Study Protocol

26/04/2024

Version 3.6

TABLE OF CONTENTS

1 Contents

Table of contents **2**

Document History **3**

LIST OF ABBREVIATIONS **6**

1. Title **7**

2. Responsible parties – study team **7**

3. ABSTRACT (Stand alone summary of the study protocol) **8**

4. AMENDMENTS AND UPDATES **11**

5. MILESTONES **12**

6. RATIONALE AND BACKGROUND **12**

7. RESEARCH QUESTION AND OBJECTIVES **15**

8. RESEARCH METHODS **19**

8.1 Study type and Study Design 19

8.2 Study Setting and Data Sources 19

8.3 Study Period 23

8.4 Follow-up 23

8.5 Study Population with inclusion and exclusion criteria 24

8.6 Variables 25

8.6.1. Exposure/s (where relevant) 25

8.6.2. Outcome/s (where relevant) 25

8.6.3. Other covariates, including confounders, effect modifiers and other variables (where relevant) 27

8.7 Study size 30

8.8 Analysis 30

8.9 Evidence synthesis 35

9. DATA MANAGEMENT **35**

10. QUALITY CONTROL **36**

11. LIMITATIONS OF THE RESEARCH METHODS **37**

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS **37**

13. GOVERNANCE BOARD ASPECTS **38**

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS **38**

15. OTHER ASPECTS **38**

16. REFERENCES **38**

17. ANNEXES **41**

DOCUMENT HISTORY

Version	Date	Description
V1.0	20.10.2023	**First Draft**
V2.0	10.11.2023	**Second Draft**
V2.1	21.11.2023	**Second Draft with format changes**
V3.0	09.01.2024	**Final archiving version**
V3.1	26.04.2024	**Amendment including change IQVIA DA Analyzer Germany per Norway**

Study Title	DARWIN EU - Effectiveness of Human Papillomavirus Vaccines (HPV) to prevent cervical cancer
Protocol version identifier	V3.6
Date of last version of protocol	26/04/2024
EU PAS register number	EUPAS1000000080
Active substance	• Bivalent HPV vaccine (types 16, 18) • Quadrivalent HPV vaccine (types 6, 11, 16, 18) • 9-valent HPV vaccine (types 6, 11, 16, 18, 31, 33, 45, 52, 58)
Medicinal product	• Cervarix • Gardasil/Silgard • Gardasil-9
Research question and objectives	Research question: What is the effectiveness of HPV vaccination in prevention of severe disease outcomes in women, including invasive cervical cancer and CIN2+ for the different licensed HPV vaccines in Europe. More specifically the study objectives are: Main objectives: 1. To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer stratified by licenced vaccine brand. 2. To assess the effectiveness of HPV vaccination in prevention of CIN2+, stratified by licenced vaccine brand. 3. To assess the effectiveness of HPV vaccination in prevention of, conization, stratified by licenced vaccine brand. Secondary objectives: • To assess the effectiveness of HPV vaccination overall for the three outcomes (i.e. invasive cervical cancer, CIN2+ and conization) • To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer, CIN2+ and conization separately in subgroups defined by number of doses, within each brand.

Country(-ies) of study	UK; Spain; Norway
Author	Daniel Prieto Alhambra, Albert Prats Uribe

LIST OF ABBREVIATIONS

Abbreviation	Name
CDM	Common Data Model
CIN	Cervical Intraepithelial Neoplasia
CPRD	Clinical Practice Research Datalink
EHR	Electronic Health Record
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
IQVIA DA	IQVIA Disease Analyzer
IRR	Incidence rate ratio
OMOP	Observational Medical Outcomes Partnership
SIDIAP	The Information System for Research in Primary Care
SNOMED	Systematized Nomenclature of Medicine

1. TITLE

Effectiveness of Human Papillomavirus Vaccines (HPV) to prevent cervical cancer

2. RESPONSIBLE PARTIES – STUDY TEAM

Table 2.1 Study Team

Study team Role	Names	Organisation
Principal Investigators	Daniel Prieto Alhambra Albert Prats Uribe	University of Oxford University of Oxford
Data Scientist / Statistician	Mike Du Edward Burn Marti Catala Sabate	University of Oxford University of Oxford University of Oxford
Epidemiologist	Daniel Prieto Alhambra	University of Oxford
Clinical Domain Expert	Albert Prats Uribe	University of Oxford

Table 2.2 Databases and data partners

Data partner	Local Study Coordinator/Data Analyst	Organisation
SIDIAP	Talita Duarte Salles	*IDIAP JGOL*
CPRD GOLD	Antonella Delmestri	*University of Oxford*
NLHR	Hedvig Marie Egeland Nordeng	*University of Oslo*
NLHR	Nhung Trinh	*University of Oslo*

3. ABSTRACT (STAND ALONE SUMMARY OF THE STUDY PROTOCOL)

Title:
DARWIN EU - Effectiveness of Human Papillomavirus Vaccines (HPV) to prevent cervical cancer.

Rationale and Background:
HPV vaccination programmes have been shown to reduce not only HPV infection but also the incidence of cervical cancer. However, uncertainty remains on the real-world effectiveness of different brands, and dose schedules.

Research Questions and Objectives:
To generate evidence from real-world data on the effectiveness of HPV vaccination in preventing severe disease outcomes in women, including invasive cervical cancer and CIN2+, for the different licensed HPV vaccines in Europe. This study will include data sources from UK, Spain and Norway.

More specifically the study objectives are:

Main objectives:

1. To assess the effectiveness of HPV vaccination in prevention of **invasive cervical cancer** stratified by licenced vaccine brand.
2. To assess the effectiveness of HPV vaccination in prevention of **CIN2+**, stratified by licenced vaccine brand.
3. To assess the effectiveness of HPV vaccination in prevention of **conization**, stratified by licenced vaccine brand.

Secondary objectives:

- To assess the effectiveness of HPV vaccination overall for the three outcomes (i.e. invasive cervical cancer, CIN2+ and conization)
- To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer, CIN2+ and conization in subgroups defined by number of doses, within each brand.

Results in both main and explanatory analyses will be further stratified by age group.

Research Methods

The target trial emulation approach will be used for this non-interventional study. The summary of the Target Trial is as follows:

1. Study primary objective: to investigate the effectiveness of HPV vaccines to prevent cervix cancer.
2. Estimand:
    a. Population: Women eligible for vaccination according to eligibility criteria in each country. More generally, females 9 years old or older any date after the launch of the vaccination programme in the corresponding country.
    b. Treatments:
        i. Placebo (unvaccinated).

ii. Vaccinated with Gardasil/Silgard.
iii. Vaccinated with Cervarix
iv. Vaccinated with Gardasil-9.

