnancy and congenital anomalies. To ensure all relevant evidence was captured, studies that investigated the use of sex hormones in: i) pregnancy diagnosis, ii) threatened/ recurrent miscarriage and iii) oral contraception were reviewed.

The different doses and administration schedules of the hormones in these three different indications as well as the different levels of risk of the treated populations for congenital

anomalies meant it would not have been appropriate to include them in the same review. Separate reviews within each of the indications were therefore conducted.

Further information on all the reviews is provided in Annexes 25 -29.

5.3.1 Data obtained

Literature searches identified published epidemiology papers with information on NETA and/or EE and congenital anomalies and included: case reports, case series, letters to journals (especially those including relevant data), research studies, published systematic reviews or meta-analyses, review articles, editorials and other commentaries. Literature searches were performed throughout May to July 2016 using Embase and Embase Alert (Medline records are loaded into Embase) databases which were accessed via the ProQuest system. No date or language restrictions were applied to any of the searches, which used 12 relevant exposure terms (eg HPTs, Primodos, norethisterone, ethinylestradiol) combined with each of 69 outcomes for congenital anomaly (relating to the terms in the EUROCAT Description of Congenital Anomaly Subgroups). All search terms are listed in Appendix 1 of Annex 25.

The results of the literature searches (titles and abstracts of all publications) were reviewed to identify suitable papers for more detailed assessment. If there was uncertainty about whether a paper should be included, a full copy (and a translation of foreign language papers where necessary) was obtained and reviewed. The initial approach to this review was broad: all potentially relevant publications were included in the first assessment round, even if the paper did not include all the specific search terms.

Hand-searching of references from individual papers and references that were identified by MA holders whose predecessors marketed an HPT product was also performed to find any other relevant articles that might have not been captured via the ProQuest search method.

Publications in the following categories were excluded from the review of epidemiology data:

- animal or other pre-clinical studies
- papers with no specific data on congenital anomalies (eg. general reviews of the efficacy and general safety of hormones)
- studies specifically examining chromosomal disorders
- papers on a completely different treatment or irrelevant study population (eg. use of NETA for menopausal hormone treatment)
- studies which investigated the relationship between oral contraceptives and pregnancy outcome, but which had no specific data on exposure during pregnancy, ie. studies that examined oral contraceptive exposure prior to conception. Papers that included data on prior exposure as well as accidental exposure during pregnancy were included.

5.3.2 General methodological considerations

In general, the interpretation of epidemiological studies to confirm or refute a causal association involves consideration of several criteria including:

- the validity of the statistical association (dependent on the robustness of the study design in relation to bias, confounding and chance)
- the strength or magnitude of the association (a larger observed association is more likely to be causal)
- consistency between different studies
- biological plausibility for the effect.

Whilst a causal association between an exposure and event can never be proven or ruled out with absolute certainty through epidemiological studies, those effects that are assessed in studies well-designed to minimise bias and confounding, and that show consistency with other studies are easier to draw conclusions from than those effects supported by studies with major limitations in study design and/or conflicting results.

A number of general limitations and biases exist which can serve to either falsely inflate results or obscure true associations, and these were common to many of the studies included in this epidemiological review, as follows:

Control or comparator group

Lack of a comparator/control group (eg. in papers describing case reports or case series) means results cannot be put into context with regard to the expected number of cases occurring naturally, in the absence of drug treatment. With a relatively high natural incidence of many anomalies it is essential to have an appropriate comparator group to assess whether, for any particular anomaly, the observed rate is higher than the background rate. The use of an inappropriate comparator may either inflate or obscure an association.

- Use of controls/comparators with a different susceptibility to the outcome is the most important limitation as it can affect the size of effect seen. It is therefore important to compare women who all choose to have a pregnancy test. This is because the risk of adverse outcomes, including congenital anomalies, in women who choose to have a test may differ from those who chose not to. This phenomenon may also be observed in women who show signs of possible miscarriage in early pregnancy and are prescribed sex hormones to support the pregnancy, and who will have a different baseline risk of anomalies than women with a history of normal, uneventful pregnancies.
- The selection of historical controls from a different time period, or controls from a different region or hospital potentially is also an important consideration as it can result in selection bias eg. prescribing practices may be different in different time periods, regions or hospitals, so patients may have different characteristics/baseline risk.
- Using healthy babies as controls increases the likelihood of recall bias because mothers of babies with congenital anomalies may be more likely to remember possible exposures to medicines than mothers of babies with no anomaly. In retrospective studies, there may also be increased pressure from investigators on mothers of babies with congenital anomalies to remember what medicines they took during pregnancy.
- The use as controls of babies with a different anomaly than the one of interest may mask a potential increased risk, if the risk affects multiple different anomaly types.

Role of confounding factors

Confounding factors are associated with both the exposure of interest and the outcome of interest. Residual confounding can lead to bias that distorts the magnitude of the relationship between the exposure and outcome of interest.

- Incomplete or no matching of cases and controls. The cases and controls are therefore not similar and may have a different baseline risk for the outcome of interest.
- Susceptibility/indication/protopathic bias due to inherently different risk in women exposed/unexposed to HPTs. For example, many women who proactively sought to

know their pregnancy status and who were given an HPT may have a different baseline risk for congenital anomalies (perhaps because of previous complications of pregnancy). Similarly, women given high dose hormones for preventing miscarriage may have been more likely to have a past history of a vaginal bleeding/spontaneous abortion.

- Residual confounding due to unidentified or unrecorded potentially important confounding factors. These factors cannot be adjusted for in any analysis if the data for them are unavailable.

Reliability of information

Accurate and complete information is important to allow reliable comparisons to be made between groups. Many older publications often lack sufficient detail on all aspects of a study to enable informed conclusions on the findings to be drawn.

- Accuracy of estimated gestational age at exposure is required in order to assess whether exposure occurred in the critical period of organogenesis for each anomaly.
- The accuracy and completeness of information recorded in medical notes or remembered by women may be lacking, particularly in retrospective studies in which variables of interest may not always have been recorded.
- Inconsistency in definitions of defects between studies makes comparing the results from different studies challenging. Some studies did not include any definitions of the specific anomalies examined.
- Interviewer bias – prior knowledge of exposure status may (unconsciously) influence the decision around whether to include a case or not and how to conduct the questioning.
- Exposure misclassification – lack of verification of exposure from maternal interviews/questionnaires with medical notes.
- Lack of blinding in ascertainment of exposure and outcome (this was often an issue in earlier studies).

Categorisation of exposure

- Inadequate analysis or interpretation of data according to pharmacologically different hormones/doses/durations/indications. Combining results for all exogenous hormones increases the power of the study to detect rarer outcomes, but any potential differences between preparations cannot then be distinguished.

Statistical robustness

Descriptive statistics (eg. in many studies, simple percentages) used to draw conclusions regarding an association may have limitations. A lack of formal statistical testing or presentation of confidence intervals around the percentage makes the data difficult to interpret.

- Statistical methodology was not always specified so the appropriateness of the method cannot be assessed.
- Statistical methods that have been used are not always appropriate eg. test for paired observations used for matched case/control analysis
- Small numbers of women exposed to HPTs within an individual study means the power to detect an association between exposure and anomalies is low. In these

circumstances, even if a real risk exists the study will be unlikely to find a statistically significant association.

- Multiplicity – if many outcomes are investigated and consequently many statistical tests performed then at least one positive association is expected through chance alone. Correction for multiple comparisons is required in such situations.

Publication bias

Publication bias occurs when the outcome of an experiment or research study influences the decision whether or not to publish it. That is, publication not only depends on the quality of the study but on the hypothesis tested and the significance and direction of the effects detected. For example, studies in support of an association may be more likely to be both submitted for publication and to be published by a journal than studies that find no association.

5.3.3 Limitations of the evidence

As may be expected from studies conducted so long ago, the limitations described above were common to many of the studies included in this review of the epidemiological evidence. The design and methodological rigour of many of the studies identified for review was not consistent with today's standards; statistical packages such as logistic regression software that became available in the 1980s were responsible for raising the standard of data analysis where, before, investigators had to rely on less sophisticated techniques such as stratification of data.

In many studies, neither the type of hormone evaluated nor its indication for use were clearly specified, with studies very often pooling data from women taking hormones for a number of reasons including bleeding during pregnancy, pregnancy diagnosis, threatened miscarriage, menstrual irregularities or contraception. As dose and timing of administration of a drug and the susceptibility of the treated population to the outcome are critical factors when considering a possible effect on the developing fetus, these limitations made interpretation of the findings from several studies especially difficult.

In many papers timing of exposure during gestation was either not reported, or was too imprecise to draw conclusions regarding the types of anomalies reported. For example, exposure was often stated as 'first trimester' and for some studies examining oral contraceptives and congenital anomalies, pre-conception exposure was combined with post-conception exposure. Studies were generally too small to detect anything other than a large marked increase in risk of adverse outcomes of pregnancy with use of NETA and EE.

Although the same sorts of biases may also occur in epidemiology studies conducted today, there is greater awareness of this issue. Methods for controlling the biases are more advanced and studies are generally reported better. Accordingly, many of the later studies did attempt to address at least some of the concerns with the earlier studies.

5.3.4 Review of sex hormones used in pregnancy diagnosis and congenital anomalies

This section presents the review of HPTs and congenital anomalies. Full assessments of the epidemiological data on the use of EE and or NETA for oral contraception or to prevent threatened or recurrent miscarriage are provided at Annexes 28 and 29.

5.3.4.1 Literature search

The ProQuest search identified 4 390 publications of which 4 227 were excluded according to the criteria listed in section 5.3.1), leaving 163 publications for further review. Hand-

searching identified a further 12 papers. A total of 175 papers were therefore reviewed in full, of which a further 78 were excluded to leave 97 that were included for further evaluation.

5.3.4.2 Development of a quality scoring system

The EWG recognised the difficulties in summarising a large number of studies, especially when comparing studies with different designs. Formal meta-analysis was not considered appropriate, because the studies were not sufficiently robust, were too heterogeneous in design and because the weighting system is usually based on study size which, given the extensive limitations of many of the studies would not have been appropriate. Similarly, a numerical scale was not explored due to the subjectivity that would be introduced when deciding on weights to be used.

Applying the scientific rigour expected for current studies as inclusion/exclusion criteria for assessment of these data would have excluded the majority of the studies that were identified for full evaluation. The data were therefore examined using a formal quality scoring system, specifically developed by tailoring the assessment criteria to the most important limitations of the studies identified from the preliminary review. These included: potential exposure misclassification/recall bias; confounding by indication; and sample size. Seven criteria, A to G, that examined different aspects of the studies, were agreed as shown in table 16. Only studies that included a comparator or control group were included.

All studies were assessed according to this pre-defined set of quality criteria, using a traffic light scale of green/amber/red to indicate for each whether the quality criterion was considered to be good, moderate or poor quality, respectively. An overall quality score for each study was not produced as the criteria are not considered to be of equal importance; to develop a weighting system for the criteria would introduce much subjectivity into the system.

5.3.4.3 Evaluation of data

Summary statistics for each study were presented using forest plots and, if not available from the publication, these were calculated from proportions data where possible. Numbers of events were also presented for each study. Absolute rates could not be calculated as the majority of studies were case-control. The plots displayed information on the following:

a) study type (cohort/case-control)

b) timing of administration of HPT exposure relative to organogenesis

c) funding (pharmaceutical industry/non-industry)

d) geographical location.

The following sections present a high-level summary of the study findings for the different anomaly types. A full description of the studies and forest plots for each of the anomaly types are provided at Annexes 26 and 27.

5.3.4.4 Evaluation of studies

Summary findings are presented below for each type of congenital anomaly. Forest plots were developed for each anomaly and may be found in Annex 27. Plots for congenital heart defects, limb reduction defects, GI defects and VACTERL are provided as examples. These plots are intended solely as a graphical representation to aid the interpretation of the study results and were considered alongside the quality assessment for each study.

In general, these studies had important limitations in their design (ie. were judged to be of poor quality with respect to at least one of the selected criteria). In addition to the quality assessment (which was based on the individual study design), careful consideration of the

individual study findings alongside other factors such as magnitude of effect and consistency with the findings of other studies was necessary to make an informed judgement about the strength of an association between exposure to NETA and EE and each anomaly type.

- Nervous system defects

Eleven studies that evaluated an association between HPTs and nervous system defects were assessed. Two studies (Oakley, 1973 and Lammer, 1986) included multiple different nervous system outcomes and therefore feature several times in the forest plot. Seven studies specifically investigated neural tube defects and four studies investigated microcephaly or CNS anomalies.

Neural tube defects: seven studies specifically investigated neural tube defects (spina bifida, anencephaly (or both), encephalocele or 'neural tube defects'). The studies show conflicting results with some showing an association between HPTs and neural tube defects and others not. The magnitude of the observed risk varies greatly, with point estimates ranging from 0.64 to 8.57.

Almost all the studies on neural tube defects were judged to be at risk of bias with a poor quality score in criteria relating to selection of controls, exposure ascertainment and/or risk of confounding making the findings difficult to interpret. The study by Lammer in 1986 (spina bifida and anencephaly) appears to be the most robust using the scoring system although exposure is grouped hormones rather than HPT-specific. Studies by Laurence (1971), Oakley (1973) and Gal (1972) also appear more robust than some of the others, although in the Gal study, controls were selected from a different population to the cases and the plausibility of the cases in terms of timing of exposure was questionable, while control for potential confounders was lacking for the studies by Laurence and Oakley. The study by Gal (1972) reported a statistically significant result whereas the others did not. The least robust studies are judged to be those by Gaal (1977) and Lammer (encephalocele, 1986).

Other nervous system defects: four studies investigated microcephaly or CNS anomalies. These also showed conflicting results and were all judged to be of overall poor quality.

- Congenital heart defects

Thirteen studies investigating a potential association with congenital heart defects were examined (Figure 2). The study by Ferencz (1980) used three different control groups and therefore appears multiple times in the plot. The majority of studies investigated all congenital heart defects as a combined outcome. Only two studies investigated more specific defects: conotruncal anomalies (Ferencz, 1980) and transposition of the great vessels (Levy, 1973).

Table 16. Quality scoring criteria.

	Score		
Quality criteria	**Green**	**Amber**	**Red**
A. Selection of controls/comparator group			
i. Cohort studies – comparator group needs to have similar susceptibility to the outcome as the treated group. Ideally all participants (exposed and unexposed) should come from the same population and loss to follow-up should be minimal	Clear evidence that all participants were selected from the same population and there was minimal loss to follow-up.	Some evidence (or suspicion) that exposed and unexposed participants were selected from different populations and/or there was some loss to follow-up.	Strong evidence (or suspicion) that exposed and unexposed participants were selected from different populations and/or there was loss to follow-up
ii. Case-control studies – control group needs to have the same chance of exposure as the cases. Ideally controls should be selected from the same population as cases.	Cases and controls were selected from the same population.	Some evidence (or suspicion) that controls were selected from a different population that the cases.	Cases and controls were selected from different populations.
B. Exposure ascertainment			
i. Exposure misclassification may be differential or non-differential. Long delay between exposure and ascertainment may result in exposure misclassification due to issues of recall. If this is non-differential, then there will be a bias towards the null. ii. Ascertainment of exposure with knowledge of the outcome (birth defect) may cause recall bias where mothers with malformed babies may be more likely to recall exposure than those with normal babies. iii. Exposure misclassification may also affect the timing of exposure in relation to the critical period of organogenesis Ideally, exposure should be ascertained from medical records.	Studies where exposure was ascertained from medical records	Studies where exposure was ascertained from maternal interview/questionnaire prior to birth	Studies where exposure was ascertained after birth of the child risking either non-differential (bias towards the null) or differential misclassification through recall bias (bias towards an increased risk).
C. Confounding factors			
i. Factors associated with both exposure and outcome that can affect study results (eg. history of anomalies) ii. These factors must be adequately recorded and controlled for in the analysis by statistical adjustment, stratification or matching Ideally, there should be extensive adjustment for potential confounding factors using appropriate statistical methodology	Extensive adjustment for potential confounding factors using appropriate statistical methodology	Some adjustment for potential confounding factors using appropriate statistical methodology	No adjustment for any potential confounding factors (including studies where information on confounders was collected but not used to adjust the analysis)

D. Definition of exposure			
i. Many different hormones have been used in early pregnancy in HPTs, oral contraceptives and to treat threatened abortion. ii. Grouping of all exogenous hormone use together even across different indications makes the results of studies difficult to interpret for specific indications/drug substances Ideally, studies should focus on exposure to HPTs only	Study focussed on exposure to HPTs only	Study includes exposure to different hormone preparations but results for HPTs are presented separately	Study groups all exogenous hormones preparations together
E. Sample size			
i. Studies must be adequately powered to detect an association. Small studies may either fail to detect an association due to small numbers of exposed cases; or be subject to large fluctuation with the inclusion of a single case in either group and have very wide confidence intervals around the point estimate that makes interpretation difficult. Ideally studies should be well powered to detect an association with a particular defect	Study was well powered to detect an association with a particular defect with either 80% power to detect a two-fold risk or 10+ exposed cases/controls	Study had reasonable power to detect an association with all anomalies but not for individual defects. Power was between 50-80% to detect a two-fold risk or 5 to 10 exposed cases/controls.	Study was too small with insufficient numbers of exposed cases and lacks power to either detect an association or results in a risk estimate with very wide confidence intervals. Power was less than 50% to detect a two-fold risk or fewer than five exposed cases/controls
F. Biological plausibility			
i. The timing of exposure is very important for assessing the biological plausibility for an exposure to cause a birth defect. If the exposure is during the critical period of organogenesis for a particular defect, then an association is biologically plausible. Ideally, exposure should be within one week of the critical period of organogenesis for the observed anomaly	Exposure was within one week of the critical period of organogenesis for the observed anomaly	Not applicable	Exposure was more than one week outside the critical period of organogenesis for the observed anomaly
G. Multiplicity			
i. If many outcomes are investigated then at least one positive association is expected through chance alone. ii. Multiplicity should be adjusted for in the analysis where many outcomes are being investigated. Ideally, study had few outcomes with no risk of multiplicity	Study had few outcomes with no risk of multiplicity	Study had multiple outcomes but controlled for multiplicity	Study had multiple outcomes but did not control for multiplicity

Most studies showed a small increased risk associated with exposure to sex hormones; most did not reach statistical significance. The point estimates ranged from 1.05-6.00. In five studies showing a non-significant risk (Goujard, 1979; Hadjigeorgiou, 1982; Nora, 1978 (cohort); Levy, 1973; and Ferencz, 1980), the sample size was inadequate to detect a statistically significant effect.

