Impact of data source diversity on the distribution of key variables in pregnancy cohorts based on the ConcePTION pregnancy algorithm leveraging a random forest imputation model

Anna Girardi, ARS Toscana (Italy)
on behalf of the Pregnancy Algorithm team

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

Pregnancy Algorithm team

Anna Girardi (1), Giorgio Limoncella (1), Giulia Hyeraci (1), Giuseppe Roberto (1), Claudia Bartolini (1), Olga Paoletti (1), Davide Messina (1), Felipe Villalobos (2), Carlo Alberto Bissacco (2), Jesse van den Berg (3), Eline Houben (3), Katia Santacà (4), Valentina Ientile (4), Ylenia Ingrasciotta (4), Gianluca Trifirò (4), Vjola Hoxhaj (5), Carlos E. Duran (5), Judit Riera Arnau (5), Patricia Garcia (6), Mar Martín (6), Consuelo Huerta (6), Ana Llorente Garcia (6), Francisco Sánchez-Sáez (7), Gabriel Sanfélix-Gimeno (7), Clara Rodríguez (7), Régis Lassalle (8), Jérémy Jové (8), Marie-Agnès Bernard (8), Nicolas Thurin (8), Sue Jordan (9), Dan Thayer (9), Hywel Turner Evans (9), Alex Coldea (9), Marco Manfrini (10), Marleen van Gelder (11), Saeed Hayati (11), Tania Schink (12), Michele Tari (13), Romin Pajouheshnia (14), Ana Sofia Afonso (15), Maryline Le Noan-Laine (16), Ditte Mølgaard-Nielsen (17), Marianne Cunnington (18), Caitlin Dodd (19), Miriam Sturkenboom (5), Hedvig Nordeng (11), Rosa Gini (1)

(1) ARS Toscana, Florence, Italy (2) IDIAP Jordi Gol, Barcelona, Spain (3) PHARMO Institute (4) University of Verona, Italy (5) University Medical Center Utrecht - UMCU, Netherlands (6) Agencia Española de Medicamentos y Productos Sanitarios - AEMPS, Spain (7) Health Services Research Unit, FISABIO, Valencia, Spain (8) Bordeaux PharmacoEpi, INSERM CIC-P 1401, Université de Bordeaux, France (9) Swansea University, Wales, UK (10) University of Ferrara, Italy (11) Pharmacoepidemiology and Drug Safety research group, Department of Pharmacy, University of Oslo, Norway, (12) Leibniz Institute for Prevention Research and Epidemiology – BIPS, Germany (13) Caserta Local Health Unit, Caserta (14) RTI Health Solutions, Spain (15) Eli Lilly, Global Patient Safety, the Netherlands (16) Eli Lilly, Global Patient Safety Ireland (17) Global Safety, Novo Nordisk, Søborg, Denmark (18) Analysis Group (19) Panalgo

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520
concePTION

Disclosure

ConcePTION project has received support from the EU/EFPIA Innovative Medicines Initiative 2 Joint Undertaking ConcePTION grant no. 821520.
https://www.imi-conception.eu/

AG, GL, GH, GR, OP, DM, RG disclose that they are employed by/consultants of a public research institution that conducts or participates in pharmacoepidemiology studies compliant with the ENCePP Code of Conduct and whose budget is partially sustained by such studies.
MS is coordinating a department that conducts studies for EMA and pharmaceutical companies (Pfizer, AstraZeneca, Janssen) using the ENCePP code of conduct.
MC works for Analysis Group that receives funds from Pharmaceutical Companies to conduct real world studies

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

Background

ConcePTION Project

Aims to establish a trusted consortium that can efficiently, systematically, and in an ethically responsible manner, **generate and disseminate reliable evidence-based information regarding effects of medications used during pregnancy and breastfeeding** to women and their healthcare providers

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

conceptION

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

Background

ConcePTION Pregnancy Algorithm

Aims to identify the **most comprehensive list of pregnancies** occurring in the population of the participating data sources, incorporating all possible scenarios in a **transparent and common framework**

The algorithm is stored in a publicly accessible GitHub repository
Diagnosis code list used in the algorithm publicly available on Zenodo