Patients are randomised to one of the above groups in a 1:1:1:1 ratio.

c. Variable/outcome: incidence of invasive cervical cancer within 5, 10 and 15 years of vaccination.
d. Summary measure: incidence rate. Comparison of interest is incidence rate ratio between every vaccinated group and the unvaccinated group.
e. Intercurrent events:
    i. For the unvaccinated: vaccination, dealt with a hypothetical strategy. To implement this, data from women in the unvaccinated group that at some point get vaccinated will be included in the analysis up to the time of vaccination. (This means in the analysis these women will be censored at the time of vaccination.)
    ii. For the vaccinated, any group: treatment discontinuation (i.e. not receiving all scheduled doses), dealt with a treatment policy strategy. To implement this strategy, all available data from these women will be included in the analysis regardless of treatment discontinuation.
f. Statistical methods: The incidence rate ratio between each vaccinated group and the unvaccinated group will be estimated.

Study Design:
New user matched cohort study.

Population:
All females born on or after 1993 (15 years old or less in 2008 – the earliest launch of the vaccine in all countries). This population will be restricted to those in observation in the database when they turned 15, and in observation in the database when they turned 9.

Further restrictions will be made in a year per year basis for the whole study period, applied on the 1st of January of each year. For each year, participants need to be in observation on the 1st of January of that year, need to have 365d of prior observation available, and need to be aged between 9 and 15 years old.

The analysis will be further restricted to matched cohorts of vaccinated and unvaccinated participants with similar baseline characteristics (see 'Data Analysis').

Data Sources:
- Primary care records from the UK (Clinical Practice Research Datalink (CPRD) GOLD) and primary care records linked to hospital records from Catalonia, Spain (Information System for Research in Primary Care (SIDIAP)); Population-based health registry data from Norway (NLHR) Norwegian Linked Health Registry data.

Study Period:

HPV national vaccination programs have different start dates. For CPRD and SIDIAP, the study period will begin on 1/1/2008 and for Norway on 1/1/2009.

For all databases, the end year of the study period is the most recent data available, sometime in 2023.

Eligibility criteria

Eligibility for vaccination in each country. More generally, females between 9 and 15 years old at any date after the launch of the vaccination programme in the corresponding country.

Exposure

Assignment procedures: Vaccination status (brand and number of doses) is assigned as seen in the data at 15 years old. Unvaccinated will be assigned as not being vaccinated at 15 years old and censored when (and if) they get vaccinated later on.

Brand: For those vaccinated, brand will be primarily assigned as brand of all the doses administered before 15. Women with heterologous brand (not the same brand for each dose) schedules will be excluded. If this information is not available, it will be inferred, when possible, using each country's vaccination schedules.

Schedules: Unvaccinated, vaccinated with 1 dose, vaccinated with 2 doses, and vaccinated with 3 doses.

Outcome

The main outcome of interest is invasive cervical cancer. Two secondary outcomes are also considered: CIN2+ and Conization. These outcomes will be phenotyped and diagnostics will be carried out.

Follow-up

Follow up will start at the moment of the administration of first dose before 15 years old. For unvaccinated, the follow up will start at the same date as their vaccinated matched counterpart. Follow-up will extend until another vaccine dose or outcome event, end of available follow-up, or death of any individual of the matched pair, whichever comes first.

Other variables

Year of birth, calendar year, age at vaccination, cytology results from smear test prior to the first dose of vaccine if available. For LASSO regression, all recorded features recorded in the database, including socio-demographics, geographic location, healthcare resource use (measured as number of visits on the prior year), comorbidity, medicine/s use, previous smear testing, and previous vaccination/s.

Data Analysis

All analyses will be conducted separately for each database, and carried out in a federated manner, with effectiveness estimates meta-analysed and the I^2 heterogeneity coefficient reported.

We will conduct a propensity score (PS) matched cohort design, where target and comparator cohort participants will be matched up to 5:1. Matching will be done based on PS, year of birth, year of first dose (for analyses not involving dose number) and geographic region using nearest neighbour matching, with calliper width 0.2 standard deviations as is standard for propensity score matching. Large-scale PS will be estimated using lasso regression to estimate the probability of being in the target cohorts, potentially including any of the covariates mentioned above.

The following matched cohorts will be compared:

Main comparisons:

Vaccinated vs unvaccinated per brand:

- Vaccinated with Gardasil/Silgard (target) (1 or more dose) vs unvaccinated (comparator)
- Vaccinated with Cervarix (target) (1 or more dose) vs unvaccinated (comparator)
- Vaccinated with Gardasil-9 (target) (1 or more dose) vs unvaccinated (comparator)

Secondary comparisons:

Vaccinated (target) (1 or more dose) (any brand) vs unvaccinated (comparator) overall.

Dose comparisons:

- Vaccinated with 2 or more doses (target) vs 1 dose (comparator) of the same brand.
- Vaccinated with 3 or more doses (target) vs 2 doses (comparator) of the same brand.

Vaccine effectiveness analyses

Incidence rates and incidence rate ratios (IRR) will be calculated for the matched cohorts and outcomes at 5, 10 and 15 years (if enough follow-up is available). Cox proportional hazard models will be used to calculate hazard ratios (HR) for time-to-event analyses.

4. AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
1	26/04/2024	Abstract, Methods, Milestones	Better specification of methods Changed IQVIA Germany DA for NLHR	Added clarifications to specify better the methods. Study in IQVIA Germany DA deemed not feasible after diagnostics.

5. MILESTONES

Additional study specific deliverables (e.g. results of validation of new disease concepts, statistical analysis plan, etc.) might be requested but this will be on an individual basis

STUDY SPECIFIC DELIVERABLE	TIMELINE
Draft Study Protocol	October 2023
Final Study Protocol	January 2024
Creation of Analytical code	January-July 2024
Execution of Analytical Code on the data	February-July 2024
Interim Study Report (if applicable)	NA
Final Study Report	31st July 2024

6. RATIONALE AND BACKGROUND

Cervical cancer ranks as the second most common cancer among women aged 15 to 44 years in the European Union (EU) and England (1, 2). Annually, there are approximately 33,000 patients diagnosed with cervical cancer in the EU, resulting in 15,000 fatalities (2). The primary cause of cervical cancer is persistent infection of the genital tract by specific strains of human papillomavirus (HPV). There are over 100 strains of HPV, 40 of which can infect the genital tract, and at least 14 of which are considered 'high risk' for cervical cancer. Around 70% of cases of cervical cancer are caused by HPV types 16 and 18 – the most common 'high risk' strains (2).