Many studies were judged to be at risk of bias due to: selection bias, exposure misclassification resulting in recall bias and/or lack of control of confounding factors, making the findings difficult to interpret. The study considered to be most robust (Heinonen, 1977) showed a statistically significant two-fold increased risk of cardiovascular anomalies but the plausibility of some of the cases in terms of the timing of exposure was subsequently questioned in other publications (Wilson, 1981; Wiseman, 1984).

- Orofacial cleft

Three studies evaluated a possible association between HPTs and orofacial clefts. Two studies (Lammer, 1986 and Oakley, 1973) both reported outcomes of cleft lip with or without cleft palate and cleft palate alone and therefore feature twice in the plot.

A small non-significant risk was shown for most analyses, with one reaching statistical significance. The study that showed a significant increased risk (Tümmler, 2014) was based on spontaneous reporting data and judged to be of very poor quality, scoring poorly for several of the assessment criteria. The other two studies had sufficient power to detect a significant risk, should it have existed, and did not do so.

- Digestive system and abdominal wall defects

Two comparative studies that investigated a possible association between HPTs and digestive system and abdominal wall defects were reviewed. Both studies (Lammer, 1986 and Oakley 1973) investigated multiple different outcomes and therefore feature several times in the forest plot (Figure 3).

Both studies showed an increased risk of oesophageal atresia, with one value reaching statistical significance. Neither study supported an association with intestinal/bowel/rectal atresia. For diaphragmatic hernia, both studies reported a small non-significant increased risk; the analysis by Lammer (1986) had reasonable power to detect a significant association, should it have existed. The studies investigating abdominal wall defects showed conflicting results.

Both studies were considered to have limitations in their design, with lack of control for confounding factors in Oakley, grouped hormone exposure rather than results specific to HPTs and no adjustment for multiple testing.

- Urinary system defects

Two comparative studies investigated a possible association between HPTs and urinary system defects (Goujard, 1979; Tümmler, 2014). The study by Tümmler included two different outcomes and therefore features twice in the forest plot.

Both studies reported statistically significant associations for three different anomalies: bilateral renal hypoplasia, renal agenesis and bladder exstrophy. Both studies were judged to have major limitations in study design with risks of selection bias, exposure misclassification and a lack of control for potential confounding factors, making the results difficult to interpret.

- Genital defects

Two comparative studies investigated a possible association between HPTs and genital defects (other than virilisation). Both reported non-significant increased risks of different genital defect outcomes. Both studies were considered to have major limitations in study design with risks of selection bias, lack of control for confounding factors and questionable plausibility for some cases.

- Musculoskeletal defects

Limb reduction defects: Five studies investigated an association with limb reduction defects (Figure 4). All showed an increase in risk, with an average increase of about two-fold. This reached statistical significance in the study that was judged to be the most robust in accounting for bias, though it included multiple outcomes, with no adjustment for multiple testing, and the result could therefore be due to chance (Lammer, 1986). The other four studies tended to lack power, limiting their ability to detect a significant increase; they were judged to suffer from a range of different biases (including selection bias and exposure misclassification resulting in recall bias) that could falsely inflate the result.

Other skeletal defects: Three studies investigated other skeletal defects and reported conflicting results. These studies were judged to be of poor quality.

- VACTERL

VACTERL describes a group of congenital anomalies which include Vertebral defects, Anal atresia, Cardiovascular anomalies, Tracheoesophageal fistula, Esophageal atresia, Renal anomalies and Limb defects. Babies born with three of more of these congenital anomalies were considered to have VACTERL syndrome. Three studies for a possible association between HPTs and VACTERL were examined (Figure 5). Two studies (Nora, 1975 and Lammer, 1986) used two different control groups and therefore feature twice in the plot.

The studies showed conflicting results with respect to the risk of VACTERL with exposure to HPTs. The two studies by Nora showed a statistically significant increase whilst the study by Lammer did not but had reasonable power to do so. All the studies were considered to have limitations including: a lack of adjustment for potential confounding factors, potential selection bias and grouped hormone exposure rather than specific exposure to HPTs.

- All congenital anomalies

Eleven studies evaluated a possible association between HPTs and all congenital anomalies. The study by Torfs in 1981 used two different control groups and therefore appears twice in the plot. Six of the studies investigated major congenital anomalies and five investigated all congenital anomalies.

The majority of studies did not show an association between HPTs and congenital anomalies despite most being adequately powered to detect an increase if one existed. One study (Greenberg, 1977) observed a statistically significant risk and is considered to be amongst the better-quality studies for the time, although it did suffer from a lack of adjustment for confounders and the comparator group were not women who also sought a pregnancy test. The study by Michaelis (1983) is also considered to be of better quality and did not observe a statistically significant increase but was well powered to do so. The study by Torfs was the only one to compare women using HPTs with women using other pregnancy tests, thus reducing confounding by indication.

5.3.5 Key observations

In general, the studies were judged to have important limitations in their design and to be of poor quality with respect to at least one (and up to five) of the seven quality scoring criteria. This made it difficult to draw robust conclusions: that is, the evidence from many of these studies was insufficiently strong to demonstrate either that there was a causal association between HPTs and congenital anomalies or conversely that there was no possibility of a causal association.

Using a pre-specified, tailor-made quality scoring system, and the definitions for strength of the evidence as outlined in section 3.3.2, the following overall observations were made:

- There was limited evidence for a weak association between the use of HPTs and congenital heart defects, limb reduction defects, and oesophageal atresia, but it was felt this could be due to chance or confounding
- The evidence reviewed did not support an association between the use of HPTs and neural tube defects, orofacial clefts, digestive system and abdominal wall defects, skeletal defects (other than limb reduction defects) or overall congenital anomalies in the fetus but the quality of the evidence is limited
- From the evidence available it was not possible to draw any conclusions about a possible association between the use of HPTs and urinary system or genital defects, nervous system defects (excluding neural tube defects), or VACTERL.

5.3.6 Overall conclusion on epidemiological data

While the quality of the available epidemiological evidence was generally very limited, no strong associations were found between the use of HPTs, including Primodos, during pregnancy and any single anomaly, or any pattern of anomalies. The weak associations that were observed could have occurred by chance or confounding.

5.4 Overall conclusion on the evidence for congenital anomalies being associated with exposure to HPTs in pregnancy

The totality of the available evidence from pharmacology, non-clinical, epidemiological and adverse event reporting data was very limited and did not, on balance, support a causal association between the use of HPTs, such as Primodos, by the mother during early pregnancy and congenital anomalies in the child.

On the basis of these findings the EWG recommended that:

1. A full up-to-date genetic clinical evaluation, in line with current best practice, should be offered to families of the Association for Children Damaged by HPTs, whose lives have been impacted by adverse pregnancy outcomes and who were given HPTs to diagnose pregnancy.

2. Electronic Yellow Card reporting should be made available at point of care, including at scanning in early pregnancy, to all those who suspect an adverse outcome of pregnancy, including a congenital anomaly, in association with exposure to any medicine in pregnancy. In particular, Yellow Card reporting should be included in relevant clinical systems and promoted in a dedicated campaign to raise awareness of this possibility.

3. There should be regular, independent review by experts of all suspected adverse drug reactions in pregnancy that are reported by healthcare professionals and women in the UK to the MHRA. The CHM should publish the findings and conclusions in their annual report.

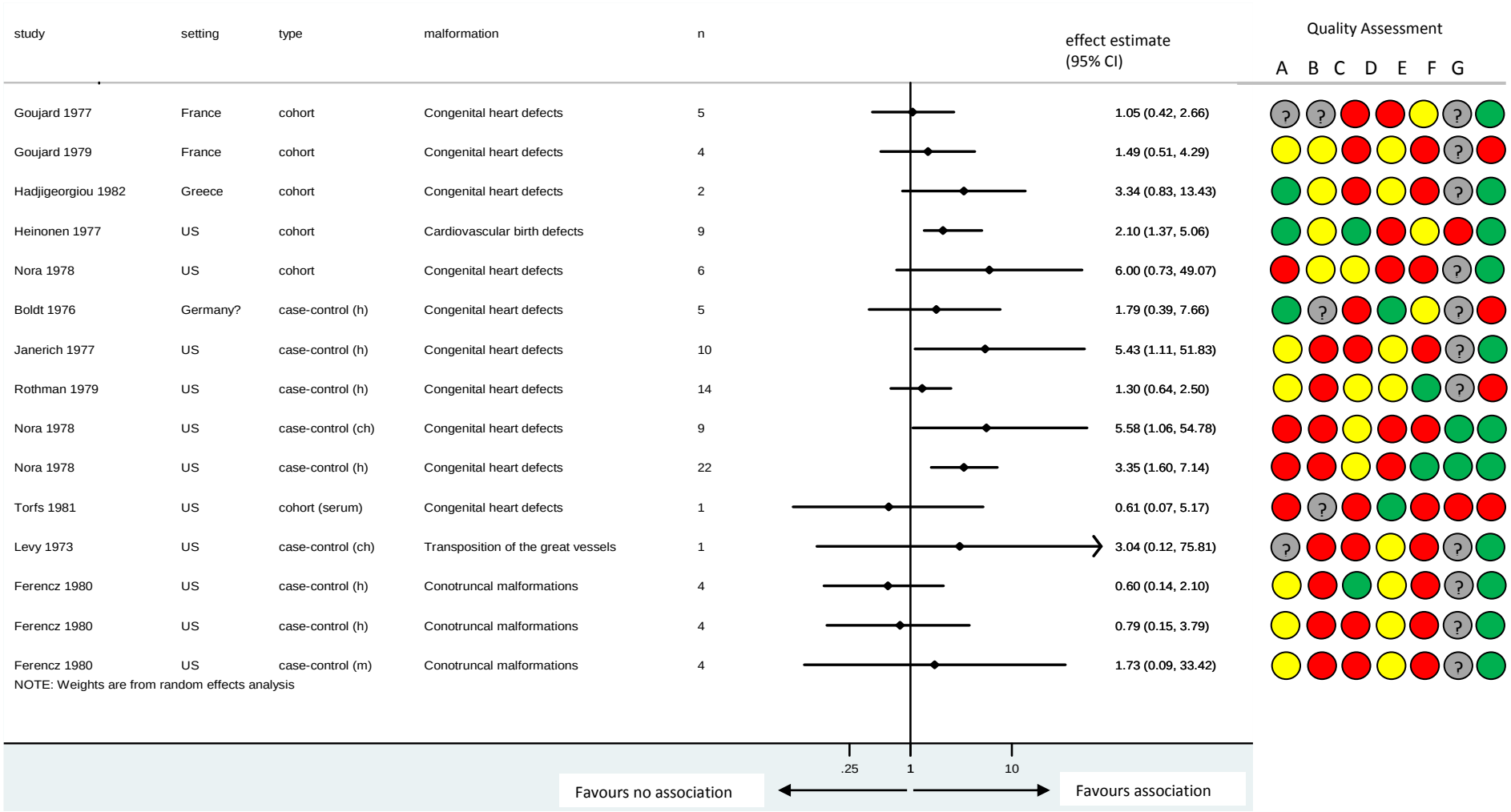
4. A scientific workshop should be held to bring together different disciplines to consider:

    a. how results from studies in pregnant animals, with individual medicines or the chemical class, can be made more accessible in order to help predict and assess the potential effects of medicines in pregnancy

    b. the feasibility of using computer modelling and molecular structure alerts to generate safety signals from animal and in vitro data and to prioritise drugs for further study.

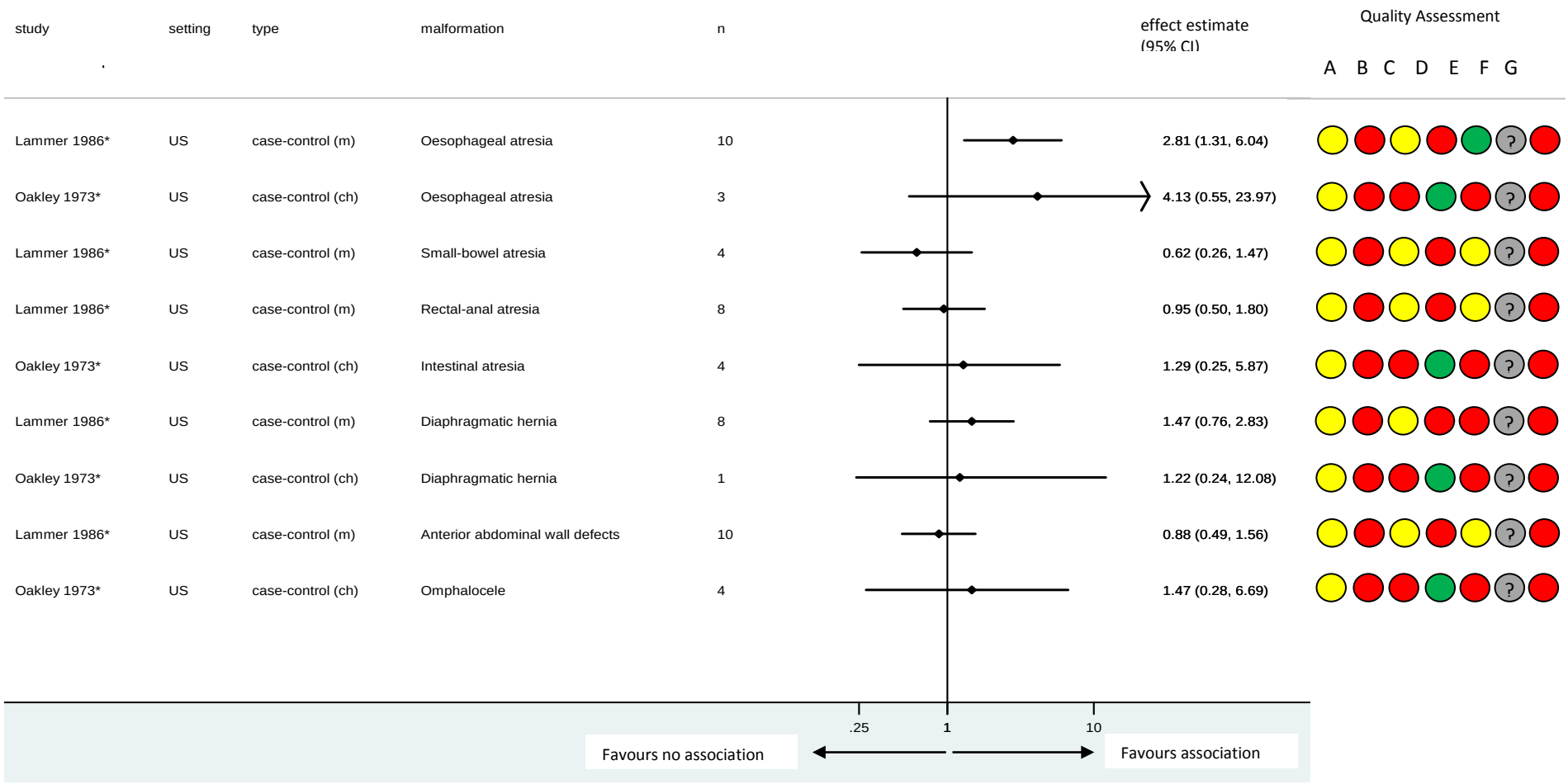
5. A strategy to co-ordinate and promote research on mechanisms of teratogenicity in early embryonic development and how the actions of and reactions to drugs vary with the individual's genes should be taken forward with appropriate experts in the field.

Figure 2. Forest plot and quality assessment of epidemiological studies of HPTs and heart defects



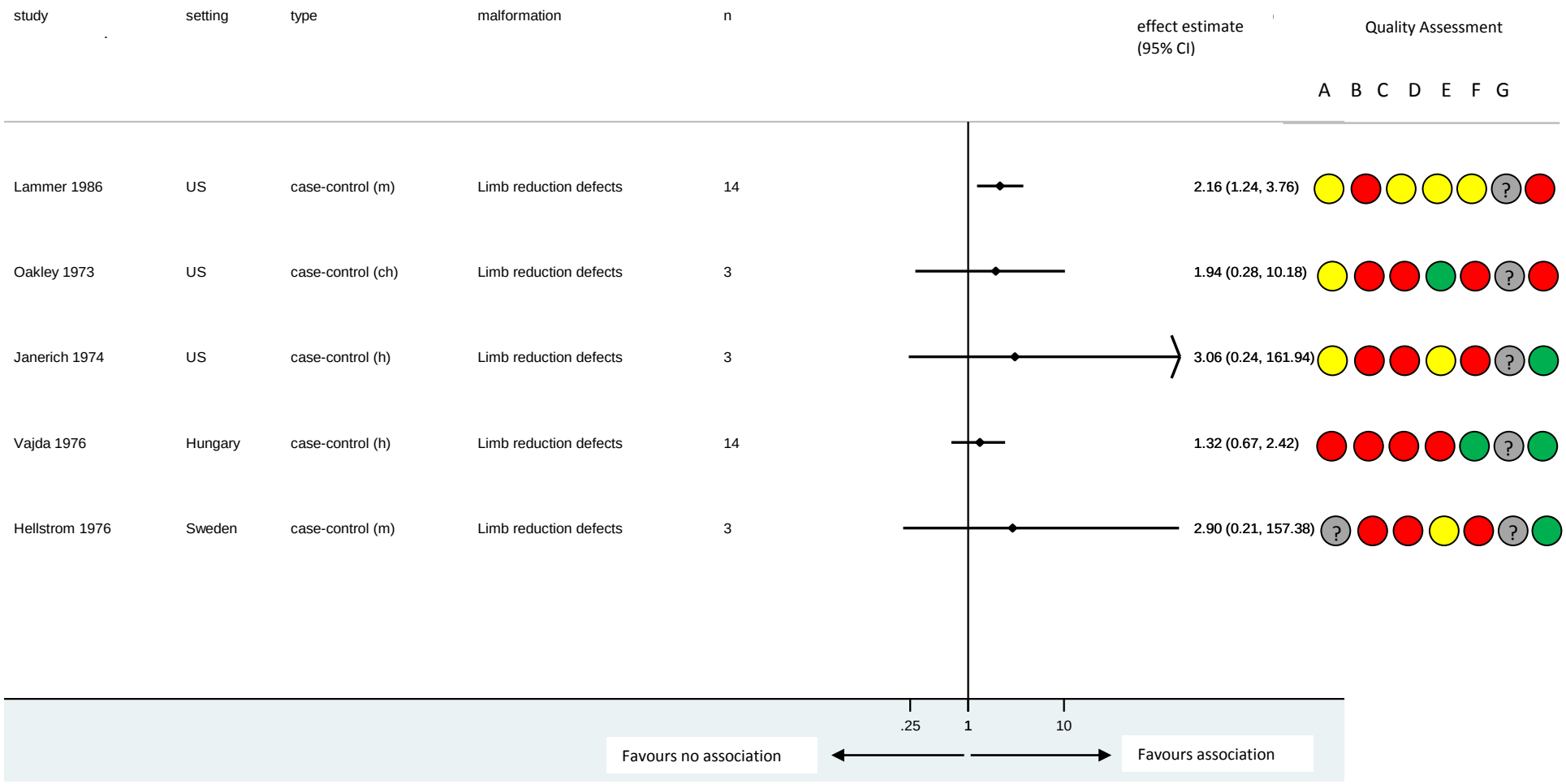
* study reports multiple outcomes, + study sponsored by industry, n = number of exposed cases, case-control (h) = healthy controls, case-control (m) = malformed controls, case-control (ch) = controls with chromosomal defects, A = selection of controls/comparator group, B = exposure ascertainment, C = confounding factors, D = definition of exposure, E = sample size, F = biological plausibility, G = multiplicity

