PASS information

Title	Analysis of pregnancy pharmacovigilance data in spontaneous reports, and literature, (Individual Case Safety Reports originating from published case series, non-interventional studies and patient support programmes); demonstration study **2.5.1** of the ConcePTION project
Protocol version identifier	1.0
Date of last version of protocol	23-12-2020
EU PAS register number	Study not registered
Active substance	For analysing spontaneous data of (national) pharmacovigilance (PV) centres, the following medicinal products (which are also part of the **2.5.2** protocol) will be covered: alemtuzumab (ATC index L04AA34), azathioprine (L04AX01), (oral) cladribine (L01AA40/L04AA40), cyclophosphamide (L01AA01), daclizumab (L04AC01), dalfampridine (N07XX07), dimethylfumarate (N07XX09), fingolimod (L04AA27), glatiramer (L04AA07), glatiramer acetate (L03AX13), interferon beta-1a (L03AB07), interferon beta-1b (L03AB08), intravenous immunoglobulin, laquinimod (N07XX10), leflunomide (L04AA13), methotrexate (L01BA01), mitoxantrone (L01DB07), mycophenolate mofetil/mycophenolic acid (L04AA06), natalizumab (L04AA23), ocreluzimab (L04AA36), ofatumumab (L01XC10), ozanimod (L04AA38), peginterferon beta-1a (L03AB13), rituximab (L01XC02), siponimod (L04AA42), sodium valproate (N03AG01) and teriflunomide (L04AA31). Other medicinal products (e.g. COVID-19 vaccines and treatments, pertussis vaccine (J07CA09)), might be identified later and added to the list where appropriate.
Medicinal product	See active substance
Product reference	Not applicable
Procedure number	Not applicable
Joint PASS	Not applicable
Research question and objectives	The main objective of demonstration study **2.5.1** is to gain insight into the nature of information on drug exposure during pregnancy from spontaneous reports and literature reports by • Describing the nature and content of spontaneous reporting systems (SRS) and literature data sources • Creating and validating a new optimized assessment tool for the quality of pregnancy data specifically • Describing the quality of reports in SRS, literature, teratology information services (TIS), registries and enhanced PV programme data • Assessing predictors of currently used teratogen signal detection techniques in international case safety reports (ICSR) databases • Exploration of cluster analysis as a possible new teratogen signal detection technique
EFPIA partners	Novartis Pharma AG Novo Nordisk
Academic partners/ENTIS centres	Kwazulu-Natal Research Innovation and Sequencing Platform (KRISP) Swiss Teratogen Information Service UK Teratogen Information Service Pharmacovigilance Centre Lareb
Authors	Yrea Weetink, MSc The Netherlands Pharmacovigilance Centre Lareb y.weetink@lareb.nl

	Eugène van Puijenbroek, MD, PhD The Netherlands Pharmacovigilance Centre Lareb e.vanpuijenbroek@lareb.nl David John Lewis, PhD Novartis Pharma GmbH, Germany University of Hertfordshire, England David-1.lewis@novartis.com Alan Moore, MSc Novartis Pharma AG, Switzerland alan.moore@novartis.com Laura Yates, MBCHB, PhD University of KwaZulu-Natal, South Africa Laura.Yates@ialch.co.za Yvonne Geissbühler, PhD Novartis Pharma AG, Switzerland Yvonne.geissbuehler@novartis.com Michael Stellfeld Novo Nordisk A/S, Denmark msfe@novonordisk.com Ursula Winterfeld, PhD Swiss Teratogen Information Service ursula.winterfeld@chuv.ch Amalia Alexe, MSc Pharm Advanced Accelerator Applications Amalia.alexe@novartis.com Bita Rezaallah, DMD PhD Novartis Pharma AG, Switzerland Bita.rezaallah@novartis.com Jonathan Richardson, PhD UKTIS Jonathan.Richardson3@nhs.net Alison Oliver, PhD UKTIS Alison.Oliver12@nhs.net T. Christopher Bond, PhD Bristol Myers Squibb christopher.bond@bms.com

1. Table of contents

2. List of abbreviations

ADR	Adverse Drug Reaction
AUC	Area Under the Curve
ATC	Anatomical Therapeutic Chemical classification system
CCSI	Company Core Safety Information
CDE	Core Data Elements
CDM	Common Data Model
EDD	Expected Date of Delivery
EFPIA	European Federation of Pharmaceutical Industries and Associations
EMA	European Medicines Agency
ENTIS	European Network of Teratology Information Services
EOP	End Of Pregnancy
HCA	Hierarchical Cluster Analysis
HCP	Health Care Professional
ICSR	Individual Case Safety Report
IMI	Innovative Medicines Initiative
LMP	Last Menstrual Period
MAH	Marketing Authorisation Holder
MedDRA	Medical Dictionary for Regulatory Activities
MS	Multiple Sclerosis
OTIS	Organization of Teratogen Information Specialists
PAES	Post Authorisation Efficacy Studies
PASS	Post Authorisation Safety Studies
PRIM	PRegnancy outcomes Intensive Monitoring
PV	Pharmacovigilance
PV-report	Spontaneous reports database of Lareb
ROC	Receiver Operating Curve
SmPC	Summary of Product Characteristics
SQL	Structured Query Language
SRS	Spontaneous Reporting Systems
STIS	Swiss Teratogen Information Service
TIS	Teratology Information Services
UMC	Uppsala Monitoring Centre
WHO	World Health Organization
WMO	Wet Medisch-wetenschappelijk Onderzoek met mensen
WP	Work Package

3. Responsible parties

The Netherlands Pharmacovigilance Centre Lareb
Goudsbloemvallei 7, 5237 MH 's-Hertogenbosch
The Netherlands

KRISP
University of KwaZulu-Natal, Durban
Nelson R. Mandela Medical School, 719 Umbilo Rd, Umbilo, Durban, 4001
South Africa

Novartis Pharma AG
Novartis Campus
CH-4002 Basel
Switzerland

Novo Nordisk
Novo Alle, DK-2880 Bagsvaerd
Denmark

Swiss Teratogen Information Service
Service de Pharmacologie Clinique, CHUV, Rue du Bugnon 17, 1011 Lausanne
Switzerland

UK Teratology Information Service
Newcastle upon Tyne Hospitals NHS Foundation Trust
16-17 Framlington Place, Newcastle upon Tyne, England, United Kingdom, NE2 4AB

4. Acknowledgement

The ConcePTION project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 821520. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

5. Abstract

Analysis of pregnancy pharmacovigilance data in spontaneous reports and literature; demonstration study **2.5.1** of the ConcePTION project.

Since limited data from studies performed pre-marketing are usually available before licensure of a medicinal product, we have to rely on post-marketing data from both primary as well as secondary data sources. The IMI funded ConcePTION project aims to enhance the way drug use during pregnancy is studied. This is in part achieved by improving the collection, analysis and interpretation of pharmacovigilance (PV) data, to allow for a more systematic analysis and exchange of data.

Work Package 2 (**WP2**) focusses on sources of primary data collection, such as spontaneous reports, data collected by Teratogen Information Services (TIS), literature, pregnancy registries, and enhanced PV studies. Tools developed for the analysis of spontaneous reports, however, were not specifically aimed at the analysis of safety information related to pregnancy. As a first step, this demonstration study will aim to gain insight into the nature of information on drug exposure during pregnancy from spontaneous reports and literature reports as filed in the ICSR databases of national PV centres and Marketing Authorisation Holders. The category literature reports therefore encompass Individual Case Safety Reports originating from published case series, non-interventional studies and Patient Support Programmes.

In order to achieve this general aim, 5 sub-studies have been designed. The first sub-study aims to describe the nature and content of spontaneous reports and literature data sources. The second sub-study aims to create and validate a dedicated assessment tool for measuring the clinical quality of pregnancy data specifically, and the third sub-study aims to use this newly developed tool in order to describe the quality of reports in spontaneous reports and literature. This substudy also takes the clinical quality of data from TIS, registries and enhanced PV programme data into account which -for all other aspects- will be studied in demonstration study **2.5.2**. Finally, we will explore possibilities to improve the analysis of pregnancy related reports in spontaneous reports databases and literature by studying the characteristics of teratogen signal detection approaches currently in use and we will explore new possibilities for analysis of this data. In order to do that, sub-study 4 will assess predictors of currently used teratogen signal detection techniques in individual case safety reports databases, and sub-study 5 aims to explore cluster analysis as a possible new teratogen signal detection technique.

6. Amendments and updates

Number	Date	Section of study protocol	Amendment or update	Reason
1				
2				
3				

7. Milestones

Milestone	Planned date
Start of data collection	Not applicable
End of data collection	Not applicable
Registration in the EU PAS register	
Final report of study results	

8. Rationale and background

8.1. ConcePTION project

Effects of drug use during pregnancy are often unknown at the time of product licensure because the inclusion of pregnant women in clinical trials is considered unethical. (1) However, treatment during pregnancy is often needed for women with pre-existing chronic conditions such as asthma, epilepsy, cancer, multiple sclerosis (MS) and depression. In daily practice the majority of women take at least one type of medicinal product during pregnancy. (2-4) Knowledge about the safety of drug use during pregnancy is therefore warranted. Since limited data from studies performed pre-marketing are usually available before licensure of a medicinal product, we have to rely on post-marketing data from both primary as well as secondary data sources. With respect to the primary data, several established approaches are currently used to compile and analyse the data on the safety profile of drug use during pregnancy, including data collected by Teratology Information Services (TIS), spontaneous reports, information from literature, organised data collection programmes and registries.

The IMI funded ConcePTION project aims to enhance the way drug use during pregnancy is studied. (4) This is done in part by improving the collection, analysis and interpretation of pharmacovigilance (PV) data, to allow for a more systematic analysis and exchange of data. **WP2** (Work Package 2) focusses on sources of primary data collection, such as spontaneous reports, data collected by TIS, literature, pregnancy registries, and enhanced PV studies.

8.2. Individual Case Safety Reports databases

Post-marketing data are essential to evaluate the safety of drugs used during pregnancy. One way to collect data is via spontaneous reports, also known as Individual Case Safety Reports (ICSRs), which are recorded in databases of national PV centres and Marketing Authorisation Holders (MAHs). ICSRs from products licenced in the European Union are forwarded to the European EudraVigilance database. (5-7) The system is operated by the European Medicines Agency (EMA) and facilitates management and analysis of reported suspected adverse reactions to medicines. (7) The information in these reports is subsequently forwarded by the EMA to the WHO (World Health organization) collaborating centre for international drug monitoring (the Uppsala Monitoring Centre (UMC)), where data can be analysed on a global level. (8) Data collected in spontaneous report databases is retrospective in nature.