In 2018 the World Health Organisation (WHO) launched the 'Cervical Cancer Elimination Initiative' which has accelerated the implementation of HPV vaccination programmes (3). As a result, HPV vaccines are now licenced in more than 100 countries worldwide. There are currently three highly efficacious prophylactic vaccines that are approved for use in Europe and the UK: a bivalent (Cervarix), a quadrivalent (Gardasil/Silgard), and a 9-valent (Gardasil-9). Clinical trials have demonstrated each of these to provide protection against HPV-associated anogenital disease including genital warts, intraepithelial neoplasia, and cervical cancer (4-6). Each of these protect against the most carcinogenic HPV strains, 16 and 18, and the quadrivalent and 9-valent vaccines provide additional protection against strains 6 and 11 which are typically responsible for non-cancerous genital warts, and the 9-valent against strains 31, 33, 45, 52 and 58 which have been associated with 20% of cervical cancers (7).

HPV vaccines provide greater advantages and enhanced protection when administered to preadolescent individuals. This is because the HPV vaccine is more effective in people who have not previously been exposed to the HPV types included in the vaccine, and research has shown that preadolescents tend to have a more robust immune response to the vaccine compared to adults (8).

Each of these 3 vaccines are approved for use in females from the age of 9 years to protect against precancerous lesions (intraepithelial neoplasia), and cervical cancer (2). Males and females aged 9-13 years (Gardasil) or 9-14 years (Cervarix and Gardasila-9) are typically given two doses; and those aged 14+ (Gardasil) or 15+ years (Cervarix and Gardasila-9) a three-dose schedule (9-11).

Because HPV vaccines only began to be approved for worldwide vaccination programmes from 2006, we can only begin to see the long-term effect of such vaccination programmes on incidence of cancerous lesions now. Furthermore, given the known lag between HPV infection and cervical lesions or cancer, longitudinal studies with long follow-up time are required to examine the impact of HPV vaccination on cancer risk.

Some observational studies have examined the impact of HPV vaccination programmes in Europe (12-15). One study in England examined the impact of the bivalent HPV vaccine in reducing incidence of HPV infection, showing substantial declines in HPV strains 16, 18, and cross-protection of strains 31, 33 and 45, 8 years following the start of the vaccination programme (14). One study in Scotland demonstrated an 89% reduction in prevalence of CIN3 or worse in girls vaccinated with the bivalent vaccine compared to unvaccinated girls, and that most protection was provided when girls were vaccinated at age 12-13 years compared to those aged 17 years (15). A meta-analysis conducted in 2021 including 65 articles from across 14 countries, including both bivalent and quadrivalent vaccines, demonstrated that between 5-8 years after the implementation of vaccination programmes prevalence of HPV strains 16 and 18 were reduced by 83% in girls aged 13-19 years, and by 66% in women aged 20-24 years. Between 5-9 years after vaccination, CIN2 or worse decreased by 51% in those screened at aged 15-19 years and by 31% in women screened at age 20-24 years (16). The first study to investigate the impact of the bivalent vaccine on incidence of cervical cancer and CIN3 used National Cancer Registry data in England, and further investigated the impact of age at vaccination (12). Three cohorts of girls vaccinated with the bivalent vaccine, Cervarix, in different calendar years were compared with unvaccinated cohorts from years prior to the vaccination programme roll-out. Girls vaccinated at age 12-13 years exhibited 87% reduction in cervical cancer rates; those vaccinated at age 14-16 years 62% reduction, and those vaccinated at 16-18 years 34% reduction (12) (note that age was classified by school year, with some overlapping ages). Rate reductions of grade 3 cervical intraepithelial neoplasia (CIN3) were even greater (97%, 75% and 39% for those vaccinated at ages 12-13 years, 14-16 years and 16-18 years, respectively).

To our knowledge only one observational study has investigated the effectiveness of the quadrivalent vaccine on cervical cancer (rather than HPV infection, and CIN2+). In a Swedish cohort of adolescent girls, the incidence rate of cervical cancer in girls receiving at least one dose of quadrivalent vaccine was compared to unvaccinated girls. Vaccination substantially reduced incidence of cervical cancer, particularly after adjusting for confounders including age at follow-up, calendar year, county of residence, and parental education, household income, mother's country of birth, and maternal disease history (13). Similarly, the quadrivalent vaccine has been demonstrated to provide protection against the development of cancers of the anus (17); and a meta-analysis of both the bivalent and 9-valent HPV vaccines showed that vaccinated individuals were 80% less likely to develop HPV-16 which is a particular risk for oropharyngeal cancer (18).

HPV vaccination has been shown to be cost-effective globally (19), though there have been suggestions that one-dose may confer comparable protection to two- and three- dose schedules, which could make vaccination programmes more cost-effective both financially and logistically. Evidence from prospective cohort studies and a few retrospective observational studies pointing to the effectiveness of a single HPV vaccine dose in providing strong protection against persistent HPV infections (20-23). For example,

Sankaranarayanan and colleagues have illustrated that the immediate protection offered by one quadrivalent HPV vaccine dose is comparable to that of two or three doses (22). This level of protection is similar to what is achieved with a full three-dose regimen. Similar findings have been reported for the bivalent vaccine (21). Additionally, some studies have modelled the clinical and economic impact of one-dose vaccine schedules compared to no- or 2-dose schedules in reducing HPV infection and cervical cancer outcomes in numerous countries worldwide (24, 25). Yet only a few observational studies have investigated the real-world impact of a single dose schedule on incidence of high-grade cervical lesions (CIN2, CIN3). A study of cancer registry and screening data in Australia has shown that one dose of the quadrivalent vaccine provides comparable effectiveness versus 2 or 3 doses in preventing CIN2 or CIN3 (26). A study in the US also demonstrated equivalent effectiveness of one, two and three doses of the quadrivalent vaccine in reducing incidence of high-grade cervical lesions (27). However, there is a dearth of research investigating these trends in Europe, and none to our knowledge that examine all vaccines approved in these regions, underscoring the need for further investigation into the dosing schedule. Reducing the dosage can lead to cost savings, streamline vaccine distribution, and enhance vaccine accessibility, all while preserving the vaccine's effectiveness in preventing severe illness (25). Recently, the UK Joint Committee on Vaccination and Immunisation (JCVI) have recommended have recommended the use of one-dose vaccination nationally [link], illustrating the relevance of and the need for research on this topic.

Based on all the above, the aim of the present study is to generate real world evidence on the effectiveness of HPV vaccination to prevent cervical cancer, including the analysis of the different licensed HPV vaccines and observed dosing regimens in Europe.

7. RESEARCH QUESTION AND OBJECTIVES

Research question: To generate evidence from real-world data on the effectiveness of HPV vaccination in preventing severe disease outcomes, i.e. invasive cervical cancer and CIN2+ , for the different licensed HPV vaccines in Europe (UK, Spain and Germany).