Figure 3. Forest plot and quality assessment of epidemiological studies of HPTs and digestive system and abdominal wall defects



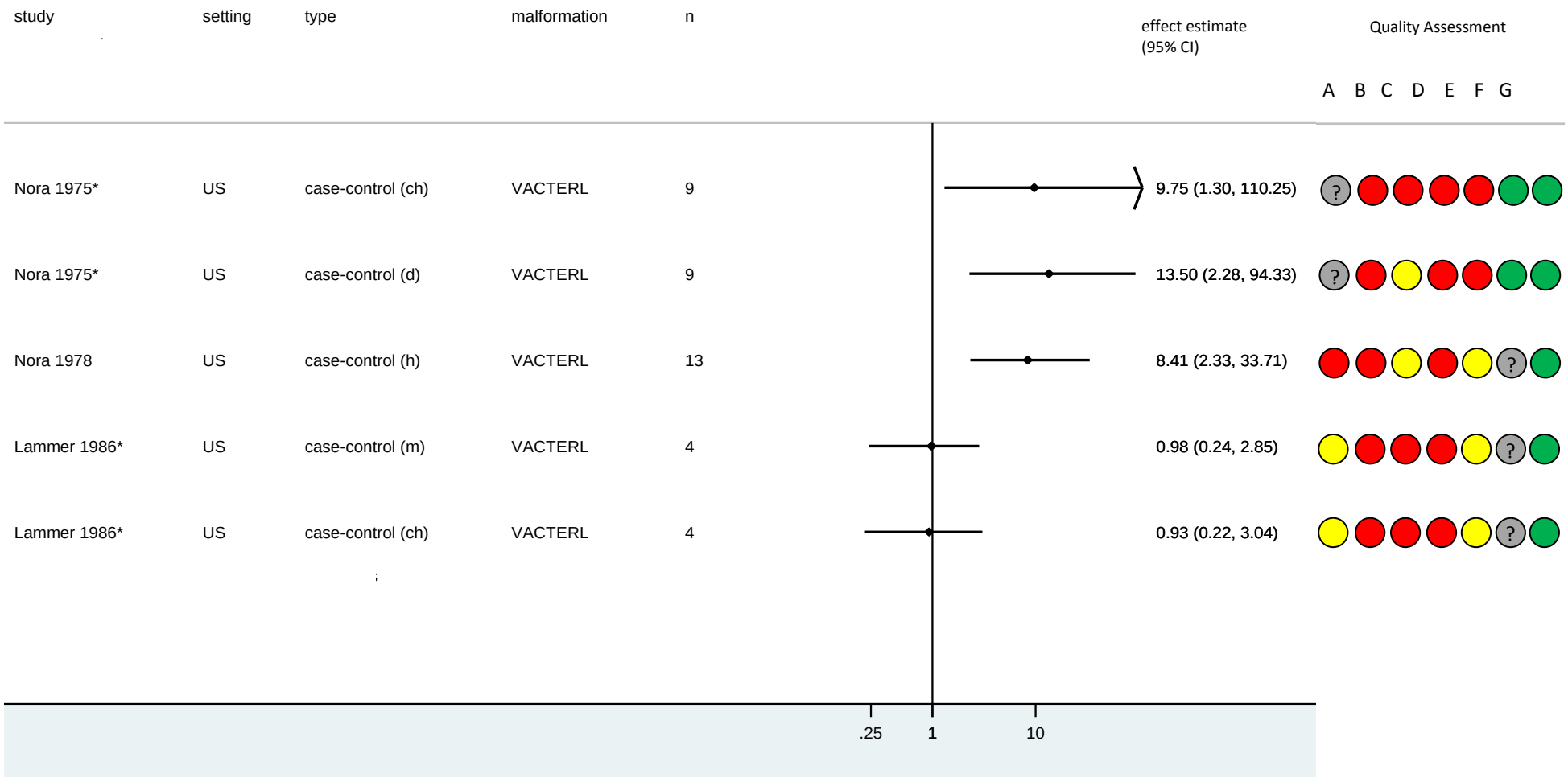
* study reports multiple outcomes, n = number of exposed cases,case-control (h) = healthy controls, case-control (m) = malformed controls, case-control (ch) = controls with chromosomal defects, A = selection of controls/comparator group, B = exposure ascertainment, C = confounding factors, D = definition of exposure, E = sample size, F = biological plausibility, G = multiplicity

Figure 4. Forest plot and quality assessment of epidemiological studies of HPTs and limb reduction defects



n = number of exposed cases,case-control (h) = healthy controls, case-control (m) = malformed controls, case-control (ch) = controls with chromosomal defects, A = selection of controls/comparator group, B = exposure ascertainment, C = confounding factors, D = definition of exposure, E = sample size, F = biological plausibility, G = multiplicity

Figure 5. Forest plot and quality assessment of epidemiological studies of HPTs and VACTERL.



* study reports multiple outcomes, n = number of exposed cases,case-control (h) = healthy controls, case-control (m) = malformed controls, case-control (ch) = controls with chromosomal defects, case-control (d) = disease controls, A = selection of controls/comparator group, B = exposure ascertainment, C = confounding factors, D = definition of exposure, E = sample size, F = biological plausibility, G = multiplicity

6. EXPOSURE TO HORMONE PREGNANCY TESTS IN EARLY PREGNANCY AND MISCARRIAGE OR ABORTION

A possible association between HPTs and miscarriage has been proposed for many years and a number of possible theories could explain this:

1. HPTs were associated with miscarriage when used at the licensed dose
2. HPTs were associated with miscarriage when used at higher than licensed doses
3. HPTs were used as an indirect abortifacient
4. The withdrawal bleeding in non-pregnancy was mistaken for miscarriage

This chapter summarises the available evidence for a possible abortifacient effect of HPTs following their administration during early pregnancy. Further information is provided in Annexes 20 and 30.

6.1 Anecdotal evidence

Various pieces of anecdotal evidence were identified in relation to the above theories during review of the extensive documentation that was submitted. While anecdotal evidence cannot confirm or refute any of the above theories it can provide a useful perspective on the knowledge or beliefs of that time.

- Associated with miscarriage when HPTs used at the licensed dose

No anecdotal evidence was identified in support of this theory. An opinion piece describing the campaign against estrogen-progestogen drugs in India, submitted through the public call for evidence, explained the background to their use as a pregnancy test and other gynaecological purposes and recommended that "*they are not essential for the treatment of menstrual disorders, unreliable as pregnancy tests, and ineffective as abortifacient".*
A letter by Harlap in the Lancet (1975) stated that some mothers in their study had taken hormonal drugs to induce abortion because there was a popular supposition among women – exploited by some gynaecologists – that such tests were abortifacient.

- Associated with miscarriage when HPTs used at higher than licensed doses

A translated German newspaper article from the Landesarchiv Berlin suggested that "*Duogynon was also an underhanded abortion drug, if one took it at sufficiently high dose.*" (Kamke, 1978). Another translated newspaper article from the Landesarchiv Berlin (author and original source unknown) stated that "*A third area of use for Duogynon was rumoured. According to rumour, the compound could terminate pregnancies if one took a high-enough amount ("after pill"). Interested women were required to obtain the compound illegally, as it requires a prescription, and the doctor could only prescribe two dragees at a time*."

- HPTs were used as an indirect abortifacient

A letter by Brogen to the Medical Journal of Australia in 1975 stated that 18 of 22 patients who were administered HPTs and had their pregnancy confirmed, did not want the pregnancy and requested a termination. This observation led the authors to speculate that the HPTs may have been used in the hope of producing a miscarriage; no evidence or data was presented in support of this.

A translated handwritten note by Schering relating to information from a trip to Korea stated "*As a rule, overdosing [with Duogynon] is used in an abortive manner. If the bleeding does not occur and the woman is pregnant, she has an abortion (semi-legal).*"

- The withdrawal bleeding in those who were not pregnant may have been mistaken for miscarriage

A document provided through the public call for information and written by the All India Drug Action Network ("The Reasons why the ban order of 1987 on high dose EP drugs should be implemented") stated that the "*Use of high dose EP drugs as abortifacient continues, even while such a use is not even recommended or mentioned in medical text books. The origin of the usage for such an indication stems from mal-practice by doctors who led the patients to believe that the withdrawal bleeding was actually abortion induced by them*."

A book chapter provided through the same route ('The case of the deadly pregnancy test', author unclear) described the campaigning in India to outlaw the use of high dose estrogen/progestogen products in India and stated "*two notable applications have been involved: one was to serve as a simple, inexpensive, dependable, and presumably safe pregnancy test; the other – never recommended publicly, never proved, and never effective – was to induce abortion. […] Moreover, since the drugs unquestionably produced bleeding in some women, word spread from woman to woman – and from physician to physician – that the high-dosage products could easily, inexpensively, and safely cause an abortion. And since the hormones could be purchased readily, with or without a prescription, from any pharmacist, women by the millions used them to detect a pregnancy or in the hope of ending it*."

6.2 Mechanistic evidence

6.2.1 Animal studies

In animals, the death or loss of an embryo after implantation in the uterus or fetus is referred to as post-implantation loss, embryo-fetal loss or embryo-lethality. However, the mechanism for expulsion of the dead or non-viable fetuses may vary between species. In rats and mice, a dead conceptus undergoes gradual degradation followed by maternal reabsorption and is referred to as a resorption; in rabbits, a dead conceptus may be reabsorbed or aborted (expelled).

Many of the animal studies described in Chapter 5 reported on rates of embryo-fetal loss as well as malformations. A substance that causes embryo-lethality does not necessarily also cause malformations; embryo-fetal loss may relate to factors that affect the survival of the embryo/fetus rather than to specific developmental effects. However, it is possible that marked embryo-lethality could mask possible developmental effects by reducing the number of foetuses that could be evaluated.

The findings for embryo-lethality in the animal studies identified are reported here.

6.2.1.1 Effect of norethisterone

Embryo-lethality was observed in studies in which NETA (and related progestogens) was given to mice, rats, rabbits and non-human primates during organogenesis. Embryo-lethal doses of NETA were generally higher than those used in Primodos. The lowest doses of NETA that had no effect on embryo-fetal survival in these studies are shown in Table 17.

Table 17. No Observed [Adverse] Effect Level (NO[A]EL) doses for embryo-fetal survival of NETA; X-fold difference to Primodos.

Species	NO(A)EL	X-fold difference vs Primodos*
Rats Schering studies #2330 (1976) and #5303 (1978); Suzuki, 1978	10 mg/kg/day by mouth	9
Mice Schering study #5304	≥48 mg/kg/day by mouth	≥32
Rabbits Schering study #2300 (1976)	<0.1 mg/kg/day by mouth	<0.1
Rhesus macaques Wharton, 1964	<25 mg/animal intramuscular 5 days/week	<25
Baboon Beck, 1982	<2.5 mg implant throughout pregnancy	–

* based on human equivalent dose calculated on a mg/m^2 basis: assume Primodos daily dose of 10 mg norethisterone acetate, 0.002 mg ethinylestradiol for a 60kg women. Based on 1 mg/kg = 4.3 mg/m^2 for a rat; 1 mg/kg = 11.8 mg/m^2 for a rabbit; 1 mg/kg = 38.8 mg/m^2 for a human.

The mechanism for the embryo-lethal effect of NETA remains unknown but may involve disruption of the fetal maternal endocrine relationship required for the maintenance of pregnancy.

6.2.1.2 Effect of ethinylestradiol

An abortifacient effect with high doses of estrogens was also observed in many studies and is now considered to be a well-recognised effect (Steinetz, 1976; Badawy, 1978; Attia, 1980; Sarkar, 1986; Matsuura, 2004). In studies conducted by Schering high doses of estrogens given to rodents and rabbits in early pregnancy caused embryo-lethality. The lowest dose at which no effect on fetal survival was observed was equivalent to about nine times the dose of EE in Primodos, when exposures are estimated based on dose per body surface area. The mechanism for an embryo-lethal effect with high-dose EE also remains unknown but, as with NETA, may involve disruption of the fetal maternal endocrine relationship required for the maintenance of pregnancy.

6.2.1.3 Effect of norethisterone and ethinylestradiol in combination

As may be expected from the effects of NETA and EE, studies in which NETA (or structurally related synthetic progestogens) combined with EE (or related estrogen formulations) was given to mice, rats, guinea pigs and rabbits consistently showed embryo-lethal effects, depending on the doses and timing of administration.

- Mice

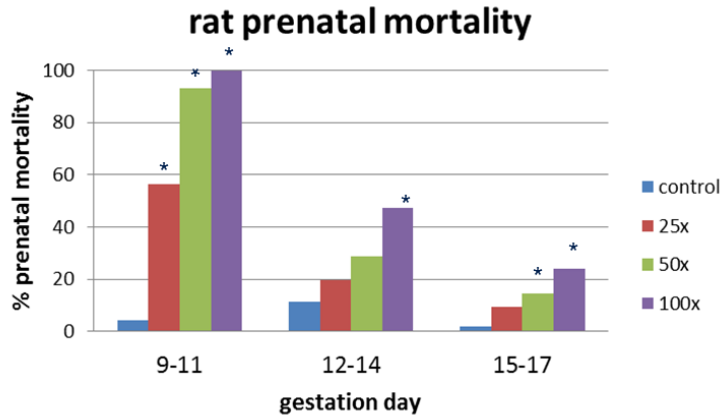
Various doses of NETA and EE at a ratio of 60:1 or 80:1 were given to mice on days 7 to 13 of gestation (Schering studies dated 20/12/1965 and 06/12/1965), and on days 6 to 15 of gestation at a ratio of 500:1 (equivalent to the ratio in Primodos, Schering study #3579 (1978)). In these studies resorption rates were consistently increased in the high dose groups compared with the low dose and control groups. In study #3579 (1978) resorption rates were significantly increased with doses of NETA (15.0 mg/kg/d) and EE (0.03 mg/kg/d)

and above (about 90-times the dose of Primodos based on a mg/kg comparison and about 10-fold the Primodos dose based on a mg/m^2 comparison).

– Rats

Three groups of rats exposed to NETA and EE at doses of 0, 25, 50 and 100 times the Primodos dose (calculated on mg/m^2 basis) on gestation days 9 to 11, 12 to 14, or 15 to 17 of pregnancy showed clearly the importance of the dose on embryo-lethality rates (Joshi, 1983; Figure **6**6). Days 9 to 11 of gestation were most sensitive to the effects of NETA and EE.

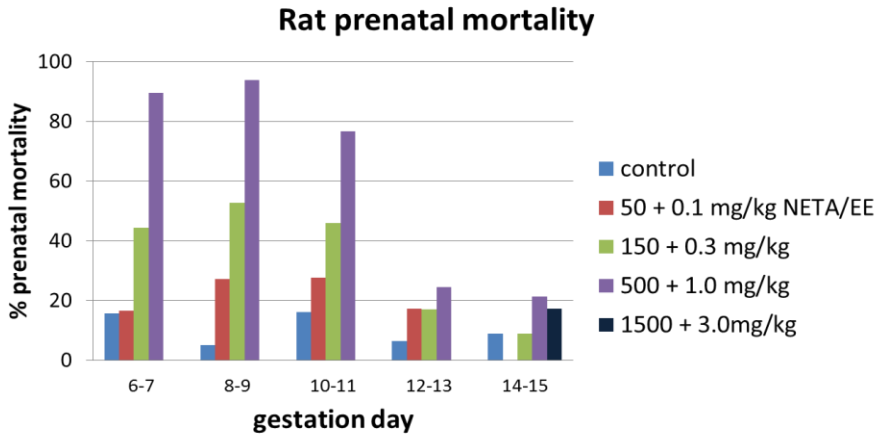
Figure 6. Effect of NETA and EE on embryo loss in rats (Joshi, 1983).



*significant findings versus control taken from publication, calculated by Mann-Whitney U test (U<0.05 for x25 d9-11 and 50x d15-17; U<0.01 for 50x d9-11 and 100x all three gestation periods)

Similar studies in five groups of rats given very high doses of NETA and EE (equivalent to between 300 and 9 000 times the dose of Primodos, based on mg/m^2) for two days of gestation per group, covering the period from days 6 to 15, showed a consistent dose-dependent increase in total prenatal mortality (Figure 7). Peak sensitivity of these rats to embryo loss with EE + NETA was from day 6 to 11 of gestation; some effects were also detected throughout the remainder of organogenesis.

Figure 7. Effect of NETA and EE on embryo loss in rats in studies conducted by Schering (studies #4037 [1979]; #4046 [1979]; #4042 [1979]; #4045 [1979]; 4044 [1979]).



The results are expressed as the total incidence of parental mortality as a % of post-implantation loss in pregnant rats from each group. No statistics were applied.

The very high doses used in these studies reflect the relative insensitivity of rats to the reproductive effects of the sex hormones.