For medicinal products available on the European market, also cases described in literature are available as ICSRs in the EudraVigilance database and they can be identified as such. (9) MAHs are responsible for monitoring the medical literature on their licensed products, and reporting individual cases of suspected adverse reactions reported in literature into EudraVigilance and national safety databases - this is a regulatory requirement. (10) MAHs must perform a literature review of widely used reference databases. They must also monitor the scientific and medical publications in local journals in countries where medicinal products have a Marketing Authorisation. Reports of suspected adverse reactions from the scientific and medical literature, published abstracts from meetings and draft manuscripts, must be reviewed and assessed by MAHs to identify and record possible ICSRs. This category encompasses also data from published case series, non-interventional studies and Patient Support Programmes. All this safety information must be included in the safety databases of the MAHs and made available in EudraVigilance. Databases are checked regularly for potential duplicate reports.

Both spontaneous reports and literature data are filed in the ICH-E2B(R3) format. (11) In this demonstration study, cases described in literature are only considered when they have been filed as ICH-E2B(R3) reports in the databases of EFPIA (European Federation of Pharmaceutical Industries and Associations) partners, EMA or national PV centres.

Other ways to collect information on drug use during pregnancy include TIS, pregnancy registries, and enhanced PV programmes. This type of information sources is mainly covered in the protocol of task **2.5.2** and will therefore not be discussed in detail in this protocol. Only in respect to the development of a tool for measuring and comparing the clinical quality of the data (sub-studies 2 and 3) these sources will be taken into account.

8.3. Relationship between demonstration studies **2.5.1** and **2.5.2**

Part of the **WP2** objectives is to develop a common data model (CDM) with a distributed data approach, allowing for one centrally defined data analysis while the data access providers maintain control of their data and their use. Whereas information from spontaneous reports and literature can already be exchanged using the ICH-E2B(R3) standard (11), automated information exchange from other primary data collections for pregnancy (e.g. TIS, registries and enhanced PV programmes) usually is not yet possible.

Following discussion among the **WP2** members, it was decided to focus demonstration study **2.5.1** on spontaneous reporting systems (SRS) and literature data only. Demonstration study **2.5.2** will consider the description of the other data sources in its protocol. The exceptions are sub-study 2 and 3 (**S2** and **S3**) where a dedicated model for the measurement of clinical quality of pregnancy related data will be developed (**S2**) and where we will study possible differences between various sources (**S3**). In these studies we will not only focus on spontaneous reports and literature data, but also the data sources of demonstration study **2.5.2** (e.g. registry data, data from TIS centres and enhanced PV programmes).

In the demonstration studies described in this protocol (like the studies described in the protocol of demonstration study **2.5.2**) special attention is paid to medicinal products registered for MS. The rationale for this is that 1. several pregnancy registries are in place for MS, both drug-specific and disease specific and 2. enhanced PV programme approaches are in place for several MS drugs. Focus on these drugs with the same indication will enable a better alignment of the outcomes among all data sources and studies in **WP2** as well as **WP1**. Apart from medicinal products indicated for MS, the analyses proposed in these demonstration studies may be extended to other medicinal products and indications when considered relevant (e.g. in the case that the power of the studies is too limited or if the generalisability is at risk).

Out of scope of this demonstration study are studies from the general area of preclinical safety, even if reporting teratogenic or mutagenic effects. Subject areas within the scope of the demonstration projects being coordinated by **WP1** (secondary data collection and evaluation), for instance pharmacoepidemiological datasets like (non-)interventional clinical studies (including Post Authorisation Efficacy Studies (PAES) and Post Authorisation Safety Studies (PASS)). This also holds for studies in respect to breastfeeding and outcomes in the infant that will not be taken into account as well as studies focussing on paternal exposure. Finally, in respect to data from literature, only information from case reports as filed in databases from national PV centres and MAHs will be taken into account. Editorials, meta-analyses and review papers will therefore be excluded.

8.4. Rationale

Tools developed for the analysis of spontaneous reports (e.g. disproportionality or other quantitative and qualitative analyses) have not specifically been designed to analyse safety information related to pregnancy. In order to improve this process, as a first step the nature and content of the currently available data in SRS will be described. Additionally, the presence and absence of data that are relevant for the analysis of exposure during pregnancy will be described, as it could show the potential and need of additional pregnancy fields to the existing ICH-E2B(R3) format. (11) Likewise, the quality of the clinical information described in the available reports is unknown; despite a large number of fields available in the ICH-E2B(R3) model, it is not clear if the information is suitable and of sufficient clinical quality to understand the nature of the events that are described. Existing methods to assess the quality were not designed to specifically assess the quality of pregnancy data, and are therefore not suitable for the assessment of reports regarding pregnancy. (12, 13) A need exists to develop a new dedicated model for the measurement of clinical quality of not only ICSRs, but also the data sources of demonstration study **2.5.2** (e.g. registry data, data from TIS centres and enhanced PV programmes).

Apart from describing the content of the available data sources, a need exists for improvement of currently used signal detection techniques. Observations from daily practice are either published in the form of case reports or submitted as (spontaneous) reports to MAHs, national PV centres and international centres like the EMA and the WHO. These observations have been proven to be a useful tool in the detection of safety information in respect to the safety of registered medicinal products used during pregnancy, but they all have their shortcomings and are also not optimized for pregnancy data. (14-18) Knowledge of which aspects play a

role in the process of generating new signals may facilitate the signal detection process. Additionally, current methods are mostly focussed on associations between a medicinal product and a single adverse drug reaction (ADR) while many ICSRs report two or more ADRs. Therefore, we would like to explore the possibilities of cluster analysis as a possible approach for teratogen signal detection.

9. Research question and objectives

The main objective of demonstration study **2.5.1** is to gain insight into the nature of information on drug exposure during pregnancy from spontaneous reports and literature reports by

1. Describing the nature and content of SRS and literature data sources (**S1**)
2. Creating and validating a dedicated assessment tool for measuring the quality of pregnancy data specifically (**S2**)
3. Describing the quality of reports in SRS, literature, TIS, registries and enhanced PV programme data (**S3**)
4. Assessing predictors of currently used teratogen signal detection techniques in ICSR databases (**S4**)
5. Exploration of cluster analysis as a possible new teratogen signal detection technique (**S5**)

The sub-studies addressing these research questions are depicted in brackets.

10. Research methods

10.1. Sub-study 1: Characterization of pregnancy data in SRS and literature reports

Aims	To describe the nature and content of SRS and literature data sources. - Describe the presence or absence of pregnancy related information in SRS and literature data sources for demonstration drugs - To describe the completeness of important pregnancy related variables in SRS and literature data sources - To describe the presence or absence of specific pregnancy related outcomes in SRS and literature data sources
Leads	Eugène van Puijenbroek, Lareb Yrea Weetink, Lareb
Collaborators	David Lewis, Novartis Other collaborators will be added during the execution of the study
Data sources	PV-report (Lareb), spontaneous and literature reports Argus Safety (Novartis), spontaneous-, literature and solicited (non-interventional studies and patient support programmes) reports

Various types of data sources contain information on the use of medicinal products during pregnancy. However, the way primary data are collected may differ, as well as the nature and content of spontaneous reports and information in literature. Knowledge of these characteristics is needed when data are going to be used and compared with other data sources in a later stage of this project.

Databases containing ICSRs of spontaneous reports and literature data are designed for the assessment and analysis of a wide spectrum of drug related issues (e.g. adverse events), but are not specifically designed to capture information on drug exposure during pregnancy. Consequently, they may lack dedicated fields for important pregnancy-related information. Apart from that, data may be incomplete or stored in the wrong fields.

Data on ICSRs are stored in a predefined format, following the ICH-E2B(R3) guidelines. Previously, a pilot study has been performed on spontaneous reports from Lareb, in which a first overview was made of the completeness of several important ICH-E2B(R3) variables for studying drug use during pregnancy. A comparison was made with the Core Data Elements (CDE), as defined as part of the deliverable of task 2.3. [unpublished data] We concluded that the availability of information varied widely (0.7% for mother's age at last menstrual period (LMP) to 100% for drug name). Non-derived CDE variables (those that could be used to populate the CDE fields directly) had higher availability compared to derived CDE variables (i.e. those variables requiring combining information from ICH-E2B(R3) data fields before they could be used to populate the CDE fields). The source of the report (e.g. health care professional or patient), receive date, origin of report (MAH or other) and several different categories of reported adverse drug reactions (ADRs) during or after pregnancy (i.e. pregnancy complications of the mother, perinatal pathology, major congenital anomalies, maternal ADRs not to pregnancy, and other/undetermined ADRs) were the most relevant predictors for the availability. No data are currently available on the validity of the reports and correctness of the pregnancy related information. This will be determined in task **2.5.5** of the ConcePTION project.

The objective for this sub-study is to describe the nature and content of SRS and literature data sources as recorded in PV databases. This is achieved by describing the presence or absence of pregnancy related information in SRS and literature data sources for selected medicinal products; to describe the completeness of important pregnancy related variables in SRS and literature data sources, and to describe the presence or absence of *specific* pregnancy related outcomes in SRS and literature data sources. These are relevant questions as they can show differences between (e.g. Lareb vs. Novartis) and within (e.g. spontaneous reports vs. literature reports) the existing data sources. It is expected that these differences exist and by showing these

the found differences may point at the additional value of adding dedicated variables for future ICH-E2B versions.

10.1.1. Study design

This study will be a retrospective observational study using spontaneous and literature reports on pregnancies exposed to medicinal products of interest (see PASS information 'active substances', page 1). Spontaneous and literature reports will be collected from databases containing ICSRs, either spontaneous reports or literature cases.

In this sub study, several outcomes will be estimated:

1. The number of pregnancy reports in the total database, stratified per origin (either SRS or literature), per drug of interest (PASS information 'active substances', page 1), and per year.
2. The number of selected variables that are completed (irrespective of the quality or correctness of information actually being provided)
3. The number of selected pregnancy related complications and outcomes

An informal comparison will be made to the results of demonstration project **2.5.2** (registries, TIS, enhanced PV programmes).

10.1.2. Setting

This study will be performed on databases of The Netherlands Pharmacovigilance Centre Lareb (PV-report) and Novartis (Argus Safety). Both databases contain both spontaneous reports and literature reports. Argus Safety additionally contains solicited (non-interventional studies and patient support programmes) reports). Data will be analysed on both data sources separately and an informal comparison of the results will be performed.

All reports collected in the timeframe from inception to the 1st of January 2021 will be eligible for inclusion. From start to finish, the study is expected to take 6 months, consisting of 2 months data collection, 2 months data analysis, and 2 months of writing of the report.

Inclusion criteria

- ICSRs have to be related to drug use during, or within 6 weeks before pregnancy. Selection of ICSRs will be done using the pregnancy flag developed by EMA. [currently in the process of being published]
- ICSRs relate to the study drugs specified in PASS information 'active substances' (page 1 of this protocol) as 'suspected' or 'interacting' drug (not applicable for the outcome "total number of pregnancy reports in the database").
- Coding and assessment of ICSRs has to be completed within the timeframe specified above.