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

conceptION

Aim of the study

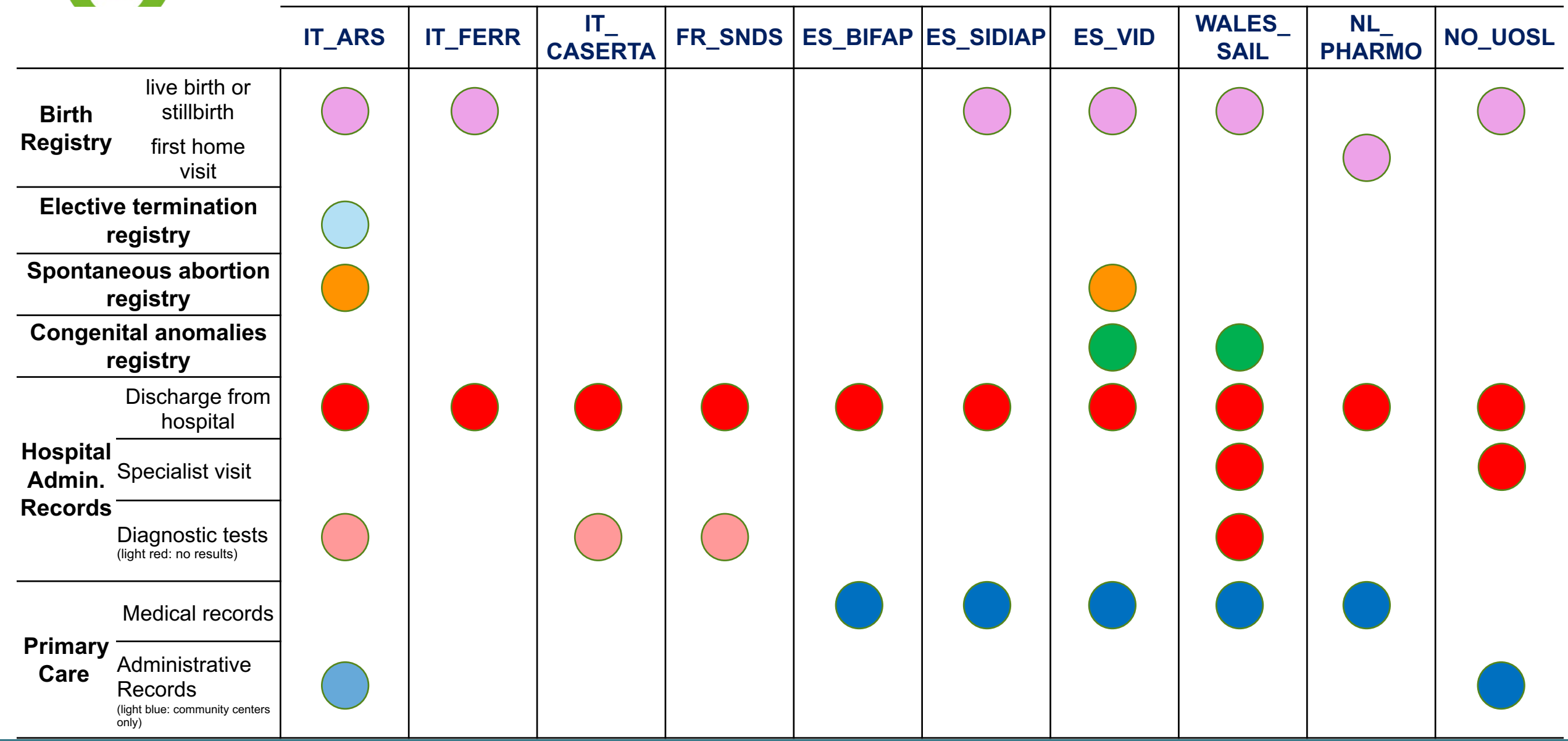
To describe the impact of data diversity of 10 European data sources on the results from the ConcePTION Pregnancy Algorithm

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

METHODS: Geographical distribution of data sources

METHODS: Diversity of data sources

analysis according to the DIVERSE framework (Gini R et al. 2024)



ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

METHODS: overall design of the meta-algorithm

1 Retrieval and labelling of records

(study period = 01/2015-12/2019)

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

A

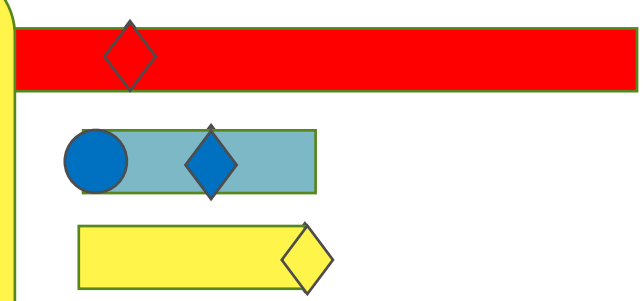
Some records imply that the pregnancy has ended on record date, and also carry information on gestational age at that date