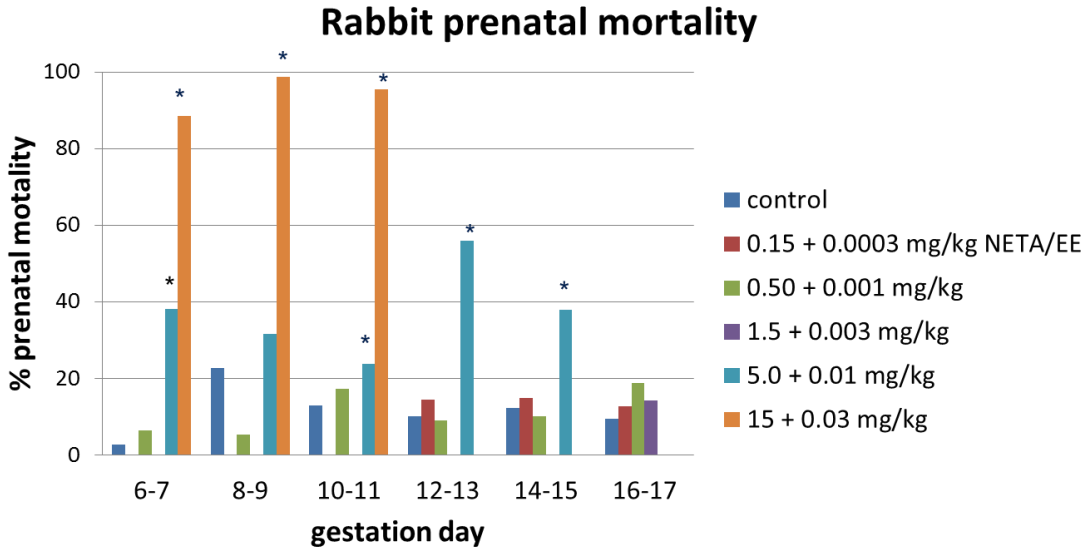
– Rabbits

The rabbit appears far more sensitive than the rat to the effects of NETA and EE on embryo-lethality. A study conducted by Schering (#1443, 1970) administered NETA and EE in a ratio of 500:1 (equivalent to the ratio used in Primodos) to two groups of pregnant rabbits between gestation day 6 to 18 at two different doses. A further group acted as the control. In the high dose group (NETA 5.0 mg/kg/day and EE 0.01 mg/kg/day) no live fetuses were delivered; in the low dose group (NETA 0.5 mg/kg/day and EE 0.001 mg/kg/day) there was no significant difference in the number of live fetuses or resorptions compared with controls.

A further experiment was therefore performed to explore the effect of NETA and EE at the lower end of the dose range (Schering study #3581, 1978). In this study, pregnant rabbits received vehicle control or NETA and EE at doses of 1.5 + 0.03 mg/kg/day, 0.5 + 0.001 mg/kg/day, or 0.1 + 0.0002 mg/kg/day between gestation days 6 to 18. There was clear embryo-fetal loss (51%) in the high dose group although a potential trend to embryo-fetal loss in the mid and low doses was observed. The lowest dose in these studies approximates to the same dose as a single tablet of Primodos, on a mg/kg basis, or to a five-times lower dose, on a mg/m^2 basis.

A parallel set of experiments to those conducted in the rat evaluated the effect of dosing on two consecutive days during organogenesis. The combination NETA + EE was studied in a set of six studies with the combination administered on gestation days 6 to 7, 8 to 9, 10 to 11, 12 to 13, 14 to 15 and 16 to 17 (Figure 8). This covers the period of major organogenesis from implantation to closure of the hard palate in the rabbit. In the rabbit, the doses of NETA + EE were also in a constant ratio of 1:500 but absolute doses were varied between studies. The lowest dose evaluated (0.15 + 0.0003) for gestation days 12 to 13, 14 to 15 and 16 to 17, approximates to the same as the Primodos dose, on a mg/kg basis, and a three-times lower dose, on a mg/m^2 basis.

Figure 8. Effect of NETA and EE on embryo loss in rabbits in studies conducted by Schering (#4036 [1978]; #4038 [1978]; #4039 [1978]; # 4040 [1978]; #4041 [1978]; #4043 [1978])



total incidence of parental mortality as a % of post-implantation loss in pregnant rabbits from each group. *significant difference versus control, as determined by Schering. Details not provided.

The NO(A)ELs for an embryo-lethal effect of NETA and EE from the above studies are summarised in Table 18.

Table 18. NO(A)ELs for embryo lethality with NETA + EE combinations given during organogenesis.

Species	Embryo-lethality NO(A)EL NETA + EE mg/kg/day	X-fold human equivalent dose *
Mouse	5.0 + 0.01 (over 10 days)	~3
Rat	5.0 + 0.01 (over 10 days)	~5
Guinea pig	<0.5 + 0.001 (over 22 days)	<0.6
Rabbit	0.5 + 0.001 (over 2 days)	~1x

* Based on a mg/m^2 comparison of a single dose of Primodos

- Non-human primates

A Schering-sponsored study specifically designed to investigate the potential for teratogenicity in non-human primates dosed rhesus monkeys, cynomolgus monkeys and baboons by oral administration with NETA and EE during days 20 to 50 of pregnancy (Hendrickx, 1987 a and b). Exposures to NETA and EE were estimated by the authors based on multiples of 1x, 10x, 100x, 300x and 1 000x the human dose used in Primodos, based on an approximate human dose of NETA 0.2 mg/kg/day and EE 0.004 mg/kg/day (based on one tablet in a 50kg woman).

Substantial embryo-lethality was seen in all three species at 100x the approximate human dose equivalent of NETA and EE, on a mg/kg/day basis. Although not statistically significant, a trend in increased embryo-lethality was observed from 1x the human dose in the rhesus monkeys and from 10x the human dose in the baboon (Table 19). Embryo-lethality was observed between gestational days 31 and 68 in the rhesus monkey, days 27 and 70 in baboons, and days 44 and 98 in the cynomolgus monkey.

In a parallel arm of the study (Hendrickx, 1987a, b) a combination of natural progesterone + estradiol benzoate given by intramuscular injection induced embryo-lethality at doses of 10x (no significant difference versus controls) and 25x (significant difference versus controls) the human dose equivalent in cynomolgus and rhesus monkeys respectively.

Table 19. Dosing studies in non-human primates and embryo-lethality (Hendrickx, 1987a, b)

Species	Human Dose Equivalent[1]	No. pregnancies[2]	No. of fetuses	Embryo-lethality No. (%)
Rhesus monkey	control	10	9	1 (10)
	1x	11	8	3 (27)
	10x	10	8	2 (20)
	100x	12	4	8 (67)*
Baboon	control	10	9	1 (10)
	1x	10	10	0
	10x	10	8	2 (20)
	100x	10	7	3 (30)
Cynomolgus monkey	control	21	17	4 (19)
	100x	21	10	11 (52)**
	300x	10	4	6 (60)**
	1,000	9	4	5 (57)

[1] Calculated by Hendrickx (1987a,b), based on a mg/kg/d comparison and approximate human dose of 0.2 mg/kg NETA and 0.0004 mg/kg EE.
[2] Treated days 20 to 50 of gestation by oral route
* $p < 0.005$ Fisher's Exact test
** $p < 0.05$ Fisher's Exact test

6.2.1.4 Key observations

- Depending on the species studied, embryo-lethal effects with NETA and EE separately or in combination were seen, most often at daily doses and durations much higher than those used in HPTs.
- Sensitivity to the embryo-lethal effect of NETA and EE was species specific with the rat being highly insensitive and the rabbit highly sensitive.
- Data from non-human primates, the most physiologically relevant animal species, suggested a small increase in pregnancy loss at around the equivalent Primodos dose when given daily for 30 days during early pregnancy.
- The mechanisms of embryo loss are not established but may relate to disruption of the normal maternal embryo-fetal hormonal relationship required for the maintenance of pregnancy; a substance that causes embryo-lethality does not necessarily also cause malformations.

6.2.1.5 Overall conclusion on animal data

Depending on the species studied, embryo-lethal effects with norethisterone and ethinylestradiol separately or in combination were seen, most often at total daily doses much higher than those used in Primodos or for longer durations or both.

6.2.2 Clinical data

A possible pathological effect of Primodos on early human pregnancy involving necrosis and subsequent bleeding in the developing placenta or a reduction in the high maternal levels of circulating progesterone necessary to support early pregnancy was investigated in a series of two studies in women seeking legal termination of pregnancy in Finland (Pulkkinen, 1984).

The first study was double-blind in design. Women who were eight to nine weeks pregnant were randomly assigned to receive NETA (20 mg) and EE (0.04 mg) – equivalent to two tablets of Primodos taken together – or placebo (n=25 women per group). Primodos or placebo were taken at time 0, with ultrasound scans at 0, 24, 48 and 96 hours before dilatation and curettage at 96 hours. The termination of pregnancy products were examined for any pathology indicative of placental damage. The second study was an open label (unblinded) study in which endogenous hormone levels in ten untreated women at 8.5 ± 0.5 weeks of pregnancy were compared to the levels in 11 women who were given NETA (20 mg) and EE (0.04 mg) at 6.8 ± 0.4 weeks of pregnancy. Limited plasma samples were taken during the first 6 hours following treatment.

None of the women who received Primodos miscarried during the observation period. The details of the pathology findings can be found in section 5.1.4.2.

6.2.2.1 Overall conclusion on clinical data

These studies were not large enough to detect any small differences in rates of miscarriage between the treatment groups and found no adverse effects of NETA and EE, at doses equivalent to two Primodos tablets taken together, on the developing pregnancy during the first trimester.

6.3 Epidemiological data

6.3.1 Literature search

Results of the ProQuest literature search used to identify studies on use of HPTs for threatened abortion and congenital anomalies (described in section 5.3) were applicable to this review because the same search terms identified papers with abortion/miscarriage as an outcome or an indication. This search identified 405 publications. After screening of the titles and abstracts 361 publications were excluded based on the criteria described in section 5.3.1. Of the 47 potentially relevant publications identified for further review a further 26 were excluded to leave 21 papers for full evaluation.

6.3.2 Results

Several studies identified for this topic are also discussed in section 5.3 because they reported the number of spontaneous abortions as well as congenital anomalies. The studies described a range of exposures and indications (Table 20) with most relating to the use of hormones for preventing threatened or recurrent miscarriage. Most studies in this indication were conducted in the 1950s and 1960s when the use of hormones to maintain an unstable pregnancy was burgeoning and the efficacy of different combinations was being investigated. The same limitations of the studies as previously set out in section 5.3.2 apply, with the added disadvantage that none of the studies set out with the objective of determining whether sex hormones had an abortifacient effect.

Table 20. Exposure and indications of epidemiological studies identified

Hormone	Indication	Number of studies
EE	Prevention of threatened abortion	4
EE + different progestogens[1]	Diagnosing pregnancy	2
EE + different progestogens[2]	Oral contraception	1
NETA	Prevention of threatened abortion	10
NETA + EE	Diagnosing pregnancy	2

[1] Excluding NETA
[2] Including unknown number with NETA

6.3.2.1 Norethisterone acetate and ethinylestradiol for diagnosing pregnancy

Four studies investigating the use of NETA and/or EE for diagnosing pregnancy were identified of which only two studies specified the use of Primodos or NETA and EE (Higgins, 1960; Kullander, 1976).

The study by Higgins (1960) investigated the efficacy and safety of Primodos in a case series of 59 women with amenorrhoea of short duration. The first twelve women were given a four-tablet test (NETA 10 mg and EE 0.05 mg) and the remaining 47 patients were given a two-tablet test (NETA 5 mg and EE 0.01 mg). The abortion rate was one out of seven pregnant patients given the four-tablet test (14.3%) and 3 out of 36 patients given the two-tablet test (8.3%). The authors note that the background incidence of abortion was 20% of live births and therefore the incidence in this case series was not high, concluding that the abortions could not be attributed definitively to the use of the test tablets. The lack of a control group in this study made the strength of any potential association difficult to determine.

The study by Kullander was large and prospective and primarily investigated the role of hormonal drugs in fetal malformation. Primodos was one of the hormonal drugs under investigation and its results were reported separately from the other hormones. In this study, 156 of 6 376 (2.5%) pregnancies were exposed to Primodos in the second month of pregnancy and the miscarriage rate was 15 out of 156 (9.6%) compared with a rate of 448 out of 6376 (7.0%) overall. The women in the comparator group were largely untreated (although some were treated with other hormones for threatened abortion). Analyses were not adjusted for confounding factors. Of the women who used Primodos 8.3% went on to have an induced abortion compared with 2.3% of women in the comparator group.

6.3.2.2 Norethisterone acetate or ethinylestradiol for threatened or recurrent abortion

In many of these studies very high doses of NETA or EE hormones were given – in one case up to 1 000 mg NETA per day (Nygren, 1975), for prolonged periods, sometimes throughout pregnancy (eg. Abramson, 1958; Hodgkinson, 1958; Thierstein, 1959).

All but one of these studies had a prospective design thus removing the risk of recall bias or exposure misclassification. Limitations included a lack of clarity of reporting of the reason/criteria for determining which patients received the different treatments, giving rise to a risk of differential baseline risks of spontaneous abortion between treatment groups. The choice of control group was also unclear or not appropriate in many cases. Some studies lacked a control group completely making it difficult to put the reported abortion rates into context; others used a comparator group treated with different hormonal preparations which reduces selection bias but may have been unable to detect an increased risk if the same risk also existed for the comparator treatment. One study included a comparator group that received different hormonal preparations and another group that received no treatment (Hodgkinson, 1958) and showed the rate of pregnancy salvage (ie prevented from

miscarriage) to be 42.3% in the NETA treated group compared with 31% in the 17-alpha-hydroxyprogesterone caproate treated group and 15.5% in the untreated group.

6.3.2.3 Norethisterone acetate alone

One small case series in nine women with a threatened abortion or poor obstetric history reported an unusually high rate of abortion in the women treated with NETA (55%) but had no control group making this finding difficult to interpret. Nine studies found favourable pregnancy salvage rates with NETA or abortion rates consistent with published background rates.

6.3.2.4 Ethinylestradiol alone

Four studies concerned the use of EE alone during pregnancy. All were investigating the effectiveness of treatment for threatened abortion/recurrent miscarriage. No evidence for an abortifacient effect was found, with the authors reporting favourable pregnancy salvage rates with EE.

6.3.3 Conclusion on epidemiological data

The available epidemiological data were limited and provide no evidence for an abortifacient effect of NETA and/or EE when given to pregnant women.

6.4 Overall conclusion on the evidence for an abortifacient effect of HPTs in early pregnancy

While administration of ethinylestradiol and norethisterone, mostly at very high doses or for prolonged periods, can result in embryo-lethal effects in animals, there was no evidence that administration of these hormones at the licensed doses used in Primodos during early pregnancy were associated with an increased risk of miscarriage.

On the basis of their findings, the EWG considered that making electronic Yellow Card reporting available at point of care, including at scanning in early pregnancy, to all those who suspect an adverse outcome of pregnancy, including miscarriage or abortion, in association with exposure to any medicine would be valuable if sensitively designed.

7. APPROACHES TO IDENTIFYING, EVALUATING AND COMMUNICATING DRUG SAFETY CONCERNS IN PREGNANCY

This chapter discusses the developments that have taken place in identifying, assessing and communicating safety concerns with medicines and highlights areas that could be further strengthened to safeguard future generations, with a particular focus on the use of medicines in pregnancy. More specifically this chapter:

- looks at the processes and tools that were available when HPTs were on the UK market
- describes the changes that have been made since then
- establishes what measures are currently in place to identify, evaluate and communicate risks when medicines are taken during pregnancy
- considers whether any areas could be further improved to support the safer use of medicines in pregnancy.

Further information is provided in Annex 31.

7.1 Background

Since HPTs were available in the UK, the relationship between healthcare professionals and patients has changed substantially and there is an extensive network of support available for pregnant women. Chapter 2 summarises the socio-medical and regulatory environment within which HPTs were developed and prescribed.

In the 1950s and 1960s, official guidance on testing of drugs was very limited. The tragedy surrounding thalidomide ensured that both legislators and regulators tightened requirements surrounding the surveillance and approval process for drugs to be sold, requiring that manufacturers establish they are both adequately safe and effective before they are marketed. Now, drug development can take between eight and twelve years and involve animal testing and tightly regulated human clinical trials.

Establishment of legislation and a regulatory framework for the authorisation and safety monitoring of medicines within the UK is described in Section 2.1.2. In more rIt years the most significant update relating to patient safety took place in 2010 with new EU legislation on pharmacovigilance throughout Europe[18]. This revision came into effect in July 2012 and its purpose was largely to address the limitations that had been identified within the existing legislation. More specifically it sought to:

- increase proactive safety monitoring
- accelerate decision-making on urgent drug safety issues
- establish a European medicines safety committee (the Pharmacovigilance Risk Assessment Committee, PRAC) with legal powers
- improve co-ordination of the communication in member states of drug safety issues
- increase transparency and openness
- include patients in decision-making

[18] Directive 2010/84/EU, Regulation No. 1235/2010 and Implementing Regulation No 520/2012, referred to in the report as the 2010 Pharmacovigilance legislation

The 2010 Pharmacovigilance legislation is underpinned by a set of detailed guidance documents on all aspects of pharmacovigilance for regulatory authorities and MA holders – the Good Pharmacovigilance Practice (GVP) guidance. The 2010 Pharmacovigilance legislation is implemented in the UK in the Human Medicines Regulations 2012.

7.2 Collecting information on safety in pregnancy before a medicine is licensed

Women take medicines in pregnancy for a number of reasons: they may have an existing condition that requires ongoing treatment (eg. epilepsy, diabetes, HIV etc); they may develop a condition during pregnancy that requires treatment (eg. gestational diabetes, infections etc); or they may be inadvertently exposed, particularly in the critical early stages when they may not be aware they are pregnant and so continue to take a medicine. It is therefore essential that any risk from a medicine to the fetus or pregnancy is identified at an early stage of product development – and preferably before the medicine is used in humans.

7.2.1 Pre-clinical/non-clinical requirements

Before a medicine can be used in humans a series of studies must be carried out by pharmaceutical companies to identify any undesirable properties that may have relevance to humans. In 2001, all existing legislation on the conduct of pre-clinical and clinical studies was consolidated into Directive 2001/83, Annex 1 of which specifies exactly what pharmaceutical companies must provide as part of a valid application to market a new medicine (Annex 1 to Directive 2001/83/EC: Analytical, pharmaco-toxicological and clinical standards and protocols in respect of the testing of medicinal products[19]). Amendments to the Directive were made in 1999 and 2003.