Exclusion criteria

- In case of duplicates only the "master"-report will be taken into account, the underlying reports will be excluded.
- ICSRs regarding drug exposure during breast feeding.
- ICSRs regarding paternal drug exposure

10.1.3. Variables

Selected variables to be assessed for completeness

The variables for the analysis of pregnancy reports are based on the CDEs, defined as 'essential'. [ConcePTION Deliverable 2.3] All essential CDEs will be assessed for whether these fields are completed. The essential CDEs are: individual case identifiers, linked mother/baby case identifier, primary reporter, primary reporter details, initial report date, mother's date of birth, mother's age at LMP, LMP, expected date of delivery (EDD), source of directly-reported EDD, date of end of pregnancy (EOP), prenatal tests, prospective or retrospective data collection (prospective status/true prospective status), drug names, drug start date, drug stop date, drug

indication(s), peri-LMP exposure, trimester 1,2,3 exposure, pregnancy/fetal/neo-natal outcome information (for categories see "pregnancy complications" below), gestational age at EOP, gestational timing of live birth, infant birth weight, infant sex, infant head circumference, death of live born infant, age at death, maternal death, maternal pre-pregnancy medical conditions, and maternal medical conditions arising in pregnancy. At this stage it is not clear if all essential CDE variables can be taken into account. This will be assessed in the early stages of the study. See also the remarks in the introduction section 9.1 of this sub-study on the pilot study that was previously carried out.

Pregnancy complications and outcomes

ICSRs code pregnancy complications (ICH-E2B(R3) field E.i.2.1b) using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.1 (November 2020). (19) In this sub-study the codes assigned will be grouped and categorized by the researchers involved. The categories will be defined as:

- Pregnancy loss (if none of the subgroups is applicable the report will be categorized as live birth)
    - Stillbirth
    - Induced termination
    - Spontaneous abortion
    - Ectopic pregnancy
    - Molar pregnancy
    - Blighted ovum
- Congenital anomalies will be classified according to the EUROCAT classification. (20, 21)
    - Minor (e.g. accessory auricle, sacral pit)
    - Major (e.g. cardiac defects, neural tube defects)
    - Not otherwise specified
    - Other
- Chromosomal defects (e.g. trisomy 21, Turner syndrome)
- Fetal complications (e.g. intrauterine growth restriction)
- Neo-natal complications (e.g. asphyxia, neonatal infection, hypoglycaemia)
- Growth related complications
    - Small for gestational age at delivery
    - Large for gestational age at delivery
    - Fetal growth restriction
- Death of live born infant
- (Neuro)developmental delay (e.g. mental retardation)
- Maternal complications
    - Death
    - Medical conditions arising in pregnancy (e.g. preeclampsia)
    - Complications during/after delivery (e.g. post-partum bleeding, placenta previa)
    - Post-partum complications (e.g. post-partum haemorrhage)
- Maternal none-pregnancy related complication (e.g. injection site inflammation)
- Other
- No adverse pregnancy outcome reported

10.1.4. Data sources

Spontaneous reports, literature reports and solicited reports will be selected from the included data sources: PV-report from the Netherlands PV centre Lareb, and Argus Safety from Novartis.

Ideally, it would be preferable to perform this study on the EudraVigilance database, where all European ICSRs are included. For practical reasons, this is to date not possible. Over the course of this project we will keep studying the possibilities to include data from EudraVigilance in this study.

Pv-report (Lareb)

The Netherlands PV centre Lareb collects and analyses spontaneous reports of suspected adverse drug reactions reported by health care professionals or patients. These reports are recorded in PV-report, the database of Lareb. Spontaneous reports as reported to Lareb (ICSRs) are forwarded to the European EudraVigilance database and the UMC. Cases that have been published in literature are recorded by MAHs and EMA and are also made available in the form of ICSRs in PV-report. (9) PV-report was set up in 1991 (22), and contains over 296000 reports in total as of November 2020. However, this number also included reports not related to pregnancy. Reports regarding drug exposure during pregnancy can be selected using the pregnancy flag developed by the EMA. [to be published over the next months by EMA]

Argus Safety (Novartis)

Novartis Pharma AG is the marketing authorisation holder for a large portfolio of medicinal products, including pharmaceuticals, biological, advanced medicinal therapeutic products and gene therapies. In accordance with European legislation, Novartis is obliged to collect, manage, and submit individual reports of suspected adverse reactions to its authorized medicinal products. Individual case safety reports are recorded in Argus Safety (version 8.1.2.3). Argus Safety comprises Argus Core, Argus Affiliate and Argus J. It is a commercially available application used for data collection of adverse events and reporting of adverse event information. It is a web-based, validated database used for the tracking and processing of all safety cases received by Novartis. Argus Safety was first deployed in 2003, and contains over 8 million case reports as of November 2020. Reports of exposure to medicines during pregnancy can be selected using validated SQL scripts based on a pregnancy flag and additional ICH-E2B(R3) fields. Pregnancy information from Argus Safety has been published (e.g. Geissbühler et al., 2018). (23)

10.1.5. Study size

All cases fulfilling the inclusion and exclusion criteria will be considered for this sub-study. Based on the pilot study previously performed in PV-report, it is estimated that in total (not limited to the demonstration drugs) approximately 4000 spontaneous reports about drug use during pregnancy are available. Upon limitation to the demonstration drugs specified in PASS information 'active substances' (page 1 of this protocol), this number will decrease considerably. It is not possible to have too low numbers in the first part of this study, as the aim is to describe what is available. If the conclusion of the first part is that the numbers are very low, all pregnancy cases (independent of demonstration drugs) could be included for the analysis of availability of variables, or multiple data sources could be combined.

For Argus Safety, 4,859 case reports for fingolimod, 331 for beta interferon, 679 for mycophenolate, and 757 for azathioprine are available. Overall, an estimated >30,000 reports of exposure to medicinal products during pregnancy in association with authorised medicinal products are included in the database (predominately exposures to the mother, but some reports of paternal exposure exist).

10.1.6. Data analysis

Data selection and analysis for the Lareb data will be performed using Microsoft SQL server 2017 and R studio version 2016. First, all pregnancy reports will be selected from the database using the pregnancy flag developed by EMA. [to be published] Descriptive statistics (absolute numbers, percentages) will be used to describe the total number of reports available, and the stratified numbers per origin (e.g. SRS, literature, MAH) and the stratified numbers per drug of interest. For the different drugs of interest, the numbers of reports reported will also be stratified per year, and time series will be used to show the trend of reporting over time (x-axis time, y-axis number of reports).

For all essential CDE variables (as described in section 9.1.3 'variables') it will be described whether the variables are available and if there is not specific variable, whether the information is described in other variables. Thereafter, the completeness of these variables as "IS NOT NULL" will be determined. Variables that have to be derived from other variables will be derived when possible, and thereafter availability of the information in the variables (how well they are filled) is depicted in percentages. In case a variable can be reported and derived, the availability is depicted for both methods separately and also combined. Results will be compared between different strata of drugs and between reports of different origins (e.g. national PV centre, MAH, literature). Reasons for missing values will be discussed.

The reported pregnancy complications will be categorized as described in section 9.1.3 'variables'. Descriptive statistics (absolute numbers, percentages) will be used to describe the number of reports per category, which will again be stratified for origin, demonstration drug and year. Assigned causality will not be taken into account in this study.

10.1.7. Limitations of the research methods

The criteria for determining whether a report is related to pregnancy or not will be based on the EMA pregnancy flag. [to be published] If, during the course of the study, an additional criterion is found to decrease misclassification, this criterion will be taken into account as well.

It is often unclear why variables are missing. It can therefore never be concluded (based on this study design) whether or not completeness of a variable is too low, it is only possible to objectively describe percentages of completeness.

10.1.8. Other aspects

Not applicable

10.2. Sub-study 2: Development and validation of a dedicated assessment tool for measuring the quality of pregnancy data

Aims	Creating, validating and testing a new optimized assessment tool for the quality of pregnancy data specifically
Leads	Eugène van Puijenbroek, Lareb Yrea Weetink, Lareb
Collaborators/tasks	Michael Stellfeld, Novo Nordisk Ursula Winterfeld, STIS Other collaborators will be added during the execution of the study
Data sources	PV-report (Lareb), spontaneous and literature reports STIS, TIS UKTIS, TIS pREGnant (Lareb), registry Gilenya Pregnancy Registry (Novartis), registry Argus Safety (Novartis), PRIM cohort enhanced PV programme

In order to assess the relationship between exposure to a drug and a reported adverse event in a reliable way, detailed clinical information of high quality is needed. Reports in which clinical information is well documented are more likely to contribute to a reliable assessment of drug safety signals since they can provide a more precise statement about the causal relationship between drug use and pregnancy outcome. However, the type and level of detail that is needed to assess the quality of a case may vary, depending on the nature of the reported event, the patient, her underlying conditions, and treatment. This information, that is clinical in nature, may vary between the reports. As an example, for a report concerning an early miscarriage in an otherwise healthy woman, other information will be needed compared to a report concerning a major congenital malformation in which the pregnant woman had been using various drugs and is known to have multiple underlying conditions.

To give an impression of the quality of information, some previous studies explored the completeness of reported information. An example is the VigiGrade completeness score that has been developed to measure the technical completeness of information provided in spontaneous reports of ADRs in general (so not only related to pregnancy), based on which specific fields are filled. (12) However, the presence or absence of information does not automatically reflect the level of clinical information, as required fields may be filled with inadequate, nonspecific, or ambiguous information. Relevant information from a clinical perspective may still be lacking, making it difficult to measure the actual clinical quality and to make a proper assessment.

Oosterhuis et al. (13) developed a tool for documenting the clinical quality taking into account whether information was indeed required for a proper assessment (ClinDoc). It includes four domains for which scores are assigned: the ADR, chronology of the ADR, the suspected drug(s) and finally patient characteristics. The final score categorizes reports into four categories: excellent, well, moderately or poorly documented. The outcome was compared with expert panel judgement. (13) Although a valuable tool for the assessment of the quality of ADRs in general, no existing quality assessment tools are available that are specifically designed for pregnancy data. Especially in these circumstances, different variables are of importance, compared to non-pregnancy data.

10.2.1. Study design

This study will create and validate a new tool (PregDoc) specifically optimized for assessment of the quality of pregnancy data. The tool should be useful for all sources of pregnancy data, i.e. spontaneous reports, literature reports, TIS data, registry data and enhanced PV programme data.

As a first step, a qualitative study will be performed based on interviews among a minimum of 5 experts from the fields of pharmacovigilance, teratology and clinical genetics. These interviews will be based on a set of

proactive and reactive 'probes' to guide the course of the interviews. They will be audio-recorded and afterwards transcribed and coded. Based on the collected data, a first model will be developed based on selected elements that play a role in defining the clinical quality. Examples of elements to be considered are the availability of information on the timing of exposure during pregnancy or sufficient information on the medical history to assess the case or outcomes of prenatal diagnostics. The way to assess and describe this information will be described as part of this first step and will be done in collaboration with the expert group that was interviewed.