More specifically the **main study objectives** are:

1. To assess the effectiveness of HPV vaccination in prevention of **invasive cervical cancer,** stratified by licenced vaccine brand
2. To assess the effectiveness of HPV vaccination in prevention of **CIN2+**, stratified by licenced vaccine brand
3. To assess the effectiveness of HPV vaccination in prevention of **conization**, stratified by licenced vaccine brand

Secondary objectives:

- To assess the effectiveness of HPV vaccination overall for the three outcomes (i.e. invasive cervical cancer, CIN2+ and conization)
- To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer, CIN2+ and conization in subgroups defined by number of doses, within each brand.

Results in both main and explanatory analyses will be further stratified by age group.

Table 7.1: Main and secondary research questions and objectives

A. Main Objective 1

Objective:	To assess the effectiveness of HPV vaccination in prevention of **invasive cervical cancer** stratified by licenced vaccine brand.
Hypothesis:	HPV vaccination is effective to prevent invasive cervical cancer
Population ***(mention key inclusion-exclusion criteria):***	All females born on or after 1993 (15 years old or less in 2008 – the earliest launch of the vaccine in all countries). This population will be restricted to those in observation in the database when they turned 15, and in observation in the database when they turned 9. Further restrictions will be made in a year per year basis for the whole study period, applied on the 1st of January of each year. For each year, participants need to be in observation on the 1st of January of that year, need to have 365d of prior observation available, and need to be aged between 9 and 15 years old. The analysis will be further restricted to matched cohorts of vaccinated and unvaccinated participants with similar baseline characteristics (see 'Data Analysis').

Exposure:	Exposure (vaccination, and brand) will be assessed at the time of 15th birthday. Where vaccine brand is not known, HPV composition (i.e. bivalent, quadrivalent, 9-valent) or calendar date will be used as a proxy for brand.
Comparator:	Unvaccinated at the 15th birthday.
Outcome:	Invasive Cervical Cancer
Time *(when follow up begins and ends):*	Follow-up will start from first dose of the vaccine or matched date for the unvaccinated peer. End of follow-up: End of their observation (i.e. date of data extraction, exit form the database, death), next vaccine dose after 15, or outcome occurrence, whichever comes first.
Setting:	Primary care electronic health records from CPRD GOLD [UK], SIDIAP [Spain],-and health registry data from Norway.
Main measure of effect:	- Incidence rate for outcomes per 100,000 person-years and IRR, overall and in 5, 10 and 15 years after vaccination - Vaccine effectiveness against invasive cervical cancer, calculated as 1 - IRR - As a secondary measure of effect time to event and Hazard Ratio (HR) - Kaplan-Meier plots and/or cumulative incidence plots will be used to illustrate time-to-event analyses.

B. Main objective 2

Objective:	To assess the effectiveness of HPV vaccination in prevention of **CIN2+** by licenced vaccine brand.
Hypothesis:	HPV vaccination is effective to prevent CIN2+ or invasive cervical cancer
Population	unchanged from Objective 1
Exposure:	unchanged from Objective 1

Comparator:	unchanged from Objective 1
Outcome:	CIN2+ or invasive cervical cancer
Time *(when follow up begins and ends):*	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1

C. Objective 3

Objective:	To assess the effectiveness of HPV vaccination in prevention of **conization** by licenced vaccine brand.
Hypothesis:	HPV vaccination is effective to prevent conization, CIN2+ or invasive cervical cancer
Population *(mention key inclusion-exclusion criteria):*	unchanged from Objective 1
Exposure:	unchanged from Objective 1
Comparator:	unchanged from Objective 1
Outcome:	conization, CIN2+ or invasive cervical cancer
Time *(when follow up begins and ends):*	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1

C. Secondary Objective 1

Objective:	To assess the effectiveness of HPV vaccination in prevention of **conization, CIN2+ or invasive cancer** (separately) overall.
Hypothesis:	HPV vaccination is effective to prevent conization, CIN2+ or invasive cervical cancer
Population *(mention key inclusion-exclusion criteria):*	unchanged from Objective 1
Exposure:	Exposure (vaccination) will be assessed at the time of 15th birthday or later.

Comparator:	unchanged from Objective 1
Outcome:	conization, CIN2+ or invasive cervical cancer
Time *(when follow up begins and ends):*	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1

C. Secondary Objective 2.1

Objective:	To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer, CIN2+ and conization in subgroups defined by number of doses, within each brand.
Hypothesis:	Incidence rates of the three outcomes are similar between dose schedules
Population *(mention key inclusion-exclusion criteria):*	unchanged from Objective 1
Exposure:	Exposure (number of doses and brand received) will be assessed at the time of 15th birthday. Where vaccine brand is not known, HPV composition (i.e. bivalent, quadrivalent, 9-valent) or calendar date will be used as a proxy for brand. Target: Vaccinated with 2 or more doses.
Comparator:	Vaccinated with 1 dose.
Outcome:	conization, CIN2+ or invasive cervical cancer
Time *(when follow up begins and ends):*	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1

C. Secondary Objective 2.2

Objective:	To assess the effectiveness of HPV vaccination in prevention of invasive cervical cancer, CIN2+ and conization in subgroups defined by number of doses, within each brand.
Hypothesis:	Incidence rates of the three outcomes are similar between dose schedules

Population *(mention key inclusion-exclusion criteria):*	unchanged from Objective 1
Exposure:	Exposure (number of doses and brand received) will be assessed at the time of 15th birthday or later. Where vaccine brand is not known, HPV composition (i.e. bivalent, quadrivalent, 9-valent) or calendar date will be used as a proxy for brand. Target: Vaccinated with 3 or more doses.
Comparator:	Vaccinated with 2 doses.
Outcome:	conization, CIN2+ or invasive cervical cancer
Time *(when follow up begins and ends):*	unchanged from Objective 1
Setting:	unchanged from Objective 1
Main measure of effect:	unchanged from Objective 1

8. RESEARCH METHODS

8.1 Study type and Study Design

Table 8.1. Description of Potential Study Types and Related Study Designs

STUDY TYPE	STUDY DESIGN	STUDY CLASSIFICATION
Vaccine Effectiveness Studies	New User Cohort	Complex

8.2 Study Setting and Data Sources

This study will be conducted using routinely collected data from 3 databases in 3 European countries. All databases were previously mapped to the OMOP CDM.

- Clinical Practice Research Datalink (CPRD GOLD), United Kingdom
- Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP) linked to hospital inpatient records (CMBD-AH for the acronym in Catalan language), Spain
- Norwegian Linked Health Registry data (NLHR), Norway

While CPRD GOLD and SIDIAP are assumed to have complete vaccine records and can contribute to all research questions and comparisons, IQVIA DA Germany will only be used for secondary objectives due to incomplete in vaccination coverage. Detailed information on data sources is described in **Table 8.2**.