7.2.1.1 Safety pharmacology studies

To obtain a licence to market a medicine pharmaceutical companies are required to demonstrate the pharmacological properties of a medicine, in both qualitative and quantitative relationship to its proposed use in humans. The measurements enable the calculation of safety margins between the doses at which any toxicity is noted in animals and the doses to be used therapeutically in humans. The specific studies required, and their design, will vary according to the individual properties and intended uses of the medicine but scientifically valid, internationally recognised methods must be used.

In recognition of the importance of good quality, consistent pharmacokinetic measurements in evaluating systemic exposure during drug development a number of guidelines have been developed by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use. However, there are currently no requirements to conduct pharmacokinetic/pharmacodynamics studies that will enable a better understanding of how pregnancy affects the levels of drug to which the mother and fetus are exposed and to develop evidence-based posology for pregnancy.

7.2.1.2 Requirements for studies in animals

Animal studies generally provide the first source of information about potential safety concerns with a medicine. In 1971 the Medicines Act 1968 (c.67) introduced a general requirement for pharmaceutical companies to conduct toxicology testing in animal models but there was very little specific guidance on precisely what studies were required to provide sufficient evidence of safety in pregnancy. The design and quality of any studies that were

[19] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0083&from=en

carried out when HPTs were developed was therefore inconsistent and not compliant with today's standards.

In relation to reproductive toxicity specifically, data on parameters now considered to be important were not routinely collected or reported and far fewer animals were exposed than are specified today – a clear limitation in detecting infrequent adverse events. In addition, there was a lack of knowledge on how to determine the human relevance of any findings in animals. In recognition of this, the ICH produced a comprehensive set of safety guidelines on how to uncover a potential risk of carcinogenicity, genotoxicity or reproductive toxicity with a medicine. To assure the quality of the data generated in support of the safety of new medicines, animal studies must be conducted in accordance with Good Laboratory Practice[20].

Despite the current legislation and guidelines, it can be difficult to predict whether safety observations made in animals will have relevance to humans – thalidomide is a good example where negative limited animal data did not translate to an absence of effect in humans. The types of studies now required and the use of both rodent and non-rodent species to investigate the potential effects of the drugs in the three segments of reproductive toxicology (fertility and embryonic development, embryo-fetal development and pre- and post-natal development) does allow for any adverse effects that could occur to be fully investigated and how this data would extrapolate to humans.

7.2.1.3 Requirements for studies in humans

Randomised controlled trials in humans are required to establish the efficacy and safety of a new medicine before it can receive a marketing authorisation.

For many years, no women (of any reproductive status) were included in clinical trials of medicines, a decision based on the experience with thalidomide. Clinical research was conducted mostly in men until 1997 when the ICH E8 European guideline "General Considerations for Clinical Trials" allowed for women to participate as long as those of childbearing potential used highly effective contraception. While it is recognised that there may be advantages in including women who are pregnant in trials, there have been no initiatives to encourage their recruitment because of the potential dangers. If, however, the product is intended specifically for use during pregnancy, a trial in pregnant women is required, with those taking part being provided with information on the potential risks before giving consent and with subsequent follow-up of the pregnancy, fetus, and child.

For products that are not specifically indicated in pregnancy but are expected to be used by pregnant women, a randomised controlled trial is not required but safety data would generally be collected after licensing through a post authorisation safety study.

7.2.2 Anticipating safety concerns

These days, at the time of licensing a considerable amount of information on the safety of a medicine is available from the pre-clinical studies and randomised controlled trials as described above. These data can be used to determine what is known about the safety of a medicine, anticipate what is not yet known, decide what concerns may need to be studied further (either before a licence can be granted or after approval, depending on the nature of the concern) and determine whether any action is required to minimise harm to patients. These four aspects are formally considered during the development of a Risk Management Plan (RMP), which pharmaceutical companies are legally obliged to submit with every marketing authorisation application. The RMP must be assessed and approved by the regulatory authority before the medicine can be granted a licence.

[20] http://www.legislation.gov.uk/uksi/1999/3106/contents/made

For products that are likely to be used in pregnancy or in women of child-bearing potential, the RMP must include information on teratogenicity findings in animals together with the outcomes of any pregnancies that occurred in the clinical development programme to ensure that any potential concerns are further studied and measures are put in place to minimise risk to the mother or fetus.

7.3 Collecting information on safety of medicines after licensing

To enable safety signals that arise once a medicine is on the market, and therefore being used in greater numbers or by those with risk factors, to be identified and evaluated further it is necessary to have some means of capturing accurately information on exposure to medicines, pregnancy status and adverse effects. The best way to collect such information is through prospective population-based pregnancy registers. Such registries, for example in Scandinavia, are typically controlled centrally, register all births, can often be linked to other population-wide databases, including those collecting data on congenital anomalies. No such population-based register currently exists in the UK; instead many individual databases exist that each record different elements of these data. The fragmented nature of these data makes it difficult to conduct large, high quality studies of adverse outcomes of pregnancy in women who are taking a medicine; for this to be possible, improved data linkage between the existing datasets would be required. The following sections outline other methods of data collection that exist in the UK.

7.3.1.1 Spontaneous reporting of suspected adverse drug reactions

Before the introduction of the Yellow Card Scheme (YCS) in the UK in 1964, there was no requirement to report or collect cases of suspected ADRs to licensed medicines. Today pharmaceutical companies are legally required to operate a system for recording suspected ADRs in association with their medicines and to report these to national regulatory authorities within strict timeframes. Suspected ADRs can also be reported by patients and healthcare professionals. Cases that report an abnormal outcome of pregnancy are classified as serious and those received by pharmaceutical companies must be expedited to the regulators. In total, there are about 40 000 ADR reports per year and the Yellow Card database now holds about 850 000 reports.

Signal detection via spontaneous ADR data has more recently been carried out using statistical software to flag any combinations of a medicine and suspected adverse event that have been reported disproportionately, relative to the rest of the reports on the database. However, despite a great many improvements to the YCS there is likely to be substantial under-reporting of adverse outcomes of pregnancy in association with a medicine. In part, this is because many months may have elapsed between taking the medicine and the observed adverse outcome, reducing the chances of a causal association being suspected, even more likely for developmental disorders that only become apparent in early childhood.

It is also more difficult to detect safety signals for conditions that occur relatively frequently in the general population, such as some of the more common congenital anomalies, and reported cases typically contain insufficient information to assess the likelihood of a possible causal association. Thus, details on the gestation period of fetal exposure, other medicines taken during pregnancy, other relevant lifestyle factors and maternal history are often lacking and positive de-challenge (ie resolution of the ADR on stopping the medicine) and re-challenge (reoccurrence of the ADR on re-starting the medicine), which are usually key criteria for assessing causality in standard ADR reports, do not apply to reports in pregnancy.

Within Europe the PROTECT project was established to develop a set of innovative tools and methods to enhance the early detection and assessment of ADRs from a range of

different data sources including spontaneous ADR reports, registries, clinical trials and other electronic health records. PROTECT also aimed to develop methods to combine the results from clinical trials, spontaneous reporting and observational data.

7.3.1.2 Systematic collection of information

- Congenital anomaly databases

In 1964, in response to the thalidomide tragedy, the National Congenital Anomaly System (NCAS) was set up to report congenital anomalies identified within seven days of birth in England and Wales and to monitor changes in reporting (Weatherall, 1978). NCAS operated until 2010. Compared with other congenital anomaly registers that were set up in defined geographical regions to establish the absolute incidence of congenital anomalies in that area, the NCAS surveillance system under-estimated the incidence of congenital anomalies, despite many attempts made to improve it (Botting, 2003). An in-depth review of the system in 1995[21] recommended the sharing of data from regional congenital anomaly registers with NCAS and in 1996 a network of the existing registers was established, the British and Irish Network Of Congenital Anomaly Registers (BINOCAR), whose role was to monitor and analyse congenital anomalies. In England, these databases covered only 40% of the population, but the data collected was considered good quality, representative of the UK population and was shared anonymously with many European registers through EUROCAT. BINOCAR provided good prevalence data for congenital anomalies and was scrutinised regularly by experts but was unable to provide accurate or detailed information on exposure to medicines in pregnancy and so was of limited use for the detection of safety signals with medicines. BINOCAR has now largely been replaced by the National Congenital Anomaly and Rare Disease Registration Service (NCARDRS).

NCARDRS is a national congenital anomalies system for England. Established by Public Health England in 2015 it incorporated the seven pre-existing congenital anomalies registers into one system. NCARDRS has established teams in areas which did not previously have registers and uses multiple sources of ascertainment; however, reporting of anomalies is not mandatory. While most notifications are received perinatally (including electronic notifications from obstetric ultrasound services), neonatally or antenatally from secondary care (predominantly midwives) data can also be collected postnatally including from children's clinics. NCARDRS is at an early stage in its development and its data are not yet sufficiently complete for data linkage. Like its predecessor registers, NCARDRS contributes data to EUROCAT.

Public Health Wales operates the Congenital Anomaly Register and Information Service (CARIS). CARIS started in 1998 and collects information about any fetus or baby who has, or is suspected of having, a congenital anomaly at any age of gestation to at least the end of the first year of life. It uses multiple sources of ascertainment, including antenatal ultrasound reports, although many of its notifications come from clinicians. CARIS can also download information on cytogenetics, paediatric cardiology and medical genetics from other databases. CARIS also collects information from medical records on medication use of the mother.

In Scotland, the Scottish Congenital Anomalies Register (SCAR) registered all congenital anomalies detected at birth or during infancy between 1988 and 1996. Data obtained from neonatal returns, the Stillbirth and Infant Death Survey and acute hospital admissions in the first year of life could be linked to create a single baby record. In recent years, the Information Services Division (part of NHS Scotland) has been working to develop a new

[21] Working EWG of the Registrar General's Medical Advisory Committee. The OPCS monitoring system: a review by the Working EWG. OPCS Occasional Paper No. 43. London: OPCS, 1995

anomaly database, the Scottish Linked Congenital Anomaly Database, along the same lines as SCAR.

There is currently no national congenital anomaly database in Northern Ireland; however, proposals for a national register were published in a report in June 2017[22].

Outside of the UK, the International Clearinghouse for Birth Defects Surveillance and Research brings together congenital anomaly programmes from around the world with the aim of conducting worldwide surveillance and research; and in 1979 EUROCAT (European Surveillance of Congenital Anomalies) was established to set common standards for data collections and reporting for registers in the European Union and to pool the data from them. EUROCAT currently includes more than 1.6 million births per years from 43 registries in 23 different countries. While the registries capture detailed and extensive outcome information on the types of anomaly, information on medicine exposure is limited. To address this, the EUROmediCAT project sought to make more systematic use of electronic healthcare databases and exposed pregnancy cohorts in combination with EUROCAT congenital anomaly data. The funded project completed in 2015[23], but the database continues to be used.

- Pregnancy registers

More systematic approaches to collection of data can also be achieved through targeted pregnancy registers. These can be set up by pharmaceutical companies, academia or clinicians to monitor use of a specific drug substance in pregnancy, or to follow pregnant women with a specific medical condition. However, recruitment of women is voluntary and so can be poor – the size of cohort established (and therefore the size of any risk that can be detected) will depend on how commonly the medicine is used in pregnancy and the willingness of women and health professionals to be involved. The accuracy and completeness of the information collected depends on whether there is access to the medical records of the mother and neonate but experience has shown that these datasets are often incomplete or the information is poorly or inconsistently recorded.

In addition, women and healthcare professionals are encouraged to report their exposure to medicines in pregnancy by registering with the 'BUMPS' programme of the UK Teratology Information Service (UKTIS). This is an entirely voluntary nationwide service which collates and reviews periodically the information received for each medicine and any pregnancy outcomes that might be reported. UKTIS is a founder member of an international network of teratology information services which provide patient specific advice and evidence based information to women and health providers, and conduct teratogen surveillance by following up for information on timing of exposure, dose, other medicines, smoking status, alcohol intake and maternal health. Information on the outcomes of pregnancies is now collected at birth and in the longer-term.

- Electronic healthcare databases

Within the UK a number of large electronic healthcare databases routinely capture health information from primary care records. Participating GP centres provide access to their electronic medical records which include exposure to prescribed medicines, pregnancy status, and medical conditions and diseases. These databases have some limitations in that they do not necessarily record the outcome of pregnancy, they may not automatically link a mother with her child, and they do not include information on medicines purchased over the counter. Examples of electronic healthcare databases include:

[22] http://uir.ulster.ac.uk/37141/1/NICAREport.pdf
[23] http://euromedicat.eu/home

- Clinical Practice Research Datalink (CPRD). Established in 1987, CPRD is an ongoing governmental primary care database of anonymised, computerised longitudinal records of patients' GP consultations and treatment from over 15 million patients (approximately 8% of the population) in the UK. Data available includes basic patient and practice demographics, medical symptoms and diagnoses, hospital referrals and in-patient and out-patient consultations, tests, drugs prescribed and immunisations. CPRD has developed a Mother-Baby Link (based on an algorithm using primary care data) with data on about 1.1 million mother-baby pairs and 780 000 mothers in the UK; however, this looks at live births only and mothers are only included if a link with her baby has been established.
- THIN: this database collects information on prescriptions (date, formulation, strength, quantity, dosing instructions, indication and events leading to stopping of a medicine) from participating GPs throughout the UK. It also collects information on pregnancy and birth and can identify mothers and children living at the same or different addresses. It currently contains data on over 10 million individuals in the UK.
- Secure Anonymised Information Linkage (SAIL). SAIL collects routine healthcare data from a sample of GPs in Wales. It currently has data linkage with the National Community Child Health Database for Wales (live and stillbirth pregnancies) and Patient Episode Database for Wales (pregnancy loss) and can be linked with the Wales Congenital Anomaly Registry (CARIS). Details on pregnancy and outcome such as date of LMP, birth weight, gestational age, congenital anomalies and developmental delay are captured.

Separate electronic databases capture information on all prescribed medicines that are dispensed to patients. These include the NHS Prescription Cost Analysis (PCA) database which contains details of all NHS prescriptions dispensed in retail pharmacies in England; the ISD Scotland Prescription Cost Analysis which similarly shows details of all NHS prescriptions that are dispensed in the Scottish community; and the HSC Business Services organisation which provides the same service for Northern Ireland. IMS MIDAS captures information on the volume of prescription drugs dispensed in hospital, as well as retail, pharmacies in the UK. Indicators that enable patients to be identified are not captured by any of these databases (except in Northern Ireland) and so they cannot be used to determine who may be pregnant when a medicine is dispensed.

7.4 Evaluating potential safety signals

7.4.1 Key principles of signal evaluation

The signal for a possible association between HPTs and spina bifida initially arose from a case control study (Gal 1967). Historical records suggest that the following steps were taken by the CSD or CSM to further evaluate this potential signal:

- discussion of the study findings with the authors
- request for manufacturers of HPTs to provide all relevant laboratory data
- request for relevant information from academics working in the field
- exploration of possible collaboration on studies with those working with congenital anomaly databases
- initiation of a long-term questionnaire study: "Maternal drug histories in babies with congenital abnormality", to examine a possible association between HPTs and cleft palate/hare lip, spina bifida and hydrocephalus and reduction deformities of limbs
- regular consultation with expert committees as new data emerged.

Based on regular evaluation of the data that was accruing, scientific uncertainty over the strength of the evidence, and the development of better pregnancy tests, the following precautionary actions were taken between 1967 (when the Gal study was published) and 1978 (when Schering withdrew Primodos from the UK market). The indication 'diagnosis of pregnancy' was removed from the Primodos datasheet by Schering in 1970, and around the same time the company stopped promoting Primodos for pregnancy testing and stopped providing free samples to healthcare professionals[24]. When the interim results of the CSM study became available in 1975 the Primodos product information was updated to include a warning about the possible risk of congenital anomaly and a contraindication in pregnancy. In 1975 prescribers were issued a warning that HPTs should not be used to diagnose pregnancy followed by a reminder in 1977. Primodos was withdrawn by Schering in 1978.

The key principles of signal evaluation have not changed fundamentally since HPTs were on the market and all the above actions with respect to data gathering, seeking expert advice and taking precautionary action would still be undertaken today. What is different is that specific pharmacovigilance legislation is now in place together with detailed guidance available to support rapid, co-ordinated, robust and impartial decision making.

7.4.1.1 Timeliness

- Ensuring action is consistent and timely

In their handling of the safety signal for HPTs, the CSD/CSM were criticised for taking too long to withdraw Primodos from the market and, in the different European countries and globally, decisions to withdraw HPT products were taken in a staggered and uncoordinated way. This is likely to have been due to several factors. Firstly, there were differences of opinion on the strength of the evidence for an association between HPTs and congenital anomalies; secondly communication channels between regulators in different countries were poorly developed at that time; and thirdly many similar HPT preparations were available (often with different names but with the same constituents).

To ensure important decisions made on drug safety issues today are timely, consistent across member states and legally binding on pharmaceutical companies, the current 2010 Pharmacovigilance legislation has strengthened coordination of any regulatory action taken for safety reasons within any country in the EU and sets out timelines for action that are consistent with the importance of the concern.

To keep abreast of global safety issues, important safety communications and actions taken on safety issues are routinely shared with regulators in other territories such as the US, Canada and Japan.

- Robustness of decision-making

Pharmacovigilance relies on judgements being made on risk, benefit and the benefit:risk balance, frequently based on data that derive from different data sources, or that may be sparse and/or subject to limitations. In recognition of this, steps have been taken at a national and EU level to ensure that decisions made on drug safety are not only timely and co-ordinated but robust and stand up to scrutiny. This means comprehensive training of assessors and quality assurance processes to ensure appropriate levels of review and sign off. For issues of public health importance where regulatory action seeks to change prescribing behaviour, independent expert advice is sought. In Europe, consideration of

[24] Today, there are strict requirements for the supply of free samples of medicines to prescribers, as set out in section 6.12 of the MHRA Blue Guide.

safety concerns by all member states and by the Pharmacovigilance Risk Assessment Committee (PRAC) ensures that any country-specific influences are minimised.

In terms of improving data quality, all aspects of epidemiological research have developed greatly since HPTs were on the UK market, but especially with respect to the recording and linking of patient records and the recognition and appropriate handling of various biases. Datasets on congenital anomalies are also now likely to be considerably 'cleaner' than before because significant advances in genetic testing has enabled adverse outcomes caused by genetic defects to be identified. Other improvements in the quality of epidemiology studies may be due in part to international initiatives to improve the design, conduct and reporting of observational studies, through initiatives such as those implemented by STROBE (Strengthening the Reporting of Observational studies in Epidemiology) and ENCePP (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance).