This first version of the quality tool will be subject to cognitive interviewing (maximum of 10 experts) to gain insight into the way PV assessors understand the questions and how they interpret the answer options, highlighting any ambiguities. Interviews will be transcribed to identify issues where the interviewee had difficulties understanding and answering the quality items. The identified issues will be coded according to a dedicated system containing the following categories: comprehension, interpretation and logical/structural problems. When needed, amendments will be made.

As a next step, the quality tool will be validated with a selected number of 90 representative reports/cases from 6 different data sources (15 per source; spontaneous reports, literature reports, European Network of Teratology Information Services (ENTIS, Switzerland and UK), 2 registries, enhanced PV programme) that will be selected by the expert panel that is involved in the first version of the pregnancy quality assessment tool. This training dataset will be used for the weighing of the parameters of PregDoc. Assigned scores will be compared with the quality grading assigned by the expert group based on global introspection and sensitivity and specificity will be calculated. The performance of the model will be calculated using the Area Under the Curve (AUC) of the Receiver Operating Curve (ROC). (16, 24) When needed, the weighing factors of the selected elements will be adjusted to optimise the performance. The time in which the experts completed the clinical assessment tool, their ability to answer all questions and interrater variability will be evaluated.

As a final step we will test the performance of the revised version of the questionnaire. A minimum of 10 experts will use the tool to assess another 90 reports/cases (15 per data source). Each expert will assess a selection of the cases in order to reduce the workload. All cases will be assessed by 2 experts, and if the results are inconsistent a third expert will decide on the final grading. This test dataset is used to provide an unbiased evaluation of PregDoc fit on the training dataset. The performance of the quality tool will be expressed in terms of specificity and sensitivity.

Although measuring different characteristics, the score of the clinical quality as determined with the newly developed PregDoc tool will be compared to outcomes of the VigiGrade model. For those cases forwarded to the Uppsala Monitoring Centre (i.e. ICSRs and ICSR-literature cases), grading according to the VigiGrade system will take place automatically. The corresponding gradings will be retrieved at the UMC via the Vigilyze system. (12) For cases from non-ICSR data sources (TIS, registries, enhanced PV programmes) the VigiGrade score will not be compared in this study. Comparison of the assigned VigiGrade score and the PregDoc score will be performed informally.

10.2.2. Setting

All data collected in the timeframe from inception to the 1st of January 2021 will be eligible for inclusion. 180 representative reports/cases from 6 different data sources (2 times 15 per source; spontaneous reports, literature reports, ENTIS, 2 registries, enhanced PV programme) will be selected by the expert panel that is involved in the first version of the pregnancy quality assessment tool. These 180 reports will be equally divided in a validation and a test-set. The cases will be selected from PV data sources of the Netherlands PV centre Lareb, TIS centres (Switzerland, UK) and from EFPIA partners (Novartis). Selection criteria for these cases will be based on the first round of interviews with the expert panel. Sufficient variety in quality and clinical picture is important.

The experts that will take part in the interviews and final assessment will be recruited from ENTIS centres and EFPIA partners collaborating in the ConcePTION project. This study is expected to take 6 months, of which 2 months for the design of the new tool, 2 months for the validation, and 2 months of publishing the results.

For all included data sources the following in- and exclusion criteria apply, wherever not already applicable:

Inclusion criteria

- ICSRs/cases have to be related to drug use during, or within 6 weeks before pregnancy. Selection of ICSRs will be done using the pregnancy flag developed by EMA. [to be published]
- Coding and assessment of ICSRs/cases has to be completed within the timeframe specified above.

Exclusion criteria

- In case of duplicates only the "master"-report will be taken into account, the underlying reports will be excluded (applicable to ICSRs).
- ICSRs/cases regarding drug exposure during breast feeding.
- ICSRs/cases regarding paternal drug exposure

Because it is preferable to develop a tool that can be used on all pregnancy cases, this study will not be limited to exposures to the study drugs as specified in PASS information 'active substances' (page 1 of this protocol).

10.2.3. Variables

Assessment new tool

For the development and assessment of the new tool, the variables necessary will be determined in the first part of this study.

Assessment VigiGrade

VigiGrade has been developed for defining the quality of spontaneous reports; the measure mainly focussed on the administrative quality and is not designed for a specific clinical situation. Grading has been done automatically when ICSRs have been uploaded to VigiBase, the database of the Uppsala Monitoring Centre (WHO collaborating centre for international drug monitoring), and can be retrieved by national PV centres (among which Lareb). No additional variables are needed for assessment of the ICSRs. (12)

10.2.4. Data sources

PV-report (Lareb), spontaneous and literature reports

The Netherlands PV centre Lareb collects and analyses spontaneous reports of suspected adverse drug reactions reported by health care professionals or patients. These reports are recorded in PV-report, the database of Lareb. Spontaneous reports as reported to Lareb (ICSRs) are forwarded to the European EudraVigilance database and the UMC. Cases that have been published in literature are recorded by MAHs and EMA and are also made available in the form of ICSRs in PV-report. (9) PV-report was set up in 1991 (22), and contains over 296000 reports in total as of November 2020. However, this number also included reports not related to pregnancy. Reports regarding drug exposure during pregnancy can be selected using the pregnancy flag developed by the EMA. [to be published over the next months by EMA]

STIS, TIS

The Swiss Teratogen Information Service (STIS) is dedicated to providing healthcare professionals evidence-based information about medications and other exposures during pregnancy and breastfeeding in Switzerland. Furthermore, STIS staff collects patient data both during initial contact and after a follow-up period covering pregnancy and breastfeeding outcome. STIS is a member of ENTIS. The STIS database consists of pregnant and breastfeeding women that have been reported by a healthcare professional. Maternal characteristics (age, tobacco use, alcohol consumption, medical and obstetric history) and information on medication exposure (indication, timing in pregnancy, duration, dose and concomitant medication) are routinely collected at initial contact and updated through the follow-up process. Follow-up pregnancy and breastfeeding outcome information is collected via postal questionnaire, which is sent to the initial enquiring healthcare professional at first contact and shortly following the expected date of delivery. Information reported at first contact can be corrected when providing outcome information. For pregnancies, collected follow-up data includes delivery

mode, pregnancy outcome, gestational age at delivery, birth weight, length, head circumference and congenital anomalies. Data collection started in 1975 and is continuous with 10506 cases registered in the database (17-11-2020).

UKTIS, TIS

UKTIS are commissioned by Public Health England to be the sole dedicated UK provider of evidence-based information regarding the fetal risks associated with pharmacological and other potentially toxic exposures which arise during pregnancy or the peri-conceptual period. UKTIS are also funded to perform national surveillance of known and emerging human teratogens across the United Kingdom. The UKTIS (United Kingdom) prospective dataset consists of exposed pregnancies that have been reported to the service for counselling advice by a general practitioner (family doctor), general practice nurse or pharmacist, community midwife, hospital-based midwife, obstetrician or other secondary care medical specialist (physician or pharmacist), or a medicines information pharmacist. Follow-up pregnancy outcome information is collected via postal questionnaire, which is sent to the initial enquiring healthcare professional, shortly following the expected date of delivery that was reported at first contact. Outcomes not provided back to the service by 24 weeks after the estimated date of delivery are considered lost to follow-up. All clinical data, including environmental exposures (medicines, occupational chemicals, radiological, biological and social/recreational substances), obstetric history, relevant medical history, and pregnancy outcome information, are reported to UKTIS by healthcare professionals through a standardized data collection procedure. Information reported to UKTIS at first contact can be corrected by the enquirer when providing pregnancy outcome information. In house data consistency calculations are applied for the maternal age at initial reporting to UKTIS, and the expected date of delivery. Clinical advice provided to healthcare professionals is checked by a senior member of the service. A random sample of clinical enquiries are periodically selected and audited for accuracy and completeness of data recording by the Assistant Head of UKTIS. Pregnancy exposure and outcome information collected by UKTIS is recognised by the Caldicott Advisory Panel of Public Health England to be vital for the identification of novel human teratogens, in recognising trends and risks associated with teratogenic exposures, protecting public health from the impact of human teratogens, and monitoring adverse reactions to vaccine and medication exposures which arise during or around pregnancy. Accordingly, UKTIS data collection and processing is covered by section 251 of the National Health Service Act 2006 and Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002 (Public Health England Approval Reference Number: 13091). Under UK law, the NHS Act 2006 and Regulation 3 of The Health Service Regulations 2002 enables the common law duty of confidentiality to be temporarily lifted. This enables healthcare professionals contacting UKTIS to disclose patient identifiable information (name, DoB, NS number) to UKTIS without being in breach of the common law duty of confidentiality. Under these regulations, UKTIS are obligated to remain compliant with all relevant legal data protection and information governance obligations. Data collection started in 1983 and is continuous with 13,840 prospective pregnancy outcomes registered in the database (04-12-2020). UKTIS are not obligated to report any suspected adverse events or reactions to the national regulatory agency (Medicines and Healthcare products Regulatory Agency).

pREGnant (Lareb), registry

The Dutch Pregnancy Register pREGnant has been set up to obtain insight into medication use among pregnant and breastfeeding women and into the potential effects on maternal and fetal health. It was launched on April 1st, 2014. Data is collected through six web-based questionnaires filled in by participating women. The data comprise current pregnancy, obstetric history, maternal lifestyle, health and medication use, delivery and infant health. The first questionnaire is sent as early in pregnancy as possible, followed by a second questionnaire in gestational week 17 (unless enrolment was after week 20, then the second questionnaire is skipped). Questionnaires 3 through 6 are respectively sent in gestational week 34, and 8, 26 and 52 weeks after the expected date of birth. All pregnant women of 18 years and older, who are proficient in the Dutch language and have access to internet are eligible for inclusion. All data is collected prospectively, although women can enrol at any time throughout the entire pregnancy. (6) As of 14 October 2020, 4856 women were enrolled.

Gilenya Pregnancy Registry (Novartis), registry

The Gilenya® Pregnancy Exposure Registry was launched in May 2011 to prospectively collect safety data on maternal, fetal and infant outcomes associated with fingolimod exposure during pregnancy and up to 8 weeks before the LMP. At time of enrollment the patient needs to be pregnant, have a diagnosis of MS, was exposed during pregnancy or up to 8 weeks before LMP and signed an informed consent form. No exclusion criteria are applied. Cases are considered prospective if at the time of enrollment, no prenatal testing had been done and pregnancy outcome was not known. Cases are considered retrospective if prenatal tests had been performed at enrollment, while the patient is still pregnant, regardless of their results. The primary outcome measure is the prevalence of major congenital malformations, defined as any structural defects with recognized surgical, medical, or cosmetic importance as per the European Surveillance of Congenital Anomalies (EUROCAT) guidelines, while anomalies were qualified as minor if they were of no serious medical or cosmetic consequence to the child. Other outcome measures included miscarriages, elective abortions, stillbirths, neonatal deaths (counted in live births), and term and preterm live births with or without congenital malformations. (23)

Congenital malformations are judged by external adjudicators (qualified pediatrician, clinical geneticist, teratologist, pediatric neurologist, nephrologist, toxicologist, or clinical pharmacologist), according to the EUROCAT or other guidelines. In cases of discrepancy, cases are classified as the more severe of the two adjudications.