Database Selection

The selection of databases for this study was performed based on data reliability and relevance for the proposed research question among those databases onboarded and available within DARWIN EU. The selected databases fulfil the criteria required for the availability of key information on exposures (i.e. complete recording of vaccines including date and brand), outcomes, and covariates, while covering different settings and regions of Europe. Records on outcomes of interest (cervical cancer/CIN/conization) and vaccination will be available in all databases, and counts were obtained during the feasibility stage and are detailed in **Table 8.2**.

Amendment:

After diagnostics of the vaccination and outcome cohorts and measure coverage of vaccination, it was deemed necessary to exclude IQVIA Germany Disease Analyzer (DA) from the study. The main reasons are the low coverage of HPV vaccine seen in the data, less than 10%, and a not expected age distribution for cervix cancer (with >10% of cancers happening in the 15-20 age groups).

We have therefore added Norway (NLHR) Norwegian Linked Health Registry data to the study, it will include all vaccinated women (with a coverage of 90%), and has information on cervix cancer and CIN23.

Table 8.2. Description of the selected Data Sources.

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data (EHR, claims, registries)	Number of active subjects	Feasibility count of exposure (if relevant)	Feasibility count of primary outcome	Data lock for the last update
United Kingdom	CPRD GOLD	Complete records on HPV vaccination, and outcome events of interest	Primary care	EHR	17M	HPV Vaccination Broad definition: ~389,000;	Cervix Cancer: ~18,000;	01/2023
Spain	SIDIAP	Complete records on HPV vaccination, and outcome events of interest	Primary care + linkage to hospital data	EHR	5.8M	HPV Vaccination Broad definition: ~383,000;	Cervix Cancer: ~15,000;	06/2022
Norway	Norwegian Linked Health Registry data (NLHR)	Expected complete records on HPV vaccination, and outcome events of interest	Primary care + linkage to hospital data + vaccination registry	Linked Health Registry	5.7M	HPV Vaccination – not mapped yet	Cervix Cancer: ~53,000;	10/2023

Clinical Practice Research Datalink GOLD, United Kingdom (University of Oxford)

The Clinical Practice Research Datalink (CPRD) is a governmental, not-for-profit research service, jointly funded by the National Institute for Health and Care Research and the Medicines and Healthcare products Regulatory Agency, a part of the Department of Health, United Kingdom (UK) (https://cprd.com). CPRD GOLD (28) comprises computerized records of all clinical and referral events in primary care in addition to comprehensive demographic information and medication prescription data in a sample of UK patients (predominantly from Scotland (52% of practices) and Wales (28% of practices). The prescription records include information on the type of product, date of prescription, strength, dosage, quantity, and route of administration. Data from contributing practices are collected and processed into research databases. Quality checks on patient and practice level are applied during the initial processing. Data are available for 20 million patients, including 3.2 million currently registered patients. Approval for this study was granted via the Research Data Governance Process.

Information System for Research in Primary Care (SIDIAP), Spain (IDIAP Jordi Gol)

SIDIAP is collected from EHR records of patients receiving primary care delivered through Primary Care Teams (PCT), consisting of GPs, nurses and non-clinical staff (29). The Catalan Health Institute manages 286 out of 370 such PCT with a coverage of 5.6M patients, out of 7.8M people in the Catalan population (74%). The database started to collect data in 2006. The mean follow-up is 10 years. The observation period for a patient can be the start of the database (2006), or when a person is assigned to a Catalan Health Institute primary care centre. Date of exit can be when a person is transferred-out to a primary care centre that does not pertain to the Catalan Health Institute, or date of death, or date of end of follow-up in the database. Drug information is available from prescriptions and from dispensing records in pharmacies. Drugs not prescribed in the GP setting might be underreported; and disease diagnoses made at specialist care settings are not included. Approval for this study was granted by both SIDIAP's Scientific and Ethics Committee.

Norwegian Linked Health Registry data (NLHR)

Norway has a universal public health care system consisting of primary and specialist health care services covering a population of approximately 5.4 million inhabitants. Many population-based health registries were established in the 1960s with use of unique personal identifiers facilitating linkage between registries. Data in these health registries are used for health analysis, health statistics, improving the quality of healthcare, research, administration and emergency preparedness. We harmonized data from the following registries: the Medical Birth Registry of Norway (MBRN), the Norwegian Prescription Registry (NorPD), the Norwegian Patient Registry (NPR), Norway Control and Payment of Health Reimbursement (KUHR), the Norwegian Surveillance System for Communicable Diseases (MSIS), the Norwegian Immunisation Registry (SYSVAK), the National Death Registry, and the National Registry (NR). Linkage between the registries was facilitated using project-specific person ID generated from unique personal identification assigned at birth or immigration for all legal residents in Norway. In brief: MBRN stores information about the pregnancy, the mother, father and child. NPR records diagnosis in secondary care (e.g., hospital). KUHR contains information about diagnosis and contact in primary care (e.g, GPs and outpatient specialists). NorPD recorded all medications dispensed outside of hospitals. MSIS collects test results of communicable diseases (e.g., Sars-Cov-2) and SYSVAK recorded vaccinations.

8.3 Study Period

The study period will start on 01/01/2008 as these dates correspond to the start date of the earliest roll-out of the HPV vaccination programme in these countries. The end of the study period will be the last available date of data collection for each contributing dataset.

8.4 Follow-up

For all analyses, follow-up time will start from index date, defined by the date when a participant receives their first dose before 15 years old. For unvaccinated, index date would be the same as their matched vaccinated pair, the date the matched vaccinated receives their first dose, and follow-up time will start then.

End of follow-up will be the end of a person's observation time (i.e. date of data extraction, death), or the date of outcome event, whichever comes first.

For main analyses (against unvaccinated), reception of first dose in the comparator cohort after the age of 15 (index date) will result in censoring of the comparator cohort. Additional analysis without this censoring will be performed to avoid differential attrition of high-risk participants in the comparator cohort. For secondary analyses the matched pairs will be censored if any of the participants receive a further dose after 15, and an additional analysis without this censoring will be performed.

8.5 Study Population with inclusion and exclusion criteria

The source population will comprise all females born on or after 1993 (aged 15 years old or less in 2008 – the earliest launch of the vaccine in all countries). This population will be restricted to those in observation in the database when they turned 15, and in observation in the database when they turned 9.

Further restrictions will be made in a year per year basis for the whole study period, applied on the 1st of January of each year. For each year, participants need to be in observation on the 1st of January of that year, need to have 365d of prior observation available, and need to be aged between 9 and 15 years old.