Some records imply that the pregnancy is ongoing on record date, but do carry information on gestational age at that date

B

Some records do not carry any specific information neither on start nor on end date of pregnancy

Some records imply that the pregnancy has ended on record date, but don't carry explicit information on gestational age at that date



Time

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

ConcePTION

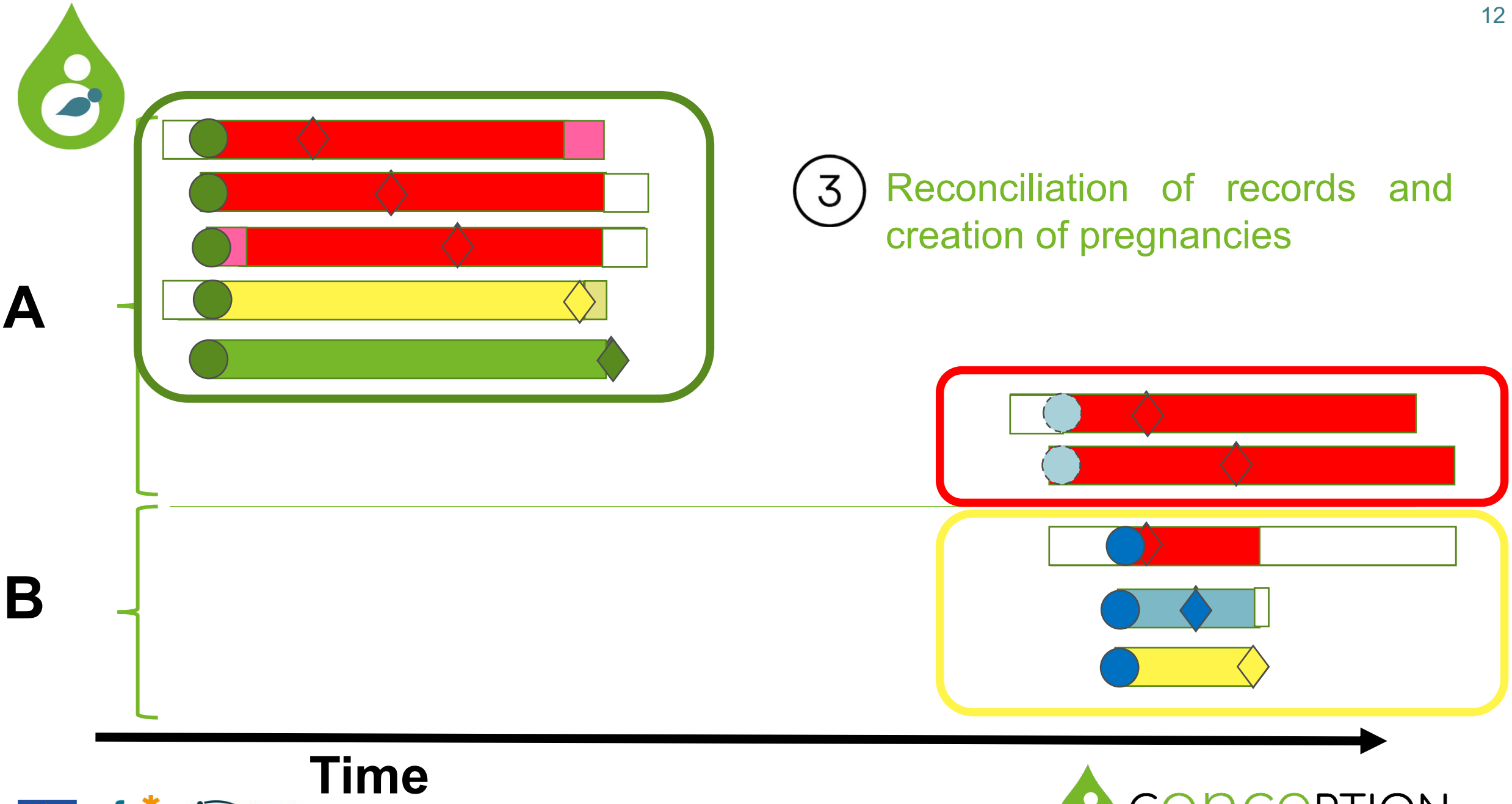
ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

METHODS: overall design of the meta-algorithm

2 Exclusion criteria

1. The ‘person identifier’ was not recorded in the list of persons included in the data source;
2. The person was not in fertile age (12-55y) at the start of pregnancy;
3. Records whose record date was not included in the periods when the person was observed in the data source

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520



3 Reconciliation of records and creation of pregnancies

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking
ConcePTION grant nº 821520