As of 28 Feb 2020, 271 women were enrolled (9 were excluded from analysis due to protocol deviations). 262 women were analysed in the most recent interim report. 177 prospective and 85 retrospective reports.

Argus Safety (Novartis), PRIM cohort enhanced PV programme

The PRIM (**PR**egnancy outcomes **I**ntensive **M**onitoring) method uses the Argus Safety database (see section 9.1.4. Data Sources) to access prospective pregnancy cases. (25) This method was first developed for fingolimod, a treatment option for MS, due to slow enrolment in the company pregnancy registry. The aim of PRIM was to enhance the process of pregnancy data collection and improve data quality, and in particular to enable estimation of the proportion of major congenital malformation and other pregnancy outcomes. In order to do this, spontaneous reports of maternal exposure to fingolimod in pregnancy or in the eight weeks immediately before the last menstrual period of patients not enrolled in the pregnancy registry were identified. Follow up checklists were sent at four time points:

1. At the time of the initial pregnancy report;
2. At the end of pregnancy;
3. When the infant attained 3 months of age, and
4. When the infant was 12 months old

These questionnaires focused on core data required for derivation of programmed analyses. From 01 Mar 2014 to 28 Feb 2018, a total of 831 prospective maternal exposures with 843 infants were reported, with fetal outcomes reported in 459/843 (54.4%) of those infants. This enabled the calculation of proportions of pregnancy cases with the main pregnancy outcomes and of fetal cases with malformation. The number of reported pregnancies was significantly higher in PRIM than in the registry, showing that structured use of pharmacovigilance data enabled more rapid assessment of risks of maternal drug exposure. (25)

10.2.5. Study size

A minimum of 180 cases related to pregnancy and drug use will be studied (2 times 15 per data source), divided in a set of 90 cases for validation and a set of 90 cases for testing. These numbers should comply with the rule of thumb for prediction models: a minimum of 10 cases should be included per category in the model. Therefore, these numbers might increase after the first round of interviews in the case that more than 9 categories are included. A minimum of 5 experts will participate in the first interviews, a minimum of 10 in the validation, and a minimum of 10 for the testing phase. Experts may be involved in various phases, given the limited number of experts available and in order to reduce the workload for the individual experts.

10.2.6. Data analysis

The development and validation of the dedicated tool to measure the quality of pregnancy related reports/cases will be qualitative in nature. The various steps and calculations that will be conducted are described in section 9.2.1. Study design. Statistical analyses will be performed in R studio version 2016.

10.2.7. Limitations of the research methods

One limitation is that a relatively small number of PV assessors will be interviewed during the content validation. Although it is to be expected that a small group will reveal most critical problems, we cannot rule out the fact that a larger number will be able to reveal additional items. In addition, we will not use a development and test set, as underlying conditions will probably be different.

10.2.8. Other aspects

Not applicable

10.3. Sub-study 3: Quality of information

Aims	To study the quality of information present in SRS, literature reports, TIS data, registry data, and enhanced PV programme data on drug exposure during pregnancy. - To compare the quality of the different data sources - To study the influence of the data source itself and the nature of the pregnancy outcome on the final quality score
Leads	Eugène van Puijenbroek, Lareb Yrea Weetink, Lareb
Collaborators/tasks	David Lewis, Novartis Michael Stellfeld, Novo Nordisk Ursula Winterfeld, STIS Yvonne GeissBühler, Novartis Other collaborators will be added during the execution of the study
Data sources	PV-report (Lareb), spontaneous and literature reports STIS, TIS UKTIS, TIS pREGnant (Lareb), registry Gilenya Pregnancy Registry (Novartis), registry Argus Safety (Novartis), PRIM cohort enhanced PV programme

Sub-study 3 **(S3)** aims to build on **S1** by assessing the quality of information of spontaneous and literature reports by means of the new grading scale developed in **S2** (PregDoc). Additionally, information from TISs, registries and enhanced PV programmes will be included (although the majority of studies on the latter three sources are described as part of demonstration study **2.5.2**).

The assessment and analysis of information on the use of drugs during pregnancy in the post-marketing phase may influence the benefit/risk balance of registered products. Information about the safety of drugs can be based on various sources. As an example, spontaneous reports can be submitted both by consumers and healthcare professionals to national PV centres or MAHs. Information from literature or studies will be also be filed as ICSRs in these databases. (9) Also information from TIS centres, registries and other data collections play an important role in the evaluation of the safety of the use of medicines during pregnancy. It is obvious that the quality of information in these sources may vary, not only depending on the nature of the source, but also on the reporter, the drug that has been used, and the reported outcome. (12, 13, 16) For this reason, knowledge of the quality of the information present in the various data sources is important.

As described before in **S2**, in order to assess the relationship between exposure to a drug and a reported adverse event in a reliable way, clinical information is needed. (13, 16) Reports in which clinical information is well documented are more likely to contribute to a reliable assessment of drug safety signals since they can provide a more precise statement about the causal relationship. However, the type and level of detail that is needed to be able to perform assessment of a case may vary between reports, depending on the nature of the reported event, the patient and underlying conditions and treatment.

In this sub-study we will analyse potential differences in the quality of the information available in SRS, literature reports, TIS data, registry data, and enhanced PV programme data, that may contribute to the detection of new safety signals. In **S2** a new tool was designed to assess the quality of pregnancy data specifically (PregDoc) that will be used to assess the quality of several data sources containing pregnancy data.

The primary aim of this sub-study is to study the quality of information present in SRS, literature reports, TIS data, registry data, and enhanced PV programme data on drug exposure during pregnancy. The secondary aim is to informally compare the quality of the different data sources, and to study the influence of the data source itself and the nature of the pregnancy complication or outcome on the final quality assessment. No a priori

hypothesis has been included as no assumptions can be made about the quality of the information in the different data sources.

10.3.1. Study design

This study will be a retrospective observational study into the quality of information of ICSRs concerning spontaneous reports and information from literature (as ICSRs in PV database), and TIS data, registry data, and enhanced PV programme data. 6 data sources are selected: spontaneous reports, literature reports, ENTIS data (Switzerland, UK), 1 or 2 registries (Lareb and/or Novartis), and an enhanced PV programme. For all individual sources, forty cases will be selected randomly for quality assessment (240 in total). The two included TIS centres will each provide 20 cases to make a total of 40 cases for the included ENTIS data. These cases will be made available in a structured format that contains the information deemed necessary for the use of the tool for quality assessment. When applicable, cases will be anonymized and potential suggestions that may point at the origin of the cases will be removed, in order to blind the researchers for the reporting source of the case.

Assessment of the cases will be performed by two researchers who will be selected during the course of the study. They will appraise the quality of the reports independently using the PregDoc method designed in **S2**. Both researchers will provide a grading of the quality as described in this approach. Based on the results of **S2**, classes representing various levels of quality will be determined. For those cases in which the assigned quality classes of the researchers differ, both assessors will reconvene and discuss their grading, in order to come to an agreement. In case agreement cannot be reached, a third researcher will assess the case and decide which value (and thus quality class) will be appointed.

10.3.2. Setting

For all six sources (SRS, literature, ENTIS, 1 or 2 registries, enhanced PV programme) 40 cases (240 in total) will be selected randomly for quality assessment. All data collected in the timeframe between 1-1-2016 and 1-7-2021 will be eligible for inclusion.

From start to finish, the study is expected to take 9 months, consisting of 4 months selection and anonymization, 3 months of assessment, and 2 months of analysis and writing of the report.

In- and exclusion criteria

The same in- and exclusion criteria as described in **S2** (section 9.2.2. Setting) will be taken into account. Additionally, all ISCRs/cases have to relate to the study drugs specified in PASS information ‘active substances’ (page 1 of this protocol) as ‘suspected’ or ‘interacting’ drug.

10.3.3. Variables

For the assessment of the PregDoc score, variables as described in the deliverable of **S2** will be used.

10.3.4. Data sources

The same data sources as described in **S2** (section 9.2.4. Data sources) will be used.

10.3.5. Study size

For each of the six data sources 40 cases (40 in total) will be randomly selected. Since no data on the quality of these type of reports exist yet, it is not possible to make a formal power calculation.

10.3.6. Data analysis

Per data source 40 cases (240 in total) will be randomly selected by means of random number generators. Based on **S2** important categories will be selected as additional selection criteria to make sure an even selection is made between the data sources, if necessary.

After converting the cases to the right format based on the data fields needed for assessment with PregDoc, the selected cases will be anonymised and identifiers that may reveal the origin of the cases will be removed, in

order to blind the assessors for the reporting source of the case. After anonymisation the selected cases will be sent to one researcher, who will blind, combine and randomise the data of the different sources.

In order to reduce workload, cases will be randomly divided over researchers who are blinded for the source of the cases. They will appraise the quality of the reports independently using the PregDoc method developed and validated in **S2**. All cases will be assessed by two experts, who will both provide a grading of the quality as described in this approach. Based on the results of **S2**, classes representing various levels of quality will be determined.. For those cases in which the assigned quality classes of the researches differ, both assessors will reconvene and discuss their grading, in order to come to an agreement. In case agreement cannot be reached, a third researcher will assess the case and decide which value (and thus quality class) will be appointed.

Differences in quality between various data-access providers will be specified, stratified for spontaneous reports, literature reports, ENTIS data, the two registries (Lareb, Novartis), or enhanced PV data. Depending on the documentation tool developed in **S2** differences in partial aspects of the clinical quality can be specified. An informal comparison between the different data sources will be done, and the influence of the data source and the nature of the pregnancy complication/outcome on the final quality assessment will be analysed via linear regression modelling. Statistical analyses will be performed in R studio version 2016.

10.3.7. Limitations of the research methods

In this study a relatively small number of cases will be studied, given the heterogeneity of data. Since the analysis will mainly focus on drugs selected as demonstration drugs, generalisability of the outcomes will be limited.