Since 2010 regulators have had the authority to ask pharmaceutical companies to conduct a post-authorisation safety study (PASS) whenever a potential safety concern with a medicine is identified, which may include a pregnancy register. This can be at the time of licensing when further characterisation of a risk is considered necessary or once the medicine is on the market in response to a new safety concern. If the PASS is imposed as a 'condition' of the marketing authorisation, the 2010 Pharmacovigilance legislation provides for possible suspension of the licence if it is not conducted. For products that are not specifically indicated in pregnancy but where exposure may occur, a PASS study would usually be required at the time of licensing.

- Transparency

A number of initiatives have been introduced in the UK to improve the transparency of Government, including the Freedom of Information Act (FOI) in 2000. Against this, Section 43 of the FOI Act provides an exemption from disclosure for information that is commercially sensitive. However, where the balance of public interest falls in favour of disclosure (ie the interest in maintaining the exemption does not outweigh the interest in disclosing the information) the Government department or other public authority is obliged to disclose the information upon request, for example where the information concerns a serious risk to public health or safety.

In the UK efforts have been made to provide more information on regulatory decisions taken throughout the lifecycle of a product and information on the safety profiles of medicines. These include: public scientific assessment reports for all licensed products granted a new national product licence, the prescribing information for all licensed products in the UK, the agendas and minutes of all expert committee meetings and ADR data received through the Yellow Card Scheme.

Within Europe the EMA publishes public scientific assessment reports for all medicines that are licensed through the European centralised system, the agendas and minutes of the monthly PRAC meetings, high-level outcomes of the PRAC's main scientific discussions and the PRAC's formal Recommendations and Advice.

7.5 Managing risk in pregnancy

7.5.1 Guidance for safe use of a medicine in pregnancy

To make sure a consistent approach to reducing potential harm in pregnancy is taken, based on the available evidence, regulators and pharmaceutical companies refer to a European guideline developed in 2008 ("Guideline on risk assessment of medicinal products on human

reproduction and lactation: From data to labelling"). This specifies the most appropriate risk minimisation measure(s) to be taken by prescribers and women to ensure safe use of the medicine (table 21). On this basis, the most appropriate risk minimisation measure for HPTs, for which there are considered to be limited data in pregnant women and animal studies that do not indicate direct or indirect harmful effects with respect to reproductive toxicity, would be to avoid the use of the medicine during pregnancy as a precautionary measure.

In addition to the measures outlined above for the SmPC and PIL, a range of additional regulatory measures can be recommended or imposed to minimise risk to the fetus, depending on the level of risk. These include:

- the development of educational materials for healthcare professionals and women
- requirement for proof of a current negative pregnancy test prior to receiving medicine
- implementation of a formal Pregnancy Prevention Plan.

7.5.2 Pregnancy Prevention Plans (PPPs)

PPPs are best described as a set of interventions aiming to minimise exposure during pregnancy to a medicine that is critical for use but that has known or potential teratogenic effects. Its aim is to ensure that women are not pregnant when starting therapy and do not become pregnant during treatment or soon after stopping. PPPs can also target men if use of a medicinal product by the father could have a negative effect on pregnancy outcome.

As an example of a set of measures to minimise risk to a pregnancy, a medicine containing thalidomide was licensed in Europe in 2008 for the treatment of multiple myeloma despite its well-recognised teratogenic effects in humans. Myeloma is typically a disease of those older than 45 years, however, for the marketing authorisation to be approved a strict set of criteria for gaining access to treatment were put in place at the time of licensing. These included:

- implementation of a PPP, including the need to show a negative, medically supervised pregnancy test before treatment, every four weeks on treatment and four weeks after the end of treatment
- a maximum of four weeks of treatment per prescription
- strict contraceptive requirements for women of childbearing potential before during and after treatment and for men use of a condom during intercourse throughout treatment and for one week after stopping
- a contraindication in pregnancy, and a contraindication in women of childbearing potential – unless all the conditions of the PPP were met
- a contraindication in patients unable to follow or comply with the required contraceptive measures
- boxed warnings in product information about the teratogenic effects of thalidomide
- circulation of a letter and provision of an Education Kit to prescribers and an education booklet and card for patients, reviewed by national patient organisations/groups
- provision of reporting forms in the event of a pregnancy.

In addition to preventing harm PPPs can also be used to systematically collect pregnancy outcome information to help further characterise a risk or to assess the effectiveness of the risk minimisation measures that have been imposed. However, if the PPP is working as intended very little data would be expected to be collected.

Table 21. Options for minimising harm in pregnancy (as specified in the "Guideline on risk assessment of medicinal products on human reproduction and lactation: From data to labelling").

Option	Strength of evidence	Recommended risk minimisation measure
1	Data in humans demonstrates that a medicine causes congenital malformations or other harmful effects.	Contraindication for the whole, or a defined period of, pregnancy Women of child-bearing potential have to use effective contraception during (and sometimes after) treatment.
2	Data in humans suggests that a medicine causes congenital malformations or animal data have shown reproductive toxicity or are insufficient with respect to reproductive toxicity.	The medicine should not be used during pregnancy unless treatment is critical; women of child bearing potential have to use effective contraception
3	Data in humans suggests that a medicine causes congenital malformations and animal data do not indicate direct or indirect reproductive toxicity.	The medicine should not be used during pregnancy unless treatment is critical; women of child bearing potential have to use effective contraception
4	No data or limited data in humans but studies in animals have either shown reproductive toxicity or are insufficient in this respect.	The medicine is not recommended during pregnancy or in women of childbearing potential unless using contraception
5	There are no or limited data (less than 300 pregnancy outcomes) in pregnant women and animal studies do not indicate direct or indirect harmful effects with respect to reproductive toxicity.	As a precautionary measure, it is preferable to avoid the use of the medicine during pregnancy
6	A moderate amount of data on pregnant women (between 300-1000 pregnancy outcomes) indicate no malformative or feto/neonatal toxicity of the medicine but animal studies have shown reproductive toxicity or are insufficient with respect to reproductive toxicity	As a precautionary measure, it is preferable to avoid use of the medicine during pregnancy
7	A moderate amount of data on pregnant women (between 300-1000 pregnancy outcomes) indicate no malformative or feto/neonatal toxicity and animal studies do not indicate reproductive toxicity	The use of the medicine may be considered during pregnancy, if necessary.
8	A large amount of data on pregnant women (more than 1000 exposed outcomes) indicate no malformative nor feto/neonatal toxicity.	The medicine can be used during pregnancy if clinically needed
9	No effects during pregnancy are anticipated, since systemic exposure to the medicine is negligible.	The medicine can be used during pregnancy

7.6 Communication of risk

The role of the regulator is to ensure that the marketing authorisation for a medicine, as described in the SmPC and PIL, reflects the available data and outlines the terms under which the balance of benefits and risks of a medicine is positive. When new data emerge that require a change to the regulatory position, this needs to be communicated to healthcare professionals and patients in a clear and targeted way. Standards set by the UK's General Medical Council (GMC) state that doctors must keep up to date with statutory information and advice.

When Schering removed the indication of Primodos as a hormone pregnancy test from its data sheet in 1970 one criticism was that this change was not actively communicated and that prescribing behaviour remained relatively unchanged. With the publication of more studies and uncertainty over the evidence, in 1975 CSM issued a precautionary warning to all prescribers advising them not to use hormonal tests for diagnosing pregnancy because of a possible risk and the increasing availability of other means of diagnosing pregnancy. Prescribing was subsequently shown to be reduced by around 60%. Then in 1977, reports that Primodos was still being used as a pregnancy test prompted CSM to issue a reminder that these products should not be used for this purpose and prescribing was shown to be reduced by a further 30%. While the temporal relationship between the CSM issuing its two warning letters to doctors in 1975 and 1977 and the observed fall in prescribing does not prove that the two events were causally associated it is a strong indication that they may have been and demonstrates the importance of effective communication.

In 1970 the absence of a legal framework for regulation of medicines, meant that pharmaceutical companies (or regulators) did not have to communicate changes to the data sheet and there was no requirement to provide information to patients. Today, a patient's need to be fully informed about their medicine is underpinned by legislation, and PILs – which accompany every medicine and contain information on all recognised adverse effects of medicines – have been a legal requirement for all medicines since 1999. The legislation[25] also makes it clear that pharmaceutical companies should communicate important safety information relating to their medicine and that consistent and co-ordinated messages should be provided throughout Europe. This is primarily achieved through circulation of Dear Healthcare Professional Communications (DHPC).

A number of advances with respect to how drug safety messages are communicated to healthcare professionals, patients and the public, have been made over the years. When there is an urgent need to deliver messages to healthcare professionals a web-based cascading system for issuing important safety messages to the NHS and independent providers of health and social care is used. This can be supplemented by a press release or press briefing, website information, use of digital and social media, or collaboration with the relevant Professional Societies to help disseminate the message to a targeted audience. An online monthly drug safety bulletin, Drug Safety Update (DSU) for healthcare professionals, has been published since 2007. Before this 'Current Problems in Pharmacovigilance' was circulated.

7.6.1 Patient engagement

Since 1999 there have been a number of initiatives to better engage with patients including involving patient representatives in regulatory decision making, to ensure the patient "voice" is given equal consideration to that of the healthcare professionals (since 2014).

[25] Commission Implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council.

A Patient EWG Consultative Forum of individuals from 50 different patient groups and research charities has been established to provide the patient perspective to understanding of a range of issues including attitudes to benefit/risk in medicines innovation and practical issues encountered by patients with regard to the packaging of medicines. More use is being made of social media, including highlighting key concerns – such as falsified medicines, through soap operas to better engage with patients and the public.

7.7 Measuring the effectiveness of regulatory action

Even after the CSM communicated the advice on HPTs to healthcare professionals in 1975 and 1977, the restriction of the indication to secondary amenorrhoea was not adhered to. The lack of effectiveness of the indication restriction in 1970 was only identified retrospectively, by which time Primodos had continued to be used to diagnose pregnancy by some for as much as eight more years. Today, pharmaceutical companies may be required to evaluate whether action taken to minimise an important risk has had the desired effect and if not why not. If the results suggest that the action has not achieved its goal, further action should be considered and its effectiveness again evaluated. When regulatory options have been shown to be insufficient in addressing the risk, consideration is given to whether there is a need to suspend or remove the product from the market.

There may be a need to conduct additional research, making particular use of the data available within the CPRD. Where advice on prescribing practice is shown not to have been followed this may be because, despite communication, doctors remain unaware of the new information. Alternatively, they may be aware but not persuaded to change their prescribing behaviour. Reason(s) why advice is not being followed should be further evaluated, and consideration given to whether different, potentially more restrictive, regulatory measures are required.

7.8 Key observations

Despite major improvements in medical and regulatory practice, the current systems in place for pharmacovigilance and pharmacoepidemiology may have limitations for identifying and managing risks relating to medicines used in pregnant women, as follows:

- Predicting whether safety observations made in animals will have relevance to human beings is challenging. The investigations and the interpretation of the results from animal studies should be related to all other pharmacological and toxicological data available to determine whether potential reproductive risks to humans are greater, lesser or equal to those posed by other toxicological manifestations.
- Significant changes in physiology take place during pregnancy but there are few data on how this may affect the absorption, distribution, metabolism and excretion of medicines.
- Randomised controlled trials remain the best way to assign causality to any relationship observed between a medicine and an adverse outcome but pregnant women are generally excluded from trials unless the medicine is specifically indicated for use in pregnancy.
- Developments in the spontaneous reporting of suspected ADRs have taken place but there remain limitations which limit the ability to detect signals of potential harm due to medicines used in pregnancy. This is particularly true for effects that manifest some years after exposure.
- The UK has several national datasets/programmes for collecting information on: i) exposure to medicines in pregnancy and ii) neonatal outcomes. A key limitation is

the inability to perform more extensive linkage between the datasets and this makes it difficult to perform pharmacoepidemiology studies to detect or evaluate potential safety signals.

- A number of useful tools and measures exist to minimise the chance of harm through inappropriate use of a medicine in pregnancy. These could be implemented more systematically.
- Progress has been made with regard to monitoring the impact of regulatory action but more could be done to evaluate changes in the rates of clinically significant outcomes and understand what factors influence a health professional to change their behaviour.
- Communication of drug safety concerns by the regulator has improved; however, ensuring that safety messages reach the target audience and effect a change in their behaviour could be improved.

7.9 Conclusion

There have been substantial and far-reaching advances in all areas of the development, legislation, regulation, study and use of medicines since HPTs were available in the UK. Nevertheless, there are clear opportunities to strengthen further how safety concerns in pregnancy are detected, evaluated, managed and communicated (see section 8.2).

On this basis, the EWG recommended that:

1. A new Working Group should be set up to advise on better ways to collect and monitor data on the safety of medicines during pregnancy. The Working Group's remit should, in particular, explore the potential for:
    - better capturing and linking of existing data on adverse outcomes of pregnancy, including congenital anomalies identified prenatally and neonatally, and developmental disorders that take longer to become apparent, to facilitate regular surveillance
    - other ways to capture relevant information from, amongst others, midwives and pregnant women on exposure to all medicines, including prescription and over-the-counter, during a pregnancy
    - improving access to all relevant data on medicines taken during pregnancy to enable studies to be conducted to support pharmacovigilance
    - improving the analytic design of studies examining drug safety in pregnancy
    - a system for the early sharing and expert review of possible signals or concerns regarding teratogenicity of a drug
    - systematic, detailed clinical and genetic evaluation of patients in whom a teratogenic effect is being queried

2. Regulators should develop specific guidance for regulators and the pharmaceutical industry to i) strengthen the capture and evaluation of data on possible safety concerns with medicines used in pregnancy, and ii) support the more systematic use of measures to reduce harm from identified risks of medicines in pregnancy.

3. MHRA should systematically monitor outcomes after taking important regulatory action to protect patients from harm, and use this information to inform further action where necessary.

4. MHRA should work with the key information providers to ensure healthcare professionals and patients receive the best available information, and are empowered to make informed decisions and ask questions about any medicines they may be prescribed in pregnancy.

5. MHRA should do more to encourage and make it easier for women, and health professionals who work with women, to report any adverse reaction they experience while taking a medicine during pregnancy through the Yellow Card Scheme.

6. MHRA should build a partnership with other bodies within the healthcare system to improve the impact of safety messages relating to medicines, to support the objectives above.

8. THE WAY FORWARD

This review of HPTs and possible associated adverse outcomes in pregnancy has considered a wide range of data from mechanistic studies, meta-analyses, observational studies and spontaneous suspected ADR reports, including reports from patients, for an effect on the development of congenital anomalies or miscarriage. This has been the most comprehensive review of this subject to date.

This chapter summarises the key findings of the review and considers what further work would be desirable to improve the safe use of medicines in pregnancy.

8.1 Overall conclusions of the review

The EWG set out to address three key issues in relation to the evidence reviewed. The main findings of the review in considering these three areas are as follows:

1. **To consider all available evidence on a possible association between exposure in pregnancy to HPTs and adverse outcomes in pregnancy (in particular congenital anomalies, miscarriage and stillbirth) including consideration of any potential mechanism of action**

   The EWG's overall finding is that the available scientific evidence, taking all aspects into consideration, does not support a causal association between the use of HPTs, such as Primodos, during early pregnancy and adverse outcomes, either with regard to miscarriage, stillbirth or congenital anomalies. All the available relevant evidence on a possible association has been extensively and thoroughly reviewed with the benefit of up-to-date knowledge by experts from the relevant specialisms.

2. **On whether the Expert Working Group's findings have any implications for currently licensed medicines**

   The findings of the review for HPTs, including Primodos, on a possible association between exposure in pregnancy to HPTs and adverse outcomes in pregnancy do not have implications for any currently licensed medicines. They are in fact considered to be reassuring for women who may inadvertently become pregnant whilst taking these hormones for contraception or gynaecological indications.

3. **To draw any lessons for how drug safety issues in pregnancy are identified, assessed and communicated in the present regulatory system and how the effectiveness of risk management is monitored**

   There have been substantial and far-reaching advances in all areas of the development, regulation, study and use of medicines in pregnancy since HPTs were available in the UK, whereas there was a lack of transparency in the past. Nevertheless, ways to strengthen further how safety concerns in pregnancy are detected, evaluated, managed and communicated should be taken forward.

8.2 Recommendations of the EWG

The EWG noted that substantial changes have taken place within the field of pharmacovigilance and pharmacoepidemiology since HPTs were available in the UK but felt that more could be done to safeguard future generations. The EWG considered that a number of steps could be taken to strengthen the systems in place for detecting, evaluating, managing and communicating risk with exposure to medicines in early pregnancy.

For the families

1. A full up-to-date genetic clinical evaluation, in line with current best practice, should be offered to families of the Association for Children Damaged by HPTs, whose lives have been impacted by adverse pregnancy outcomes and who were given HPTs to diagnose pregnancy.

Optimising collection of, access to and use of data on medicines in pregnancy

2. A new Working Group should be set up to advise on better ways to collect and monitor data on the safety of medicines during pregnancy. The Working Group's remit should, in particular, explore the potential for:
    - better capturing and linking of existing data on adverse outcomes of pregnancy, including congenital anomalies identified prenatally and neonatally, and developmental disorders that take longer to become apparent, to facilitate regular surveillance
    - other ways to capture relevant information from, amongst others, midwives and pregnant women on exposure to all medicines, including prescription and over-the-counter, during a pregnancy
    - improving access to all relevant data on medicines taken during pregnancy to enable studies to be conducted to support pharmacovigilance
    - improving the analytic design of studies examining drug safety in pregnancy
    - a system for the early sharing and expert review of possible signals or concerns regarding teratogenicity of a drug
    - systematic, detailed clinical and genetic evaluation of patients in whom a teratogenic effect is being queried

3. Electronic Yellow Card reporting should be made available at point of care, including at scanning in early pregnancy, to all those who suspect an adverse outcome of pregnancy in association with exposure to any medicine in pregnancy. In particular, Yellow Card reporting should be included in relevant clinical systems and promoted in a dedicated campaign to raise awareness of this possibility.

4. There should be regular, independent review by experts of all suspected adverse drug reactions in pregnancy that are reported by healthcare professionals and women in the UK to the MHRA. The CHM should publish the findings and conclusions in their annual report.

5. A scientific workshop should be held to bring together different disciplines to consider:
    a) how results from studies in pregnant animals, with individual medicines or the chemical class, can be made more accessible in order to help predict and assess the potential effects of medicines in pregnancy

b) the feasibility of using computer modelling and molecular structure alerts to generate safety signals from animal and in vitro data and to prioritise drugs for further study.