People will then be matched up to 1:5 based on PS (nearest neighbour within a 0.2 calliper width), and exact-matched by calendar year (of index date), and geographic region or GP practice (where available).

8.6 Variables

Preliminary concept/code lists used for the identification of exposure/s and/or outcomes are included as Supplementary Documents in Appendix I. These will be refined during the study execution following the DARWIN EU® Phenotyping standard processes, which involve the review of code lists by clinical experts, and the review of phenotypes after their execution in the participating databases.

8.6.1. Exposure/s (where relevant)

HPV Vaccination

HPV vaccination exposure status will be defined by number of doses received (0, 1, 2, or 3 or more) before 15 years old. Vaccine exposure will be examined overall, as well as stratified by vaccine brand (Bivalent: Cervarix; Quadrivalent: Gardasil/Silgard; and 9-valent: Gardasil-9).

8.6.2. Outcome/s (where relevant)

Primary outcomes of interest:

1. Diagnosis of Invasive Cervical Cancer
2. CIN2+
3. Conization

The operational definition of the outcomes is presented in **Table 8.6**.

Preliminary code lists are available in Appendix I of this protocol for all outcomes.

Table 8.6. Operational Definitions of Outcomes

Outcome name	Details	Primary outcome?	Type of outcome	Wash out window	Care Settings	Code Type	Diagnosis Position	Applied to study populations:	Measurement characteristics/ validation	Source of algorithm
CIN2+	Clinical diagnosis of CIN2+	Yes	Binary, time-to-event	Anytime prior to ID	IP, OP	SNOMED	Any	All study population	n/a	n/a
Conization	Record of conization of the cervix or cold knife cone (CKC) or loop diathermy					SNOMED				
Invasive Cervical Cancer	Clinical diagnosis of invasive cervical cancer					SNOMED				

ID = Index date, IP = inpatient, OP = outpatient, n/a = not applicable

8.6.3. Other covariates, including confounders, effect modifiers and other variables (where relevant)

All these variables will be used for characterisation of study cohorts, matching (e.g. geography), stratification (eg. by age), and to minimise confounding through their inclusion as potential covariates in large-scale propensity scores.

Demographics
Age at index date will be calculated.

Geographic region
Where available (CPRD GOLD and SIDIAP) we will use broad geographic regions (England [split into 9 distinct regions]; Northern Ireland, Scotland, Wales), and finer regions defined by GP surgery to account for local vaccination rates, infection, and testing rates.

Cytology results
Cytology results indicating HPV status from smear test prior to index date will be accounted for.

Healthcare resource use
Prior number of visits to GPs or any other specialists as recorded in the year before index date will be used as a proxy for healthcare resource use.

Health conditions pre-index date
Individuals´ history of the comorbidities will be identified over three time periods prior to the index date, and will be used for summary characterisation and calculation of large-scale propensity scores:

1) 30 days prior to one day prior index date,
2) 365 days prior to one day prior index date,
3) all available days observed up to one day prior to index date.

A range of health conditions will be assessed using the time windows above.

Medications pre-index date
Pre-existing medication use will be identified using 2 time windows, defined as 365 days to one day prior to index date, and 30 days to 1 day prior to index date, and will be used to provide summary characterisation for patients and calculation of large-scale propensity scores.

HIV status pre-index date
Presence/absence of HIV/AIDS any time in history prior to index date will be recorded and included as covariate in analyses.

Previous Papanicolaou smear Testing
Number of Papanicolaou smear test (cytology tests) anytime in history will be used as a covariate.

Previous Vaccinations
Number of vaccinations (any vaccine) any time in history will be used as a covariate.

Operational definitions of covariates are provided in **Table 8.7.**

Table 8.7. Operational Definitions of Covariates

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position[2]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
Demographics	Age at index date (first vaccine or matched date for unvaccinated)	Numeric, binary	All history	OP	n/a	n/a	All study population	n/a	n/a
Geographic location	Location identifier of GP surgery, healthcare regions (England [split into 9 distinct regions]; Northern Ireland, Scotland, Wales)	Categorical	All history	OP	n/a	n/a	All study population	n/a	n/a
Cytology results pre-index date	HPV infection status from cytology Papanicolaou test pre-index date	Categorical	All history prior to index date	OP	SNOMED	n/a	All study population	n/a	n/a
Healthcare resource use	Prior number of visits to GPs recorded in the patient record will be used as a proxy for healthcare resource use.	Numeric	All history prior to index date	OP	SNOMED	n/a	All study population	n/a	n/a
Health conditions pre-index date	Conditions of interest prior to index date	Binary	[-365,-1], [-30 –1], All history prior to index date	OP	SNOMED	n/a	All study population	n/a	n/a
Medication pre-index date	Drug prescriptions prior to index date	Binary	[-365,-1], [-30 –1],	OP	RxNorm	N/A	All study	n/a	n/a

			All history prior to index date				population		
HIV status pre-index date	Diagnosis of HIV/AIDS anytime in history prior to index date will be used as a covariate.	Binary	All history prior to index date	OP	SNOMED	N/A	All study population	n/a	n/a
Previous Papanicolaou smear Testing	Number of Papanicolaou smear tests (cytology tests)	Numeric, binary	All history	OP	LOINC	N/A	All study population	n/a	n/a
Previous vaccinations	Number of vaccinations (any vaccine) any time in history will be used as a covariate.	Numeric	All history	OP	RxNorm	N/A	All study population	n/a	n/a

ID = Index date, OP = outpatient, n/a = not applicable

8.7 Study size

For each database, all individuals that satisfy the eligibility criteria for a study cohort will be included.

Assuming a vaccine effectiveness against cervical cancer of 88%, with 60% vaccination coverage (a mean ratio unvaccinated to vaccinated of 0.67), a 10-year cumulative incidence of 94/100,000 based on a previous study (13), and for a 95% CI we calculated sample size needed for different precision values: (30)

IRR	Lower limit of 95%CI	Upper limit of 95%CI	Relative precision (%)	Sample size total
0.12	0.11	0.13	9	884,672
0.12	0.10	0.14	20	201,492
0.12	0.09	0.16	33	80,930
0.12	0.08	0.18	50	40,740
0.12	0.07	0.21	71	23,055
0.12	0.06	0.24	100	13,940
0.12	0.05	0.29	140	8,738
0.12	0.04	0.36	200	5,550

Contributing data sources range from 40,000 to 80,000 people vaccinated against HPV, so we would expect a relative precision of 33-50.

8.8 Analysis

All analyses will be conducted separately for each database, and will be carried out in a federated manner, allowing analyses to be run locally without sharing patient-level data.