10.3.8. Other aspects

Not applicable

10.4. Sub-study 4: What predicts a signal?

Aims	Assessing predictors of currently used teratogen signal detection techniques in ICSR databases
Leads	David Lewis, Novartis
Collaborators/tasks	Amalia Alexe, Novartis Bita Rezaallah, Novartis Eugène van Puijenbroek, Lareb Michael Stellfeld, Novo Nordisk Yrea Weetink, Lareb Other collaborators will be added during the execution of the study
Data sources	PV-report (Lareb), spontaneous and literature reports Argus Safety (Novartis), spontaneous-, literature and solicited (non-interventional studies and patient support programmes) reports

The aim of the analysis of Individual Case Safety Reports (ICSRs) in pharmacovigilance is the early detection of previously unrecognized adverse reactions, and obtaining information on new aspects of previously known associations (those listed in the company core safety information, CCSI). Signal detection historically relied on the analysis, by trained PV assessors of the content of the ICSRs submitted by health care professionals (HCPs) and consumers. (26) Potential signals will undergo a detailed analysis in which other reports in the database and other sources of information are taken into account. Over the past decades, the number of ICSRs increased considerably, so additional approaches were developed to support the signal detection process. An example is disproportionality analysis, which enabled the screening of large amounts of data. Nevertheless, it is obvious that this approach has its shortcomings from an epidemiological point of view. (16-18) For this reason, the outcome of disproportionality analysis will be combined with other information such as the number of reports in the database, the source of the reports, if an association is labelled or not, or a possible pharmacological mechanism. This information will be used to select those associations that should be studied in more detail to assess if they represent a true safety signal. Although the contribution of various elements for the detection of safety signals in pharmacovigilance in general has been studied in detail (16), it is not clear which factors may contribute to the actual selection of safety signals in case of drug exposure during pregnancy. **S4** aims to improve the teratogen signal detection approach by studying which factors contribute to the selection of safety signals in cases related to teratogen signal detection.

10.4.1. Study design

This study will be a retrospective observational study determining predictors for medicinal product-ADR signal generation. The model will be based on ICSRs in the ICH-E2B(R3) format. (11) All included medicinal product-ADR associations expressed as MedDRA/ATC with a minimum of 3 reports will be assessed for whether they resulted in a signal or not. With all resulting medicinal product-ADR associations a prediction model will be made.

For this study, a reliable gold standard needs to be developed. An option is to use a reference list with associations with positive and negative controls that are based on the product information of the demonstration drugs (see PASS information 'active substances', page 1). Another option might be to take the moment on which association were highlighted as positive controls into account, since it cannot be ruled out that at this moment the number of reports were increased due to notoriety bias. It has to be checked if this information is only available to the MAHs of the product or also in the public domain. This may limit the amount of drugs that can be taken into account in this study. True signals could be based on company decision points (e.g. decision recorded by Medical Safety Review Committee) or on regulatory actions taken (e.g. CCSI or SmPC change).

10.4.2. Setting

This study will be performed on databases of The Netherlands Pharmacovigilance Centre Lareb (PV-report) and the Novartis PV database (Argus Safety). Both databases contain both spontaneous reports and literature reports; in addition the company database includes solicited reports (from non-interventional studies and patient support programmes). Data will be analysed on all data sources separately and an informal comparison of the results will be performed.

All reports collected in the timeframe from inception to the 1st of January 2022 will be eligible for inclusion. From start to finish, the study is expected to take 6 months, consisting of 2 months data collection, 2 months data analysis, and 2 months of writing of the report.

In- and exclusion criteria

The same in- and exclusion criteria as described in **S1** (section 9.1.2. Setting) will be taken into account.

10.4.3. Variables

The outcome variable of the prediction model will be whether the association between the medicinal product and the ADR led to a signal or not. Expected possible predictors are aspects that are usually involved in the process of signal detection. The possible predictors that will be included are: absolute number of ICSRs on the drug-ADR combination, lower limit of the 95% confidence interval (LL) of the reporting odds ratio (ROR) with the full database as the comparison group, the LL of the ROR with pregnancy reports as a comparison only, the percentage of the ICSRs on the drug-ADR association reported by HCPs, the percentage of ICSRs on the drug-ADR association derived from MAHs, whether the drug-ADR association is classified as serious, the time the drug has been on the market, the presence or absence of important variables as described in **S2**, the presence or absence of important variables as described in **S1**, the type of pregnancy complication/outcome (see classification in section 9.1.3. Variables). Additionally, it could be of interest to see how other variables could be taken into account in this study. An example are the Bradford Hill criteria, that play a role in the assessment of the causal relationship between the drug and the reported event. (26) It is possible that some of these criteria can also be used as predictors in this study. The feasibility of applying these predictor variables will be considered in more detail over the course of the project. The final choice for number and type of variables will after all depend on the number of associations that will be studied.

10.4.4. Data sources

Both spontaneous reports and literature reports will be selected from the included data sources: PV-report from the Netherlands PV centre Lareb, and Argus Safety from Novartis.

Ideally, it would be preferable to perform this study on the EudraVigilance database, where all European ICSRs are included. For practical reasons, this is to date not possible. Over the course of this project we will keep studying the possibilities to include data from EudraVigilance in this study.

Pv-report (Lareb)

The Netherlands PV centre Lareb collects and analyses spontaneous reports of suspected adverse drug reactions reported by health care professionals or patients. These reports are recorded in PV-report, the database of Lareb. Spontaneous reports as reported to Lareb (ICSRs) are forwarded to the European EudraVigilance database and the UMC. Cases that have been published in literature are recorded by MAHs and EMA and are also made available in the form of ICSRs in PV-report. (9) PV-report was set up in 1991 (22), and contains over 296000 reports in total as of November 2020. However, this number also included reports not related to pregnancy. Reports regarding drug exposure during pregnancy can be selected using the pregnancy flag developed by the EMA. [to be published over the next months by EMA]

Argus Safety (Novartis)

Novartis Pharma AG is the marketing authorisation holder for a large portfolio of medicinal products, including pharmaceuticals, biological, advanced medicinal therapeutic products and gene therapies. In accordance with

European legislation, Novartis is obliged to collect, manage, and submit individual reports of suspected adverse reactions to its authorized medicinal products. Individual case safety reports are recorded in Argus Safety (version 8.1.2.3). Argus Safety comprises Argus Core, Argus Affiliate and Argus J. It is a commercially available application used for data collection of adverse events and reporting of adverse event information. It is a web-based, validated database used for the tracking and processing of all safety cases received by Novartis. Argus Safety was first deployed in 2003, and contains over 8 million case reports as of November 2020. Reports of exposure to medicines during pregnancy can be selected using validated SQL scripts based on a pregnancy flag and additional ICH-E2B(R3) fields. Pregnancy information from Argus Safety has been published (e.g. Geissbühler et al., 2018). (23)

10.4.5. Study size

All cases fulfilling the inclusion and exclusion criteria will be considered for this sub-study. Based on the pilot study previously performed in PV-report (described in **S1**), it is estimated that in total (not limited to the demonstration drugs) approximately 4000 spontaneous reports about drug use during pregnancy are available. Upon limitation to the demonstration drugs specified in PASS information 'active substances' (page 1 of this protocol), this number will decrease considerably.

For Argus Safety, 4,859 case reports for fingolimod, 331 for beta interferon, 679 for mycophenolate, and 757 for azathioprine are available. Overall, an estimated >30,000 reports of exposure to medicinal products during pregnancy in association with authorised medicinal products are included in the database (predominately exposures to the mother, but some reports of paternal exposure exist).

If the numbers are too low to perform this study, multiple data sources could be combined, the study could be expanded to inclusion of other drugs, and/or pre-signals could be included as an outcome.

10.4.6. Data analysis

Data will be collected using Microsoft SQL Server 17, and data analysis will be performed using R studio version 2016. All drug-ADR associations of the demonstration drugs mentioned in PASS information 'active substances' (page 1 of this protocol) consisting of at least 3 reports that are eligible for inclusion will form the basis for the development of the prediction model. A multivariable logistic regression based prediction model (outcome signal yes or no) will be developed and validated using the list with validated safety signals as the gold standard for the outcome.

First, contingency tables will be made in order to have a first look at the possible predictors of safety signals. This will be done for all predictors (as described in section 9.4.3. Variables). As a second step, univariable logistic regression models will be made with the same predictors. Results of the models will be reported in table format including the prediction values, p-values and 95% confidence intervals.

In the case that more of the tested predictors seem to be relevant from the contingency tables and univariable models, multivariable logistic regression models will be made. Backwards step-wise selection will be used with $P < 0.05$ (Wald test) as a selection criterion. For all possible predictors the variance inflation factor will be calculated to investigate collinearity, using 4 as a cut-off value. In case collinearity is found, this will be reported and only one variable will be included in the prediction model. Results of the models will be reported in table format including the prediction values, p-values and 95% confidence intervals.

Performance of the model will be calculated using the Area Under the Curve (AUC) of the Receiver Operating Curve (ROC). (16, 24) It should be noted that, although its' name may suggest differently, the prediction model that will be developed is not a substitute for the clinical, pharmacological and teratological analysis of the underlying cases. This approach only allows for an initial impression of the performance of the various candidate predictor variables in terms of sensitivity and specificity and their combined performance versus the gold standard that will be applied in this study.

If possible, the models will be validated on other (larger) datasets, such as EudraVigilance. (7)

10.4.7. Limitations of the research methods

A limitation of this study is the risk of bias due to selective reporting. Because PV databases contain well established associations, it is reasonable to assume that these associations are reported more frequently than unknown associations, which may influence the predictors in the model. Another limitation is that predictors will be assessed, but not tested in a separate dataset. Since a limited number of drugs is used, generalisability will be limited.

10.4.8. Other aspects

Not applicable

10.5. Sub-study 5: Cluster analysis as a possible new teratogen signal detection technique

Aims	Exploration of cluster analysis as a possible new teratogen signal detection technique
Leads	Eugène van Puijenbroek, Lareb Yrea Weetink, Lareb
Collaborators/tasks	David Lewis, Novartis Michael Stellfeld, Novo Nordisk Other collaborators will be added during the execution of the study
Data sources	PV-report (Lareb), spontaneous and literature reports Argus Safety (Novartis), spontaneous-, literature and solicited (non-interventional studies and patient support programmes) reports

As described in section 7.4. Rationale, we will explore new methods of teratogen signal detection techniques, specifically cluster analysis. The starting point in the detection of safety signals in spontaneous reporting usually is the association between a suspected medicinal product and a single reported term (product-event combination). This also holds for the analysis of ICSRs related to exposure to drugs during pregnancy. In reality however, the majority of reports are coded with multiple ADRs. As an example, over 60% of pregnancy related reports at Lareb have 2 or more reported terms. Often, there will be a reason for co-reporting these terms. For example, these terms express the same underlying phenomenon (e.g. nausea and vomiting, itching and rash, bruising and tendency to bleed), they represent an underlying syndrome (e.g shortness of breath, low blood pressure, urticaria), or have a similar common cause (diabetes and myocardial infarction), etc. Finally, coding one or multiple events in a single patient is impacted by coding standards and coding conventions. (27) although multiple factors may influence coding practices, focussing on drug-single ADR association in the analysis of ICSRs only, will ignore the link between the co-reported terms. This aspect is currently not routinely taken into account in the analysis of ICSRs.