6. A strategy to co-ordinate and promote research on the following should be taken forward with appropriate experts in the field:
    a) mechanisms of teratogenicity in early embryonic development and how the actions of and reactions to drugs vary with the individual's genes
    b) drug transporter expression in the placenta, particularly in early pregnancy; how it differs between individuals; and how it is affected by maternal disease.

Safeguarding future generations

7. For medicines used commonly in pregnancy, particularly the first trimester, pharmacokinetics and pharmacodynamics studies in pregnant women should be performed, where possible, to understand better how pregnancy affects the levels of drug to which the mother and fetus are exposed and to develop evidence-based dosing and frequency of administration for use in pregnancy.

8. In support of the above recommendation opportunities should be provided for obstetricians to receive training in pharmacology.

9. Regulators should develop specific guidance for regulators and the pharmaceutical industry to i) strengthen the capture and evaluation of data on possible safety concerns with medicines used in pregnancy, and ii) support the more systematic use of measures to reduce harm from identified risks of medicines in pregnancy.

10. MHRA should systematically monitor outcomes after taking important regulatory action to protect patients from harm from medicines, and use this information to inform further action where necessary.

Informing and engaging healthcare professionals, patients and the public

11. MHRA should work with the key information providers to ensure healthcare professionals and patients receive the best available information, and are empowered to make informed decisions and ask questions about any medicines they may be prescribed in pregnancy.

12. MHRA should do more to encourage and make it easier for women, and health professionals who work with women, to report any adverse reaction they experience while taking a medicine during pregnancy through the Yellow Card Scheme.

13. MHRA should build a partnership with other bodies within the healthcare system to improve the impact of safety messages relating to medicines, to support the objectives above.

REFERENCES

ABPI Data Sheet Compendium 1974. Drug Ther Bull;12(11):43-44.

Abramson D (1958). Clinical observation on the use of 17-ethynyl-19-nortestosterone in repeated and threatened abortions. Ann N Y Acad Sci.; 71(5):759-61.

Attia MA, Joosten HF, Zayed I and Hoekstra A. (1980) The role of the pituitary-ovarian axis in the termination of early pregnancy in rabbits treated with ethinyl-oestradiol. Arch Toxicol.; 43:163-76.

Badawy S, Abdul-Karim R. (1978) Effect of oestrogen on fetal survival in the rabbit. J Reprod Fertil.; 53:237-9.

Beck LR, Pope VZ. (1982). Demonstration of an early abortifacient effect of norethisterone (NET) in the primate (baboon). Contraception; 25: 97-105.

Botting BJ (2003). Evaluation of strategies to improve case ascertainment and data quality in the National Congenital Anomaly System. PhD University of London)

Brogan WF. (1975). Letter: Cleft lip palate and pregnancy tests. The Medical Journal of Australia; 1: 44.

Calzolari E, Barisic I, Loane M. et al. (2014). Epidemiology of multiple congenital anomalies in Europe: a EUROCAT population-based registry study. Birth Defects Research; 100: 270-276.

Commission Implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council.

Coomarasamy, A, Williams H, Truchanowicz et al. (2015). A randomized trial of progesterone in women with recurrent miscarriages. N Engl J Med; 373:2141-2148.

DeSesso JM. Comparative features of vertebrate embryology (2006). In Hood RD, ed Chapter 6 in Developmental and reproductive toxicology: A practical approach. 2nd ed. Boca Raton, FL: CRC Press, Inc: 147-197.

Directive 2010/84/EU of the European Parliament and of the Council of 15 December 2010 amending, as regards pharmacovigilance, Directive 2001/83/EC on the Community code relating to medicinal products for human use.

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on methodological standards in pharmacoepidemiology, Rev 6, 2017. http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

EUROCAT European surveillance of congenital anomalies. EUROCAT Guide 1.4 and reference documents (last version update 20/12/2016). http://www.eurocat-network.eu/content/Full%20Guide%201%204%20v5%2020_Dec2016.pdf

Ferencz C, Matanoski GM, Wilson PD et al. (1980). Maternal hormone therapy and congenital heart disease. Teratology; 21(2): 225-239.

Firth H. (1997) Chorion villus sampling and limb deficiency—cause or coincidence? Prenat Diagn; 17: 1313 – 30.

Gaal A and Ukvari E. (1977). Frequency of the treatment with female sex hormones during pregnancy in the case of infants born with neural tube anomalies (Hungarian). Magyar Noorvosok Lapja; 40: 257-259 (Hungarian).

Gal I, Kirman B, Stern J (1967). Hormonal pregnancy tests and congenital malformations. Nature; 216: 83.

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

Gal I (1978). Teratological adverse drug effects: review of evidence implicating hormonal pregnancy tests. Report to Sir Roland Moyle, Minister of State, Department of health and Social Services and to Jack Ashley, Esq., CH. Member of Parliament, House of Commons.

Glasziou P, Vandenbrouke J and Chalmers I (2004) Assessing the quality of research. BMJ;328(7430):39-41.

Goldzieher G (1994). Pharmacology of the Contraceptive Steroids, Raven Press. Part III. Studies in Laboratory Animals (Chapters 4-8, 28).

Goujard J, Rumeau-Rouquette C. (1977). First-trimester exposure progestogengen/oestrogen and congenital malformations. Lancet; 1: 482-483.

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 (French).

Greenberg G, Inman WHW, Weatherall, JAC et al. (1975). Hormonal pregnancy tests and congenital malformations. BMJ; 2: 191-192.

Greenberg G, Inman WHW, Weatherall, JAC et al. (1977). Maternal drug histories and congenital anomalies. BMJ; 2: 853-856.

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

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

Harlap S, Prywes R, Davies A Michael. (1975). Letter - Birth defects and oestrogens and progesterones in pregnancy. Lancet; 1: 682-683.

Heinonen OP, Slone D, Monson RR et al. (1977). Cardiovascular birth defects and antenatal exposure to female sex hormones. N Engl J Med; 296(2): 67-70.

Hendrickx AG, Korte R, Leuschner F et al. (1987a). Embryotoxicity of sex steroidal hormone combinations in nonhuman primates: I. Norethisterone acetate + ethinylestradiol and progesterone + estradiol benzoate (Macaca mulatta, Macaca fascicularis, and Papio cynocephalus). Teratology; 35:119-27.

Hendrickx AG, Korte R, Leuschner F et al. (1987b) Embryotoxicity of sex steroidal hormones in nonhuman primates: II. Hydroxyprogesterone caproate, estradiol valerate. Teratology; 35:129-36.

Higgins GL, Sadler WR. (1960). A two-tablet oral pregnancy test. Practitioner; 185:677-80.

Hodgkinson CP, Igna EJ, Bukeavich AP (1958). High potency progesterone drugs and threatened abortion. Am J Obstet Gynecol.; 76(2): 279-287.

Human Medicines Regulations 2012, as amended (S.I. 2012/1916)

Humpel M. (1982). Pharmacokinetics and biotransformation of norethisterone in animals: a review. Contraception; 28:83-95.

Huppertz B and Kingdom JCP (2007). The placenta and fetal membranes. Dewhurst's textbook of Obstetrics and Gynaecology; Ch 3, pp19-27.

Iqbal M, Audette M, Petropuolos S et al. (2012). Placental drug transporters and their role in fetal protection. Placenta;33:137-142.

Joshi A, Vaidya S, St-Pierre M et al. (2016). Placental ABC transporters: Biological impact and pharmaceutical significance. Pharmaceutical Research;33:2847-2878.

Joshi NJ, Ambani LM, Munshi SR. (1983). Evaluation of teratogenic potential of a combination of norethisterone and ethinylestradiol in rats. Indian J Exp Biol.; 21:591-6.

Kamke J (1978). Pregnancy test with consequences. Is Duogynon responsible for birth defects in thousands of children? Die Zeit; 32: 4th August (German).

Kazmierski MR (1976) A study of the teratogenic potential of oral hormonal pregnancy tests in the white rat. SoDM Masters Theses. 132. hhpt://digitalcommons.uconn.edu/sodm_masters/132

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

Lammer EJ and JF Cordero (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.

Lemmen JG, Broekhof L, Kuiper GG et al. (1999). Expression of estrogen receptor alpha and beta during mouse embryogenesis. Mech Dev; 81(1-2): 163-7.

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

Manco-Johnson MJ (2005). Development of hemostasis in the fetus. Thrombosis Research; 115: S55-63.

Markiewicz L., Gurpide E (1994). Estrogenic and progestagenic activities coexisting in steroidal drugs: quantitative evaluation by in vitro bioassays with human cells J. Steroid Biochem Molec Biol; 48 (1): 89-94.

Matsuura S, Itakura A, Ohno Y, et al. (2004). Effects of estradiol administration on feto-placental growth in rat. Early Hum Dev.; 77:47-56.

Medicines Act 1968 Chapter 67. http://www.iracm.com/wp-content/uploads/2013/01/sommaire-medicines-act-1968-modifi%C3%A9-3342.pdf

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

Mueller RF and Young ID (1998). Emery's elements of medical genetics. 10th Edition.

Nora, A. H. and J. J. Nora (1975". A syndrome of multiple congenital anomalies associated with teratogenic exposure. Arch Environ Health 30(1): 17-21.

Nora JJ, Nora AH, Blu J et al. (1978). Exogenous progestogen and estrogen implicated in birth defects. JAMA; 240(9): 837-843.

Nygren KG and Johansson EDB (1975). The effect of norethindrone and some other synthetic gestagens upon the peripheral plasma levels of progesterone and estradiol during early, human pregnancy. Acta Obster. Gynec. Scand; 54: 57-63.

Oakley GP, Flynt JW. (1973). Hormonal pregnancy tests and congenital malformations. Lancet; 2: 256-257.

Olszynko-Gryn J. (2014). Pregnancy testing in Britain, c.1900-67: laboratories, animals and demand from doctors, patients and consumers. Dissertation submitted for the Doctor of Philosophy.

Orme M, Back D, Ball S. (1989) Interindividual variability in the pharmacokinetics of ethinylestradiol. Pharmacology and therapeutics; 43:251-60.

Pulkkinen M, Dusterberg B, Hasan H et al. (1984). Norethisterone acetate and ethinylestradiol in early human pregnancy. Teratology; 29:241-249.

Regulation (EU) No 1235/2010 of the European Parliament and of the Council of 15 December 2010 amending, as regards pharmacovigilance of medicinal products for human use, Regulation (EC) No 726/2004 laying down Community procedures for the authorization and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency, and Regulation (EC) No 1394/2007 on advanced therapy medicinal products.

Sarkar K, Kinson GA, Rowsell HC. (1986). Embryo resorption following administration of steroidal compounds to rats in mid pregnancy. Can J Vet Res.; 50:433-7.

Schering study (1963). Effect of NEA on intrauterine foetal growth, the placenta and uterus after S.C. administration in various stages of pregnancy. Schering AG 1963. Report of 15.03.1963.

Schering study 06/12/1965. Test for embryotoxic action in mice after oral administration (NEA:EE ratio 80:1).

Schering study 20/12/1965. Test for embryotoxic action in mice after oral administration (NEA: EE ratio 60:1).

Schering study 983 (1973): ZK 4.944 Prufung auf embryotoxische Wirkung an Ratten. Reprotox order number 11. Study report dated 17.4.1973.

Schering study 1443 (1970). Test for embryotoxic action in rabbits after oral administration (NEA:EE ratio 500:1).

Schering study 1631 (1970). Prufung auf embryotoxische wirkung an Ratten. Conducted 1969.

Schering study 2221 (1971). Prufung auf embryotoxische wirkung an Ratten. Study conducted 1969.

Schering study 2300/Protocol No. 494/75 (1976). Test for embryotoxic action in rabbits after intragastric administration.

Schering study 2330/Protocol No 493/75 (1976). Test for embryotoxic action in rats after intragastric administration.

Schering study 3578 (1978). Test for embryotoxic action in rats after intragastric administration (NEA:EE ratio 500:1).

Schering study 3581 (1978). Test for embryotoxic action in rabbits after intragastric administration (NEA:EE ratio 500:1).

Schering study 3579 (1978). Test for embryotoxic action in mice after intragastric administration (NEA:EE ratio 500:1).

Schering study 4036 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 175) of April 23rd, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 14 and 15 of pregnancy on foetuses of the rabbit. Kommentar- Nr. 4036, vom 4.12.1979 [full study report also included].

Schering study 4037 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 166) of March 7th, 1979: ZK 4 . 944 + ZK 5. 422 (1 + 500). Effect of intragastral administration on day 6 and 7 of pregnancy on foetuses of the rat. Kommentar- Nr. 4037, vom 4.12.1979 [full study report also included].

Schering study 4038 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 174) of April 17th, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 12 and 13 of pregnancy on foetuses of the rabbit. Kommentar-Nr. 4038, vom 4.12.1979 [full study report also included].

Schering study 4039 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 176) of April 25th, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 16 and 17 of pregnancy on foetuses of the rabbit. Kommentar- Nr. 4039, vom 4.12.1979 [full study report also included].

Schering study 4040 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 173) of April 2nd, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 10 and 11 of pregnancy on foetuses of the rabbit. Kommentar-Nr. 4040, vom 4.12.1979 [full study report also included].

Schering study 4041 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 172) of March 29th, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 8 and 9 of pregnancy on foetuses of the rabbit. Kommnetar- Nr. 4041, vom 4.12.1979 [full study report included].

Schering study 4042 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 168) of April 6th, 1979: ZK 4.944 + ZK 5.422 (1 + 500) - Effect of intragastral administration on day 10 and 11 of pregnancy on foetuses of the rat. Kommentar- Nr. 4042, vom 4.12.1979 [full study report also included].

Schering study 4043 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 171) of March 12nd, 1979: ZK 4.944 + ZK 5.422 (1 + 500). Effect of intragastral administration on day 6 and 7 of pregnancy on foetuses of the rabbit. Kommentar- Nr. 4043, vom 4.12.1979 [commentary and study report included].

Schering study 4044 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 170) of April 9th, 1979: ZK 4.944 + ZK 5.422 (1 + 50 0). Effect of intragastral administration on day 14 and 15 of pregnancy on foetuses of the rat. Kommentar- Nr. 4044, vom 4.1.1979 [full study report also included].

Schering study 4045 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 169) of April 10th, 1979: ZK 4. 944 + ZK 5. 422 (1 + 500) - Effect of intragastral administration on day 12 and 13 of pregnancy on foetuses of the rat. Kommentar- Nr. 4045, vom 4.12.1979 [full study report also included].

Schering study 4046 (1979). Commentary on the report of Reprotox/HRC Germany (Order no. 167) of April 4th, 1979: ZK 4. 944 + ZK 5. 422 (1 + 500) - Effect of intragastral administration on day 8 and 9 of pregnancy on foetuses of the rat. Kommentar - Nr. 4046, vom 4.12.1979 [full study report also included].

Schering study 4282 (1978). Commentary on the report of Lee Newman Research Associates, Barnstaple, Devon, England, of March 24th,1980: "ZK 4.944 + 5.422 (1 + 500) Pathology report of protocol 97/78 - Embryotoxicity study in rats" (Supplement to report no. 3578 of December 22nd,1978) - Kommentar- Nr. 4284 vom 21.5.1980.

Schering study 5303 (1978). Effect of ENTA [sic] given orally during the period of major organogenesis of the rat progeny. ZK No. 5422. Bericht-Nr. 1/2/5303 von N. Kodama, Nihon Schering, vom 26.12.1978 [commentary and full report].

Schering study 5304 (1978). Effect of ENTA given orally during the period of major organogenesis of the mouse progeny. Bericht- Nr .1/2/5304 von N. Kodama, Nihon Schering, vom 26.12 .1978 [commentary and full report].

Effect of ENTA given orally during the period of major organogenesis of the mouse progeny. Bericht- Nr .1/2/5304 von N. Kodama, Nihon Schering, vom 26.12 .1978 [commentary and full report]

Schering study 3861/Protocol No.98/78 (1979). Test for embryotoxic action in mice after intragastric administration. Histological studies of the thorax (NEA:EE ratio 500:1).

Schering study 4136 (1980). ZK 4.944 - Embryotoxicity study in rats after daily intragastrical administration during days 6 – 15 of gestation. Pharma-Forschungs- Bericht Nr. 4136 vom 22.1.1980 [Commentary and full report].

Schindler A, Campagnoli C, Druckman R et al. (2003). Classification and pharmacology of progestins. Maturitas; 46S1: S7-S16.

Slater BCS, Watson GI and McDonald JC (1964). Seasonal variation in congenital abnormalities. Preliminary report of a survey conducted by the Research Committee of Council of the College of General Practitioners. Brit J Prev Soc Med;18:1-7.

Slikker W, Bailey JR, Newport D et al. (1982). Placental transfer and metabolism of 17α-ethynylestradiol-17β in the rhesus monkey. J Pharmacol Exp Ther.; 223:483- 489.

Soskin S, Wachfel H and Hetchter O (1940) The treatment of delayed menstruation with prostigmine. A therapeutic test for early pregnancy. JAMA 114(21) 2090-2091.

Steinetz BG, Butler MC, Sawyer WK, et al. (1976) Post-implantation termination of pregnancy in rats. Contraception; 14(5):487-505.

Suzuki MR; Eguchi K, Yoda M (1978) Effects of some gestagens on the fetuses of rat and mouse Japanese). Pharmacometrics; 15(5): 955-960.

Tauber U, Humpel M (1984). Distribution of norethisterone and ethinylestradiol in the gravid rat and in the feto-placental unit of rhesus monkeys following intragastric administration of norethisterone acetate in combination with ethinylestradiol. I. Direct determinations. Arzneimittel-Forschung; 34(4):440-6.

Thierstein ST, Rodler DE. (1959). Norlutin and stilbestrol compared to progesterone and stilbestrol in treating threatened and repeated abortion. Nebr State Med J.; 44: 433-441.

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

Tümmler G, Ribmann A, Meister R and Shaefer C (2014). Congenital bladder exstrophy associated with Duogynon hormonal pregnancy tests – Signal for teratogenicity or consumer report bias? Reproductive Toxicology; 45: 14-19.

Van Gelder MMHJ, van Rooil IALM, Miller RK et al. (2010). Teratogenic mechanisms of medical drugs. Human Reproduction Update;16:378-94.

Weatherall JAC 1978. Congenital malformations: surveillance and reporting. Population Trends; 11: 27-29

Webster WS, Abela D (2007). The Effect of Hypoxia in Development Birth Defects Research (Part C); 81:215–228.