Before sharing the study package, test runs of the analytics will be performed on a subset of the data sources and quality control checks will be performed. After all the tests are passed (see section 11 Quality Control), the final package will be released in a version-controlled study repository for execution against all the participating data sources.

The data partners will locally execute the analytics against the OMOP-CDM in R Studio and review and approve the default aggregated results. They will then be made available to the Principal Investigators and study team in secure online repository (Data Transfer Zone). All results will be locked and timestamped for reproducibility and transparency.

Table 8.9. Description of Study Type and Type of analysis

STUDY TYPE	STUDY CLASSIFICATION	TYPE OF ANALYSIS
Vaccine Comparative Effectiveness Studies	Complex (C3)	New cohort design: - Large-scale characterisation of participants in the target and comparator cohorts - Large-scale propensity scores (LSPS) will be estimated - Incidence rate/s of each of the outcomes of interest in the target and comparator cohorts - Diagnostic/s: Covariate balance, Equipoise, Power - Incidence Rate Ratios or Hazard Ratio/s and 95% confidence intervals using Poisson or Cox models respectively

8.8.1 Patient privacy protection

Cell counts <5 will be suppressed to comply with the database's privacy protection regulations.

8.8.2 Descriptive statistics

For each analysis, summary descriptive analyses will be conducted including age, sex, key variables for matching and conditions and medication pre-index date for characterisation.

8.8.3 Comparative-Effectiveness Analyses

We will use a PS-matched cohort design, where target and comparator cohort participants will be matched up to 5:1 based on propensity scores, and exact-matched on year of birth, calendar year of vaccination or index date, and geographic region (broad by region and granular by GP identifier, where available).

Large-scale propensity scores will be estimated using lasso regression to estimate the probability of being in the target cohorts potentially including any of the covariates mentioned above (SECTION 8.6.3). The resulting equations will be manually inspected by two clinical epidemiologists to identify any strong instrumental variables.

The following steps will then be followed for the curation of the necessary cohorts:

1. First, all subjects in the target cohort will be exact-matched by year of birth, calendar year and geographic region to all potential matches not belonging in the target cohort
2. Second, a PS will be estimated as follows:
    a. For vaccinated vs unvaccinated (main objectives): a PS will be calculated at the beginning of the calendar year for both the target and comparator cohorts. Up to 5 matches will be found in the target cohort for each participant in the comparator cohort using PS matching with nearest neighbour matching with a calliper width of 0.2. Matches will be sampled from the

pool of target cohort participants identified as potential matches in the first step. Then, the index date of the target cohort participant (or the mean time point if more than one) will be

b. applied to all the identified comparator cohort matches, and large-scale PS will be re-calculated on this index date.
c. For secondary objectives involving dose schedule: PS will be calculated on the index date for target and comparator cohorts, defined by the index vaccination, and PS matching will proceed using nearest neighbour matching with the same calliper as above (width <0.2) and variable 5:5 matching.

The following matched cohorts may (depending on sample size) be compared:

Main comparisons (Primary objectives):

Vaccinated vs unvaccinated per brand:

- Vaccinated with Gardasil/Silgard (target) (1 or more dose) vs unvaccinated (comparator)
- Vaccinated with Cervarix (target) (1 or more dose) vs unvaccinated (comparator)
- Vaccinated with Gardasil-9 (target) (1 or more dose) vs unvaccinated (comparator)

Secondary comparisons (Secondary objectives):

Vaccinated (target) (1 or more dose) (any brand) vs unvaccinated (comparator) overall.

Dose comparisons:

- Vaccinated with 2 or more doses (target) vs 1 dose (comparator) of the same brand.
- Vaccinated with 3 or more doses (target) vs 2 doses (comparator) of the same brand.

In all matched cohorts, people will be followed up from their index date until the earliest of:

- End of their observation (i.e. date of data extraction, death)
- Follow-up will be censored for the matched pair if one counterpart is censored
- Participants in the comparator cohort will be censored when and if they fulfil the entry criteria for the target cohort

8.8.4 Vaccine effectiveness analyses and study diagnostics

Incidence rates and incidence rate ratios (IRR) will be calculated for the matched cohorts and outcomes at 5, 10 and 15 years after vaccination using Poisson regression. Cox proportional hazard models will be used to calculate hazard ratios (HR) for the outcomes of Invasive Cervical Cancer, CIN2+ and conization.

Two study diagnostics will be used to minimise the risk of reporting biased results. First, any analyses with evidence of observed confounding after matching, as defined by ASMD>0.1 for any covariate will be inspected manually by two clinical epidemiologists. If any of these variables is deemed as a confounder, all subsequent analyses will stop, and vaccine effectiveness results will not be reported. Additionally, negative control outcomes will be used to identify residual (unobserved) confounding. A previously validated list of negative control outcomes will be utilised and refined to identify potential outcomes not associated with outcome risk, but sharing similar confounders as the association between HPV vaccination and outcomes. A preliminary list can be found in Appendix I Table 6.

Kaplan-Meier plots will be used to illustrate time-to-event analyses. Log-log plots will be visually inspected to identify scenarios with a violation of the proportional hazards assumption. If these plots show evidence of violation we will not report the results from the Cox regression, and will only report incidence rates and incidence rate ratios.

Table 8.10: Primary, secondary, and subgroup analysis specification

A. Primary analysis

Hypothesis:	HPV vaccine will decrease the risk of CIN2+, conization, and invasive cervical cancer
Exposure contrast:	HPV Vaccine(each brand) vs unvaccinated
Outcome:	CIN2+, conization, and invasive cervical cancer
Analytic software:	R
Model(s):	Incidence rates, incidence rate ratios, Cox proportional Hazards models, Kaplan-Meier Time-to-event.
Confounding adjustment method:	Among those participants in the target and comparator cohorts who met the inclusion criteria, target participants will be matched 5:1 to a comparator participant, based on year of birth, calendar year of vaccination, geographic region, and large-scale propensity scores using the nearest neighbour matching, with calliper width 0.2 standard deviations. Large-scale propensity scores will be estimated using lasso regression to estimate the probability of being in the target cohorts. Covariates will include all recorded features in the database, including socio-demographics, geographic location, healthcare resource use, comorbidity, medicine/s use, previous Papanicolaou testing, and previous vaccination/s. Among those, covariates with a prevalence below 0.5% in the study population will be omitted. Logistic regression with LASSO regularization will then be used for variable selection. The list of selected covariates will be manually screened by 2 epidemiologists/clinical domain experts to exclude potential instrumental variables.
Missing data methods:	None
Subgroup Analyses	***List all subgroups***

A. Secondary analysis 1

Hypothesis:	HPV vaccine will decrease the risk of CIN2+, conization, and invasive cervical cancer
Exposure contrast:	HPV Vaccine(overall) vs unvaccinated
Outcome:	Unchanged