In this sub-study we will examine if the use of cluster analysis will allow for a meaningful classification of reported adverse events in ICSRs related to pregnancy in order to support the signal detection. Cluster analysis seeks to identify natural subgroups in data with closer resemblance between items within a subgroup than between items in different subgroups. It represents a form of unsupervised learning, where the classes and their profiles are derived from data without additional guidance. (28) However, the approach requires dedicated knowledge to interpret the outcomes of the analysis based on for instance the clinical background and pathogenesis to determine if the results are indeed meaningful. For this reason, this study is exploratory in nature. As cluster analysis is an assumption free method no a priori hypothesis can be included.

The WHO collaborating centre UMC has previously used a new approach in cluster analysis, specifically designed for the purpose of signal detection of pharmacovigilance data (28), but the exact statistical approach remains unpublished, but is expected early 2021. In this study we will start exploring the possibilities of regular clustering methods such as Hierarchical Cluster Analysis (HCA) in data sources of Lareb and Novartis. (29) We will explore the possibility of performing cluster analysis in VigiBase (UMC) in this protocol. Additionally, as described in **S1** and **S4**, we will continue studying possibilities for collaboration with the EMA EudraVigilance database as well.

10.5.1. Study design

This study will be a retrospective observational cohort study of pregnancy related cases of products mentioned in PASS information 'active substances' (page 1 of this protocol). Cases will be analysed at the MedDRA Preferred Term (PT)-level and ATC code of the suspected drugs. Drugs will be taken into account individually and as classes, in order to find the most optimal method in this exploratory analysis.

Several approaches for cluster analysis are in use. In this study we will explore the possibilities of HCA as a teratogen signal detection technique. In the case that a collaboration with the UMC/WHO can be formalised, the possibility of consensus clustering [to be published] will be assessed as well.

Obtained clusters will be reviewed by clinicians to determine the backgrounds of the cluster (e.g. pharmacological profile, clinical scenario, etc.). For all clusters it will be assessed whether the association was known by means of comparison to the SmPC and consultation of teratology experts.

10.5.2. Setting

This study will be performed on databases of The Netherlands PV Centre Lareb (PV-report) and Novartis (Argus Safety). Both databases contain both spontaneous reports and literature reports; the Novartis system also includes solicited reports (from non-interventional studies and patient support programmes). Data will be analysed from both data sources separately and an informal comparison of the results will be performed.

All reports collected in the timeframe from inception to the 1st of January 2022 will be eligible for inclusion. From start to finish, the study is expected to take 6 months, consisting of 2 months data collection, 2 months data analysis, and 2 months of writing of the report.

In- and exclusion criteria

The same in- and exclusion criteria as described in **S1** (section 9.1.2. Setting) will be taken into account.

10.5.3. Variables

Spontaneous reports code pregnancy complications (ICH-E2B(R3) field E.i.2.1b) using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.1 (November 2020). (19) The reported MedDRA PT fields will be taken into account as well as the ATC codes for the suspected drugs as mentioned in PASS information 'active substances' (page 1 of this protocol). Other variables that will be taken into account are: maternal age, duration of pregnancy, source of the reports, and trimester of exposure.

10.5.4. Data sources

Both spontaneous reports and literature reports will be selected from the included data sources: PV-report from the Netherlands PV centre Lareb, and Argus Safety from Novartis.

Pv-report (Lareb)

The Netherlands PV centre Lareb collects and analyses spontaneous reports of suspected adverse drug reactions reported by health care professionals or patients. These reports are recorded in PV-report, the database of Lareb. Spontaneous reports as reported to Lareb (ICSRs) are forwarded to the European EudraVigilance database and the UMC. Cases that have been published in literature are recorded by MAHs and EMA and are also made available in the form of ICSRs in PV-report. (9) PV-report was set up in 1991 (22), and contains over 296000 reports in total as of November 2020. However, this number also included reports not related to pregnancy. Reports regarding drug exposure during pregnancy can be selected using the pregnancy flag developed by the EMA. [to be published over the next months by EMA]

Argus Safety (Novartis)

Novartis Pharma AG is the marketing authorisation holder for a large portfolio of medicinal products, including pharmaceuticals, biological, advanced medicinal therapeutic products and gene therapies. In accordance with European legislation, Novartis is obliged to collect, manage, and submit individual reports of suspected adverse reactions to its authorized medicinal products. Individual case safety reports are recorded in Argus Safety (version 8.1.2.3). Argus Safety comprises Argus Core, Argus Affiliate and Argus J. It is a commercially available application used for data collection of adverse events and reporting of adverse event information. It is a web-based, validated database used for the tracking and processing of all safety cases received by Novartis. Argus Safety was first deployed in 2003, and contains over 8 million case reports as of November 2020. Reports of exposure to medicines during pregnancy can be selected using validated SQL scripts based on a pregnancy flag

and additional ICH-E2B(R3) fields. Pregnancy information from Argus Safety has been published (e.g. Geissbühler et al., 2018). (23)

10.5.5. Study size

All cases fulfilling the inclusion and exclusion criteria will be considered for this sub-study. Based on the pilot study previously performed in PV-report (described in **S1**), it is estimated that in total (not limited to the demonstration drugs) approximately 4000 spontaneous reports about drug use during pregnancy are available. Upon limitation to the demonstration drugs specified in PASS information 'active substances' (page 1 of this protocol), this number will decrease considerably. If the numbers are too low to perform this study, multiple data sources could be combined.

For Argus Safety, 4,859 case reports for fingolimod, 331 for beta interferon, 679 for mycophenolate, and 757 for azathioprine are available. Overall, an estimated >30,000 reports of exposure to medicinal products during pregnancy in association with authorised medicinal products are included in the database (predominately exposures to the mother, but some reports of paternal exposure exist).

10.5.6. Data analysis

Data selection and analysis will be performed using Microsoft SQL server 17, R studio version 2016 and/or IBM SPSS statistics 22. All pregnancy reports eligible for inclusion will be selected from the database using the pregnancy flag developed by EMA. [to be published]

A cluster analysis method often used is HCA, which generates a series of models with cluster solutions from 1 (all cases in one cluster) to n (each case is an individual cluster). HCA also allows for working with specific variables as opposed to cases and can handle various types of data (e.g. nominal, ordinal etc.); however it is not recommended to mix different levels of measurement. The use of this method for teratogen signal detection will be explored in this study, by means of using various scenarios in order to find the most optimal one. (29) For example, stratification will be performed for known teratogens/medicinal products with a minimum of 10 years information available vs. more recent products.

All obtained clusters will be independently reviewed by two different clinicians to determine whether clusters are probably based on the pharmacological profile, clinical scenario or for other reasons. Differences will be discussed and in case of disagreement a third reviewer will make the final decision (Delphi method). (30) The outcomes will be informally compared with the available knowledge in the Summary of Product Characteristics (SmPC) of the various products under study, and will be discussed with teratology experts as it is expected that the SmPCs do not contain a lot of information on safety of drug use during pregnancy.

10.5.7. Limitations of the research methods

This study is exploratory in nature, which means that no definite conclusions can be made. Apart from that, the outcomes of cluster analysis will strongly depend on the interpretation of the experts involved. As this is part of the method, it will not limit this study, but it might limit the possibilities of cluster analysis as a teratogen signal detection technique.

10.5.8. Other aspects

Not applicable

11. Quality control and data management

11.1. Data recording and source data

Quality monitoring for completeness and correctness of data retrieved from the included data sources will be performed.

11.2. Confidentiality and coding

Project data are handled with uttermost discretion and are only accessible to authorized personnel who require the data to fulfil their duties within the scope of the research project. Data will not include any patient identifier. For **S2** and **S3** clinical information will be needed for the comparison of the quality. Though data will be provided in an anonymised approach, confidentiality agreements will be signed by researchers involved with the data.

11.3. Data collection and transmission

Data collection of sub-studies can be combined in the studies where the same inclusion criteria apply. For **S2** and **S3** STIS and UKTIS will be combined as ENTIS data in RedCap. Data collection from the pREGnant registry is labour intensive as the analysis database is not yet ready. In the case that the research or the timelines are restricted because of this reason, pREGnant data may not be taken into account.

Data transmission is only applicable for **S2** and **S3**, where data of different sources will be combined. This will be done via a standard template or table that has to be filled for all data sources. Data will be anonymized and confidentiality agreements will be signed by researchers involved with the data.

11.4. Retention and destruction of study data

Data stay under the responsibility of each data access provider. Hence, study data will be handled according to local requirements.

Lareb

Processed data will be stored for preservation and further sharing, because it is the actual data that produces the research results. Data will be preserved for a period of 10 years after the project ends. Data will be archived in the Lareb archive.

Novartis

Processed data will be archived for reference and further sharing, because the research results are applicable to a specific dataset. Data will be managed and maintained for a period of 10 years after the completion of the IMI ConcePTION project. Data will be archived in a secure and private records management system which complies with GxP standards.

STIS

Data will be preserved for a period of 10 years after the project ends. Data will be archived in the STIS archive.

UKTIS

Data will be preserved for a minimum period of 10 years after the project ends, and will be securely stored in accordance with the local data protection agreements of the Newcastle upon Tyne Hospitals NHS Foundation Trust IT network.

11.5. Local quality control measures

PV-report (Lareb)

Data are collected and filed according to European Regulations following the Good Pharmacovigilance Guidelines (GVP) as published on the EMA website. (10) Data are stored in ICH-E2B(R3) format following data validation rules as specified in ICH. (11) Data are exchanged with the EudraVigilance database following the specifications as mentioned in the GVP guidelines. (10) Data cleaning and coding of the information is carried out by trained assessors of pharmacovigilance centre Lareb.

UKTIS

All clinical data, including environmental exposures (medicines, occupational chemicals, radiological, biological and social/recreational substances), obstetric history, relevant medical history, and pregnancy outcome information, are reported to UKTIS by healthcare professionals through a standardized data collection procedure. Information reported to UKTIS at first contact can be corrected by the enquirer when providing pregnancy outcome information. In house data consistency calculations are applied for the maternal age at initial reporting to UKTIS, and the expected date of delivery. Clinical advice provided to healthcare professionals is checked by a senior member of the service. A random sample of clinical enquiries are periodically selected and audited for accuracy and completeness of data recording by the Assistant Head of UKTIS.