Wharton LR, Scott RB (1964). Experimental production of genital lesions with norethindrone. Am. J. Obstet. Gynecol.; 89: 701–715.

Wilson, JG and Brent RL (1981). Are female sex hormones teratogenic? American Journal of Obstetrics and Gynecology;141: 567-580.

Wiseman, RA and Dodds-Smith IC (1984). Cardiovascular birth defects and antenatal exposure to female sex hormones: a re-evaluation of some base data. Teratology;30(3): 359-370.

Xia B, Heimbach T, Gollen R et al. (2013). A Simplified PBPK Modeling Approach for Prediction of Pharmacokinetics of Four Primarily Renally Excreted and CYP3A Metabolized Compounds during Pregnancy. AAPS Journal;15(4):1012-1024.

ANNEXES

Annex 1:	HPT EWG paper: Introduction and background to the review of Hormonal Pregnancy Tests (1st meeting)
Annex 2:	HPT EWG paper: Structure and remit of the Expert Working Group on Hormonal Pregnancy Tests (1st meeting)
Annex 3:	HPT EWG paper: Updated chronology of events: Hormonal Pregnancy Tests (7th meeting)
Annex 4:	HPT EWG paper: Programme of work, including: summary of information collated; proposed topics to cover in future meetings (1st meeting)
Annex 5:	HPT EWG paper: Updated schedule of information on HPTs (7th meeting)
Annex 6:	Conflict of interest policy for HPT EWG
Annex 7:	Code of Practice for Chairmen and Members of the Commission on Human Medicines, Certain Committees and Expert Advisory Groups
Annex 8:	EWG participant declarations of interest (as of 20/7/2017)
Annex 9:	HPT EWG paper: A brief history of pregnancy testing, socio-medical landscape and medicines regulation in the UK: a summary of background information
Annex 10:	HPT EWG paper: Normal human embryonic development and congenital anomalies
Annex 11:	Presentation: A current update on congenital anomalies, Dr D Wellesley
Annex 12:	Report from Dr I Gal to Sir Roland Moyle (197–) - Teratological adverse drug effects: Review of evidence implicating hormonal pregnancy tests.
Annex 13:	Report by RA Wiseman. A study of sales of HPTs and congenital malformations. Undated.
Annex 14:	Primodos data sheet, 1974
Annex 15:	CSM warning letter, 1975
Annex 16:	Primodos data sheet, 1976
Annex 17:	CSM reminder warning, 1977
Annex 18:	HPT EWG paper: Possible effect of NETA/EE on the developing fetus: evidence from pharmacological data
Annex 19	HPT EWG paper: Pharmacokinetics and pharmacology of norethisterone and ethinylestradiol, the steroid components of Primodos hormonal pregnancy test
Annex 20:	HPT EWG paper: Review of non-clinical evidence of reproductive and developmental toxicity for norethisterone acetate and ethinylestradiol

Annex 21:	HPT EWG paper: Review of the evidence for vascular disruption during pregnancy in association with Hormone Pregnancy Tests (HPTs)
Annex 22:	HPT EWG paper: Assessment of spontaneous reporting data for adverse reactions relating to Hormone Pregnancy Tests
Annex 23:	HPT EWG paper: Further analysis of spontaneous reports with Hormone Pregnancy Tests
Annex 24:	HPT EWG paper: Further analysis of spontaneous reports with Hormonal Pregnancy Tests – second corrected.
Annex 25:	HPT EWG paper: Review of the epidemiological evidence for an association between use of ethinylestradiol and norethisterone acetate in early pregnancy and congenital anomalies or miscarriage
Annex 26:	HPT EWG paper: Review of the epidemiological evidence for an association between use of Hormonal Pregnancy Tests in early pregnancy and congenital anomalies
Annex 27:	HPT EWG paper: Re-analysis of the epidemiological evidence for a possible association between HPT use and congenital anomalies
Annex 28:	HPT EWG paper: Review of the epidemiological evidence for an association between use of oral contraceptives in early pregnancy and congenital anomalies
Annex 29:	HPT EWG paper: Review of the epidemiological evidence for an association between use of norethisterone for threatened/recurrent miscarriage and congenital anomalies
Annex 30:	HPT EWG paper: Review of the epidemiological evidence for an association between use of norethisterone and/or ethinylestradiol preparations in early pregnancy and miscarriage
Annex 31:	HPT EWG paper: Lessons learnt with respect to identifying, assessing and communicating drug safety concerns in pregnancy

GLOSSARY OF TERMS

Term	Meaning
2010 Pharmacovigilance legislation	Directive 2010/84/EU, Regulation No. 1235/2010 and Implementing Regulation No 520/2012
Affinity (Ki and Kd)	Measure of strength of binding of a ligand to its receptor: Kd is measured by direct binding of compound with a radiolabelled attached. Ki is measured from indirectly, from displacement of binding of another labelled ligand.
Agonist	Substance that initiates a biological response when it binds a receptor.
Anencephaly	Absence of a major portion of the brain, skull and scalp that occurs during embryonic development.
Anophthalmia	Absence of one or both eyes.
Antagonist	Substance that blocks or reduces a biological response when it binds to a receptor.
Anti-gonadotropic	Suppression of the activity and/or downstream effects of one or both the gonadotropins, follicle-stimulating hormone and luteinizing hormone. This results in an inhibition of the hypothalamic-pituitary-gonadal axis, and a decrease in the levels of the androgen, estrogen, and progestogen sex steroids in the body.
Anti-androgenic	Preventing androgens such as testosterone and dihydrotestosterone from mediating their biological effects in the body by blocking the androgen receptor and/or inhibiting or suppressing androgen production.
Anti-mineralocorticoid	Anti-mineralocorticoid action, or aldosterone antagonism, causes a diuretic effect through antagonising the action of aldosterone via mineralocorticoid receptors.
Atrial septal defect	A congenital heart defect in which blood flows between the atria (upper chambers) of the heart, which are normally separated by a dividing wall.
Backbench Business Committee	Set up in 2010 this committee is responsible for determining the business before the House of Commons for approximately one day each week.
Bias	In an epidemiological study, any systematic errors in the research methodology that result in an incorrect estimate of the association between exposure and risk of outcome.
Bioavailability	The fraction of an administered dose of unchanged drug that reaches the systemic circulation.

Commission on Human Medicines (CHM)	The CHM advises ministers on the safety, efficacy and quality of medicinal products. It is an advisory non-departmental public body, sponsored by the Department of Health.
Conception	Fertilisation of the oocyte.
Conceptus	Products of conception or fertilisation including the embryo or fetus, placenta and membranes.
Confounding	Factors that are associated with the exposure of interest and also with the outcome of interest. Residual confounding can lead to bias that distorts the magnitude of the relationship between the exposure and outcome of interest.
Congenital anomaly	Also known as a congenital disorder or birth defect, this is a condition existing at or before birth that is characterised by a structural deformity. Congenital anomalies vary widely in cause and symptoms. Any substance that causes birth defects is known as a teratogen.
Data sheet	A document, originally provided by the licence holder for use by healthcare professionals, that summarises the conditions for safe and effective use of a medicine. This has been superseded by the Summary of Product Characteristics (SmPC).
Department of Health	The Ministerial Department of the United Kingdom Government responsible for government policy on health and adult social care matters in England, along with a few elements of the same matters which are not otherwise devolved to the Scottish Government, Welsh Government or Northern Ireland Executive. It oversees the English National Health Service (NHS). Some of its work is carried out through arms-length bodies including MHRA.
Development days	Duration of embryonic/fetal growth.
Developmental delay	When a child does not reach their developmental milestones at the expected times. It is an ongoing (rather than temporary) major or minor delay in the process of development.
Developmental toxicity	The ability of a chemical or physical agent to cause any of the manifestations of adverse developmental outcome (i.e. malformation, embryo lethality, growth retardation, functional deficit) either individually or in combination.
Disposition	Knowing the fate of a drug and how it is absorbed, distributed, metabolised, and excreted.
Ethinylestradiol	Ethinylestradiol is a synthetic estrogen, derived from estradiol, which acts through the estrogen receptor.
Embryo	A new organism in the earliest stage of development. In humans, this is defined as the developing organism

	from the fourth day after fertilization to the end of the eighth week. After that the unborn baby is usually referred to as the fetus.
Embryo-fetal lethality	Death of an embryo or fetus, which can result in abortion/resorption of the embryo/fetus back into the mother's body.
Embryogenesis	The process by which the embryo forms and develops in pregnancy
Embryotoxicity	Toxicity during the embryo developmental process.
Endogenous	Naturally occurring/produced by the body (as for sex steroid hormones).
Endometrium	Lining of the womb/uterus
Enzyme	A biological catalyst that accelerates chemical reactions within the cells of the body.
Epidemiological studies	Studies on human populations, which attempt to link human health effects to a cause.
Estradiol	Most commonly occurring natural estrogen.
Ethisterone	A synthetic progestogen derived from testosterone. Also known as pregneninolone, 17α-ethynyltestosterone, and 19-norandrostane.
EUROCAT	European network of population-based registries for the epidemiologic surveillance of congenital anomalies, established in 1979 and surveying over 1.7 million births per year in Europe across 43 registries in 23 countries.
Exencephaly	A disorder in which the brain is located outside of the skull.
Exogenous	Originating from outside the body.
Exposure	Total amount of a substance to which the body is exposed.
Fertilisation	The union of an egg and sperm.
Fetotoxicity	Refers to toxicity during the fetal developmental process.
Fetus	Unborn offspring – in humans more than 8 weeks after conception.
First pass effect	Where the concentration of a drug is reduced (mostly through metabolism by the gut wall and liver) before it reaches the systemic circulation.
Forest plot	A forest plot is a graphical display of estimated results from a number of scientific studies addressing the same question, along with the overall results. The left-hand column lists the names of the studies; the right-hand column is a plot of the measure of effect for each of these studies incorporating confidence intervals represented by horizontal lines. A vertical line through

	1, representing no effect, is also plotted. If the confidence intervals for individual studies overlap with this line, it demonstrates that at the given level of confidence their effect sizes do not differ from no effect for the individual study
Gestation	Duration of pregnancy.
Good Vigilance Practice	Detailed European regulatory guidance documents underpinning pharmacovigilance legislation.
HM Regulations 2012	Human Medicines Regulations 2012, as amended (S.I. 2012/1916) – can be found at: http://www.legislation.gov.uk/
Hormone pregnancy test	Products containing an estrogen and progestogen (or rarely a progestogen only) given during the first three months to diagnose pregnancy – available in the UK in the 1950s to 1970s.
ICH	The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) was established in 1990 discusses scientific and technical aspects of drug authorisation with a view to achieving greater harmonisation worldwide to ensure the development of safe, effective, and high quality medicines.
Implantation	Adherence of the embryo to the wall of the uterus in early pregnancy.
In utero	In the uterus, before birth.
In vitro	Referred to as test tube experiments, these are studies conducted outside of the usual biological context.
In vivo	Studies performed in a living organism.
Ligand	A molecule that binds to a receptor.
Lipophilic	Able to dissolve in lipids and fats.
Kinetics	The absorption, distribution, metabolism and excretion of a substance in the body.
Limb reduction defect	Upper and lower limb reduction defects occur when a part of or the entire arm or leg of a fetus fails to form completely during pregnancy and so is reduced in size or missing completely.
Macroscopic	Large enough to be visible without magnifying optical instruments.
Malformation	Defective or abnormal formation; deformity: a permanent structural deviation.
Mammary	Relating to the female breast.
Marketing authorisation holder	Pharmaceutical company owning a marketing authorisation, or product licence, for a medicine.
MedDRA	Clinically validated international medical terminology dictionary used by regulatory authorities in the

	pharmaceutical industry during the regulatory process. It is the adverse event classification dictionary endorsed by the ICH.
Medical Data Index	Prescription data compiled historically by an independent market research organisation, Intercontinental Medical Statistics Limited (IMS)
Medicines and Healthcare products Regulatory Agency (MHRA)	MHRA regulates medicines, medical devices and blood components for transfusion and is responsible for ensuring their safety, quality and effectiveness. MHRA is an executive agency of the UK Department of Health.
Microphthalmos	Condition of having unusually small eyeballs.
Microscopic	So small as to be visible only with magnifying optical instruments.
Miscarriage	Spontaneous loss of an embryo or fetus, usually in the first 23 weeks of pregnancy. Also known as spontaneous abortion or pregnancy loss.
The National Archives (TNA)	A non-ministerial government department that is the official archive of the UK government and for England and Wales.
Non-clinical studies	In drug development, preclinical development, also named preclinical studies and nonclinical studies, is a stage of research that begins before clinical trials (testing in humans) can begin, and during which important feasibility, iterative testing and drug safety data are collected.
Norethisterone	A synthetic progestogen derived from 19-nortestosterone. Also known as norethindrone.
Observational study	A type of study in which individuals are observed or certain outcomes are measured. No attempt is made to affect the outcome (for example, no treatment is given).
Offspring	A person' child or children or an animal's young.
Oocyte	Immature ovum, or egg cell.
Open label study	A type of clinical trial in which both the researchers and participants know which treatment is being administered, that is it is unblinded.
Organogenesis	In humans, development of the internal organs in the uterus in weeks 3 to 8 of embryonic development.
Patient Information Leaflet	The patient information leaflet (PIL) guides patients or their carers on the safe and effective use of a medicine. PILs include all the information provided to healthcare professionals in the Summary of Product Characteristics but in lay language. PILs are produced by the marketing authorisation holder and approved by MHRA. Unless all the information is on the pack, all medicines must be accompanied by a PIL.

Pharmacokinetics	The branch of pharmacology that studies how the body absorbs, distributes, and eliminates a drug.
Pharmacology	Branch of medicine/science relating to the uses, effects, and modes of action of drugs.
Phocomelia	A rare congenital anomaly in which the hands or feet are attached close to the trunk, the limbs being grossly underdeveloped or absent.
Placebo	A tablet or liquid that contains no medically active substance but is otherwise indistinguishable from the medicine.
Placenta	An organ that connects the developing fetus to the wall of the uterus to allow nutrient uptake by the fetus, waste elimination from fetal blood, and gas exchange via the mother's blood supply. The placenta also produces hormones to support pregnancy.
Post-natal	After birth
Power	The power of a study is the likelihood that it will distinguish an effect of a certain size from pure chance. A low-powered study might detect a substantial risk from a medicine, but is less likely to detect a small difference.
PRAC	The PRAC is the European Medicines Agency's committee responsible for assessing and monitoring the safety of human medicines.
Pre-clinical development	In drug development, preclinical development, also named preclinical studies and nonclinical studies, is a stage of research that begins before clinical trials (testing in humans) can begin, and during which important feasibility, iterative testing and drug safety data are collected.
Pregnancy prevention plan	Set of measures implemented to minimise exposure during pregnancy to a medicine with known teratogenic effects.
Prenatal development	Development of the embryo and fetus after fertilisation and before birth.
Progesterone	Endogenous steroidal sex hormone involved in the menstrual cycle, pregnancy and embryogenesis.
Progestogen	A synthetic hormone that acts at progesterone receptors to mimic its actions.
Receptor	A region of tissue, or a molecule in a cell membrane, which responds specifically to a particular neurotransmitter, hormone, antigen, or other substance.
Reliable	In epidemiology – ability to reach an unbiased answer to the question that was being asked
Reproductive toxicity	Reproductive toxicity studies investigate whether drugs

	can have adverse effects on sexual function and fertility in the adult male or female, as well as developmental toxicity in the offspring.
Resorption	Early death of the embryo during the fetal period with maternal assimilation of the products of conception (usually in animals).
Robust	A finding or conclusion that is considered to be reliable because it is based on high quality supporting data that are well designed to minimise potential biases.
Secondary amenorrhoea	Menstruation has previously occurred but it has stopped - definitions vary as to how long but it is usually taken as at least six consecutive months (longer where menses have previously been infrequent).
Situs inversus	Congenital condition in which the major internal organs are reversed or mirrored from their normal positions.
Somatic cell	Every cell type in the mammalian body, apart from the sperm and ova (eggs), and undifferentiated stem cells (eg all internal organs skin, bones, blood and connective tissue)
Spina bifida	A birth defect believed to be due to a combination of genetic and environmental factors, where there is incomplete closing of the backbone and membranes around the spinal cord.
Spontaneous abortion	The natural death and vaginal loss of an embryo or fetus - usually in the first 23 weeks of pregnancy. Also known as complete miscarriage or pregnancy loss.
Statistical significance	The likelihood that a relationship between two or more variables is caused by something other than random chance.
Stillbirth	The delivery of a dead baby, usually after 24 completed weeks of pregnancy.
Strong association	A strong association between possible cause and observed effect generally has a larger effect size (relative risk/odds ratio) and is unlikely to be due to chance or bias.
Summary of Product Characteristics	Document provided by the marketing authorisation holder and approved by MHRA that details the terms for the safe and effective use of a medicine.
Systemic	Throughout the whole body
Teratogen	A substance that can disturb the development of and/or cause malformations in an embryo or fetus.
Testis	Organ that produces sperm.
Threatened miscarriage	Some vaginal bleeding but pregnancy continues
Uterus	Womb
Variation	In teratology – a divergence beyond the usual range of

	structural constitution or development that may not adversely affect survival or health.
Ventricular septal defect	Opening in the wall (septum) that separates the two lower chambers (ventricles) of the heart.
Vigibase	The WHO's database that contains adverse drug reactions reported globally by member states enrolled under the WHO's international drug monitoring programme.
Virilisation	The abnormal development of male sexual characteristics in a female.
Weak association	A weak association between possible cause and observed effect generally has a smaller effect size (relative risk/odds ratio) and could be due to chance or bias.
Yellow Card Scheme (YCS)	The Yellow Card Scheme is the UK system for collecting information on suspected ADRs to medicines. The scheme was founded in 1964 and allows the safety of the medicines and vaccines that are on the market to be monitored. It is run by the Medicines and Healthcare Products Regulatory Agency (MHRA) and the Commission on Human Medicines (CHM).
Yolk sac	The yolk sac provides nourishment to the embryo, which is attached. The placenta takes over this role a few weeks later.
Zebrafish	The zebrafish is a widely used vertebrate model organism in scientific research. It is particularly notable for its regenerative abilities, and has been modified to produce many transgenic strains. Zebrafish embryos have proven to be a rapid, cost-efficient, and reliable teratology assay model. Drug screens in zebrafish can be used to identify novel classes of compounds with biological effects, or to repurpose existing drugs for novel uses.