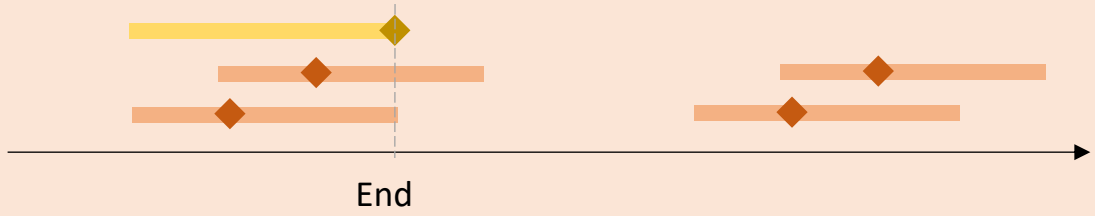
ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

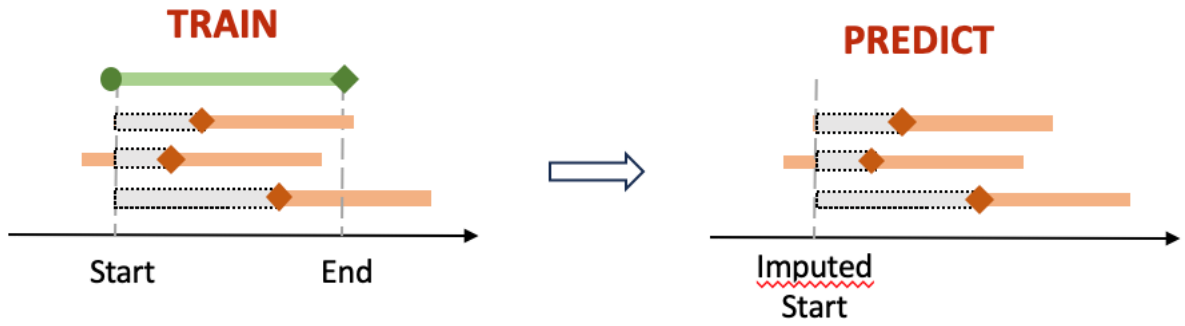
METHODS: overall design of the meta-algorithm

4 Predictive model for pregnancy start date

- In data sources with green records (such as the birth registries records)
- Based on a Random Forest model

We predict start of pregnancy for yellow and red records using green and blue records





The researc
ConcePTION grant nº 821520

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

METHODS: overall design of the meta-algorithm

5 Output

- ✓ Final dataset of pregnancies
  - Main variables: start date, end date, type of pregnancy end
  - Other variables to track the choices made by the algorithm
- ✓ Descriptive markdown report
- ✓ Sample of pregnancies for manual verification of choices

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520
conceptION

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

Types of pregnancy end in the ConcePTION Pregnancy Algorithm

LB	Live birth	the pregnancy ended in a live birth
SB	Stillbirth	fetal death before or during the delivery, after gestational week 22 or week 24 in UK
SA	Spontaneous abortion	pregnancy loss before 22 weeks' gestation (24 weeks in UK)
T	Elective termination	legal termination of pregnancy /medical abortion
ECT-MOL	Ectopic or molar pregnancy	the fertilized egg implants outside the uterus or there is evidence of abnormal product of conception
ONGOING	Pregnancy ongoing	the estimated date of end of pregnancy is after the date on which data are extracted
UNK	Unknown	the imputed or observed date of end of pregnancy is before the cutoff date of the data, therefore the pregnancy has surely ended, but the type of end could not be established
UNF	Unfavorable Unspecified	non-live birth pregnancy with observed end date, but outcome unspecified
LOSTFU	Lost to follow-up	the estimated date of end of pregnancy is after the end of the observation period (i.e. a continuous period of inclusion in the underlying population of the data source) of the pregnant person

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520

ConcePTION

ispe International Society for Pharmacoepidemiology 2024 ISPE ANNUAL MEETING August 24-28 BERLIN, GERMANY

CONCLUSIONS

- ✓ **Creation of a meta-algorithm** that handle in a transparent manner data source diversity
- ✓ Tested on **10** diverse European data sources
- ✓ **Distribution of quality and types of pregnancy end** influenced by data diversity
- ✓ **Decision on which pregnancies are included in the study** should be based on the data source and on the research question, to minimize bias

efpia imi innovative medicines initiative
The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520
concePTION

Thanks to all the research partners

The research leading to these results has received support from the EU/EFPIA Innovative Medicines Initiative [2] Joint Undertaking ConcePTION grant nº 821520