STIS

The STIS database has been registered with the « Préposée à la protection des données et à l'information du canton de Vaud » (local authority for data protection) and the Swiss Health Observatory. Access to the database is limited to professionals in charge of the STIS activity (secure and limited computer access). The professionals involved in STIS activities are subject to the respect of legal confidentiality. Quality monitoring of data entry into the database is regularly performed.

pREGnant (Lareb)

All data is collected through web-based questionnaires, which allow for technical checks, validations and internal controls to minimize false entries of data. Completed questionnaires are checked by assessors in order to make sure all data is documented in the correct fields. Medication use is ATC coded, while maternal morbidities, pregnancy and partus related data and data on congenital anomalies is coded in ICD-10. Coding of anomalies is based on EUROCAT standards and a team of clinicians is available for consultation if necessary. In case of congenital anomalies additional clinical information, with permission of a participant, is requested from the paediatrician or family physician. If additional information is needed from other data sources (e.g. pharmacy, clinical records), participants can be contacted to obtain permission for these additional data requests.

Argus Safety (Novartis)

The Quality System for the Novartis Pharmacovigilance system operates in accordance with the guidance set-down in GVP Module 1 'Pharmacovigilance systems and their quality systems'. (31) This quality system is described in detail within the Novartis Pharmacovigilance System Master File (v27.0, 15 Oct 2020). Data are collected and filed according to European Regulations following the Good Pharmacovigilance Practice Modules (GVPs) as published on the EMA website. (10) Data are stored in an extended ICH-E2B(R3) format following data validation rules as specified in ICH (11) and in accordance with 21 CFR Part 11. (32) Data are submitted to the EudraVigilance database following the specifications within the GVP guidelines. (10) Data cleaning and coding of the information is carried out by trained pharmacovigilance specialists in accordance with GPVP and GCP guidelines. AUDITS INSPECTIONS

12. Protection of human subjects

The protection of human subjects is dependent on national regulations. As this project is performed in various settings, the applicable laws per location will be discussed, including data privacy.

The Netherlands

This study does not need approval by a medical ethical committee according to the Dutch regulations ('niet-WMO plichtig' Wet Medisch-wetenschappelijk Onderzoek met mensen).

United Kingdom

This study does not require review by Research Ethics Committees for sites in England, Scotland, Wales or Northern Ireland. This was verified by use of the tool @ http://www.hra-decisiontools.org.uk/ethics/.

Switzerland

In Switzerland this study using anonymized data does not require ethics committee approval.

13. Management and reporting of adverse events/adverse reactions

In case signals will be detected during the course of these demonstration studies signals will be handled according to the current and local rules and regulations. (33)

14. Plans for disseminating and communicating study results

Results of this study will be used in the ConcePTION project and published in a peer reviewed scientific journal if possible. Authorship in scientific publication will have to satisfy the conditions of the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (www.icmje.org/icmje-recommendations.pdf.)

15. References

1. Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. American Journal of Medical Genetics Part C: Seminars in Medical Genetics. 2011;157(3):175-82.
2. Lupattelli A, Spigset O, Twigg MJ, Zagorodnikova K, Mardby AC, Moretti ME, et al. Medication use in pregnancy: a cross-sectional, multinational web-based study. BMJ Open. 2014;4(2):e004365.
3. Mitchell AA, Gilboa SM, Werler MM, Kelley KE, Louik C, Hernandez-Diaz S, et al. Medication use during pregnancy, with particular focus on prescription drugs: 1976-2008. Am J Obstet Gynecol. 2011;205(1):51 e1-8.
4. ConcePTION full proposal. 2018.
5. Benevent J, Montastruc F, Damase-Michel C. The importance of pharmacoepidemiology in pregnancy-implications for safety. Expert Opin Drug Saf. 2017;16(10):1181-90.
6. Vorstenbosch S, te Winkel B, van Gelder MMHJ, Kant A, Roeleveld N, van Puijenbroek E. Aim and Design of pREGnant, the Dutch Pregnancy Drug Register. Drug Safety. 2018;42(1):1-12.
7. EMA. EudraVigilance [Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/pharmacovigilance/eudravigilance.
8. Lindquist M. Vigibase, the WHO Global ICSR Database System: Basic Facts. Drug Information Journal. 2008;42:409-19.
9. EMA. Medical literature monitoring [Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/medical-literature-monitoring.
10. EMA. Guideline on good pharmacovigilance practices (GVP) Module VI - Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 2017.
11. EMA. ICH guideline E2B (R3) on electronic transmission of individual case safety reports (ICSRs) - data elements and message specification - implementation guide. 2013.
12. Bergvall T, Noren GN, Lindquist M. vigiGrade: a tool to identify well-documented individual case reports and highlight systematic data quality issues. Drug Saf. 2014;37(1):65-77.
13. Oosterhuis I, Rolfes L, Ekhart C, Muller-Hansma A, Härmark L. First experiences with a tool to measure the level of clinical information present in adverse drug reaction reports. Expert Opinion on Drug Safety. 2017;17(2):111-5.
14. Lester J, Neyarapally GA, Lipowski E, Graham CF, Hall M, Dal Pan G. Evaluation of FDA safety-related drug label changes in 2010. Pharmacoepidemiol Drug Saf. 2013;22(3):302-5.
15. Raine JM. Risk management - a European regulatory view. New York: Wiley; 2007.
16. Scholl JHG, van Hunsel F, Hak E, van Puijenbroek EP. A prediction model-based algorithm for computer-assisted database screening of adverse drug reactions in the Netherlands. Pharmacoepidemiol Drug Saf. 2018;27(2):199-205.
17. Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. Drug Saf. 2015;38(6):577-87.
18. van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf. 2002;11(1):3-10.
19. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Safety. 1999;20(2):109-17.
20. EUROCAT. EUROCAT Guide 1.4: Instruction for the registration of congenital anomalies. EUROCAT Central Registry, University of Ulster; 2013.
21. Lechat MF, Dolk H. Registries of Congenital Anomalies: EUROCAT. Environmental Health Perspectives Supplements. 1993;101:153-7.
22. van Grootheest AC, van Puijenbroek EP. Geneesmiddelenbewaking in Nederland. 2008.

23. Geissbuhler Y, Vile J, Koren G, Guennec M, Butzkueven H, Tilson H, et al. Evaluation of pregnancy outcomes in patients with multiple sclerosis after fingolimod exposure. Ther Adv Neurol Disord. 2018;11:1756286418804760.
24. Safari S, Baratloo A, Elfil M, Negida A. Evidence Based Emergency Medicine; Part 5 Receiver Operating Curve and Area under the Curve. Evidence Base Medicine. 2016;4(2):111-3.
25. Geissbuhler Y, Rezaallah B, Moore A. An alternative to product-specific pregnancy registries? PRIM; PRegnancy outcomes Intensive Monitoring. Reprod Toxicol. 2020;94:13-21.
26. Hill AB. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine. 1965;58(5):295-300.
27. Berlin C, Blanch C, Lewis DJ, Maladorno DD, Michel C, Petrin M, et al. Are all quantitative postmarketing signal detection methods equal? Performance characteristics of logistic regression and Multi-item Gamma Poisson Shrinker. Pharmacoepidemiology and Drug Safety. 2012;21(6):622-30.
28. Chandler RE, Juhlin K, Fransson J, Caster O, Edwards IR, Noren GN. Current Safety Concerns with Human Papillomavirus Vaccine: A Cluster Analysis of Reports in VigiBase((R)). Drug Saf. 2017;40(1):81-90.
29. Zhang Z, Murtagh F, Van Poucke S, Lin S, Lan P. Hierarchical cluster analysis in clinical research with heterogeneous study population: highlighting its visualization with R. Ann Transl Med. 2017;5(4):75.
30. Rowe G, Wright G. The Delphi technique as a forecasting tool: issues and analysis. International Journal of Forecasting. 1999;15:353-75.
31. EMA. Guideline on good pharmacovigilance practices (GVP) Module I - Pharmacovigilance systems and their quality systems. 2012.
32. FDA. 21 CFR Part 11 Electronic records; Electronic signatures - Scope and Application. Guidance for industry. 2003.
33. EMA. Guideline on good pharmacovigilance practices (GVP) Module IX – Signal management (Rev 1). 2017.

Annex 1. List of stand-alone documents

Number	Document reference number	Date	Title
1			
2			

Annex 2. ENCePP checklist for study protocols

Doc.Ref. EMA/540136/2009

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Analysis of pregnancy pharmacovigilance data in spontaneous reports, and literature, (Individual Case Safety Reports originating from published case series, non-interventional studies and patient support programmes); demonstration study 2.5.1 of the ConcePTION project

EU PAS Register® number: **Study reference number (if applicable):**

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	10.x.2
1.1.2 End of data collection[2]	☒	☐	☐	10.x.2
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☐	☐	☒	
1.1.6 Final report of study results.	☒	☐	☐	10.x.2

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Comments:

Section 2: Research question		**Yes**	**No**	**N/A**	**Section Number**
2.1	Does the formulation of the research question and objectives clearly explain:	☐	☐	☐	
	2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	8.4 and 10.x
	2.1.2 The objective(s) of the study?	☒	☐	☐	9
	2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	10.x.2
	2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
	2.1.5 If applicable, that there is no *a priori* hypothesis?	☒	☐	☐	10.x

Comments:

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	10.x.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	8
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☐	☐	☒	
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☐	☒	
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	10.x.4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	10.x.2

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.2.2 Age and sex	☒	☐	☐	10.x.2
4.2.3 Country of origin	☒	☐	☐	10.x.4
4.2.4 Disease/indication	☒	☐	☐	10.x.2
4.2.5 Duration of follow-up	☐	☐	☒	
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	10.x.2

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☐	☐	☒	
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	
5.3 Is exposure categorised according to time windows?	☐	☐	☒	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☐	☒	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	10.x.1 and 10.x.3
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	10.x.3
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☐	☒	

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☐	☐	☒	
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☐	☒	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☐	☐	☒	

Comments:

Wherever confounding of bias is a possibility it is considered and described, however, bias in the conventional sense is not included in this study as it is mainly descriptive and exploratory.

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☐	☒	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	10.x.4
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	10.x.4
9.1.3 Covariates and other characteristics?	☒	☐	☐	10.x.4
9.2 Does the protocol describe the information available from the data source(s) on:				

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☐	☐	☒	
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☐	☐	☒	
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☐	☐	☒	
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	10.x.3 and 10.x.4
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	10.x.3 and 10.x.4
9.3.3 Covariates and other characteristics?	☒	☐	☐	10.x.3 and 10.x.4
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	10.x.1 and 10.x.6

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	10.x.6
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	10.x.5
10.3 Are descriptive analyses included?	☒	☐	☐	10.x.6
10.4 Are stratified analyses included?	☒	☐	☐	10.x.6
10.5 Does the plan describe methods for analytic control of confounding?	☐	☐	☒	
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☐	☒	
10.7 Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8 Are relevant sensitivity analyses described?	☐	☐	☒	

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	11
11.2 Are methods of quality assurance described?	☒	☐	☐	11.5
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☒	
12.1.2 Information bias?	☐	☐	☒	
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	11

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	12
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	11 and 12

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	6

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	14
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	14

Comments:

Name of the main author of the protocol: Yrea Weetink, MSc (Lareb)

Date: 30/November/2020

Signature: _____