Analytic software:	Unchanged
Model(s):	Unchanged
Confounding adjustment method:	Unchanged
Missing data methods:	Unchanged
Subgroup Analyses	***List all subgroups***

A. Secondary analysis 2

Hypothesis:	Higher number of doses will decrease more the risk of CIN2+, conization, and invasive cervical cancer
Exposure contrast:	HPV Vaccine with 1 vs 2 or more doses (secondary objective 1); HPV Vaccine with 2 vs 3 or more doses (secondary objective 2)
Outcome:	Unchanged
Analytic software:	Unchanged
Model(s):	Unchanged
Confounding adjustment method:	Unchanged
Missing data methods:	Unchanged
Subgroup Analyses	***List all subgroups***

Table 8.11 Sensitivity analyses – rationale, strengths and limitations

	What is being varied? How?	Why? (What do you expect to learn?)	Strengths of the sensitivity analysis compared to the primary	Limitations of the sensitivity analysis compared to the primary
	Study population and follow-up, not censoring unvaccinated subjects who are vaccinated after index date	To asses the potential impact of selection bias related to the censoring of subjects vaccinated after age 15	Does not exclude potentially higher risk subjects vaccinated in later life	Misclassification of exposure (vaccination status)
	In databases with incomplete vaccine data completeness, we will restrict the population to women in practices/region and birth cohort with more than 60% coverage.	We restrict the analyses where we believe vaccine data is close to complete, so we reduce the risk of exposure misclassification	Excludes subjects without information on vaccination.	Overestimation of exposure prevalence, selection bias.

Two sensitivity analyses will be performed. In the first one, we will not censor unvaccinated people once they get vaccinated after 15 years old, to mimic an intention to treat analysis.

A second sensitivity analysis will be performed in CPRD, where we suspect we have incomplete information on vaccination for some GP practices for some birth cohorts.

For this analysis, we will restrict the analysis to women vaccinated in GP practices that have more than a 60% coverage of HPV for their year of birth cohort. We decided to use this threshold by establishing a minimal coverage reported by the local public health agency (Public Health England/UKHSA/OHID) on their Fingertips database [cite]. This number was arrived to by getting the minimum coverage achieved by area, using the smallest area data available (Upper tier local authorities), before the COVID pandemic. This coverage was for Kensington and Chelsea in 2014-15, of 67.6% 95%CI(63.6% - 71.3%). We then decided to truncate the figure to 60%, to account for the possible variability introduced by us having smaller areas.

If we suspect similar potential misclassification of exposure in other databases we will define a similar threshold based on the available public health data on coverage by area.

8.9 Evidence synthesis

We will report analyses separately for each database and outcome. Additionally, we will pool the effect estimates across databases using random effect meta-analyses, I^2 coefficient for heterogeneity will be reported. Forest plots will be used to show results from meta-analyses.

9. DATA MANAGEMENT

9.1 Data Management

All databases have previously mapped their data to the OMOP common data model. This enables the use of standardised analytics and using DARWIN EU tools across the network since the structure of the data and the terminology system is harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org

The analytic code for this study will be written in R, and will use standardized analytics wherever possible. Each data partner will execute the study code against their database containing patient-level data, and then return the results (csv files) which will only contain aggregated data. The results from each of the contributing data sites will then be combined in tables and figures for the study report.

Methods for data collection, retrieval, and preparation. Statistical software(s) to be used in the study should be specified.

9.2 Data storage and protection

For this study, participants from various EU member states will process personal data from individuals which is collected in national/regional electronic health record databases. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All databases used in this study are already used for pharmaco-epidemiological research and have a well-developed mechanism to ensure that European and local regulations dealing with ethical use of the data and adequate privacy control are adhered to. In agreement with these regulations, rather than combining person level data and performing only a central analysis, local analyses will be run, which generate non-identifiable aggregate summary results.

The output files are stored in the DARWIN Remote Research Environment. These output files do not contain any data that allow identification of subjects included in the study. The RRE implements further security measures in order to ensure a high level of stored data protection to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/20161 in the various member states.

10. QUALITY CONTROL

General database quality control

A number of open-source quality control mechanisms for the OMOP CDM have been developed (see Chapter 15 of The Book of OHDSI http://book.ohdsi.org/DataQuality.html). In particular, data partners will have run the OHDSI Data Quality Dashboard tool (https://github.com/OHDSI/DataQualityDashboard). This tool provides numerous checks relating to the conformance, completeness and plausibility of the mapped data. Conformance focuses on checks that describe the compliance of the representation of data against internal or external formatting, relational, or computational definitions, completeness in the sense of data quality is solely focused on quantifying missingness, or the absence of data, while plausibility seeks to determine the believability or truthfulness of data values. Each of these categories has one or more subcategories and are evaluated in two contexts: validation and verification. Validation relates to how well data align with external

benchmarks with expectations derived from known true standards, while verification relates to how well data conform to local knowledge, metadata descriptions, and system assumptions.

Study specific quality control

Vaccine exposure status, clinical diagnoses and conization procedures will be identified from the data using code-lists that will be reviewed by clinicians. When defining conditions for outcomes of interest, i.e. CIN2+, cervical cancer, a systematic search of possible codes for inclusion will be conducted using CodelistGenerator R package (31). This software allows the user to define a search strategy and using this will then query the vocabulary tables of the OMOP common data model so as to find potentially relevant codes. Clinicians will review the resulting code lists to exclude irrelevant codes, such as for persisting disease or complications. In addition, vaccine coverage and cohort diagnostics will be run if needed to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error.

11. LIMITATIONS OF THE RESEARCH METHODS

The study will be informed by routinely collected health care data and so data quality issues, and adequate capture of the variables of interest in primary care data, must be considered. There may be incomplete exposures in some databases (e.g. IQVIA DA Germany, CPRD), and incomplete outcomes in all databases. This can lead to misclassification of exposure and outcome, and if informative, to bias. In SIDIAP, for example, sensitivity for cervix cancer using primary care ICD-10 codes has been reported to be very low (32). However, for this study, we plan to address that by adding hospital discharge data for all patients. The inclusion of the CIN 2+ and conisation outcomes could also help detect in situ carcinomas that may have not been codified as ICD-10 cancers.

Vaccinations programmes have varied the number of doses and brands in time, but for a certain point in time is possible that only one schedule of brand and doses was administered. That could result in the impossibility to perform most of the stratified or interaction analyses as number of doses and brands will be very colinear between them.

Using conisation as a proxy for CIN2+ may not be sensible in some settings, where clinicians may decide not to treat CIN2, especially in younger females. In addition, conization practice would vary by institution and health care system and will have impact on outcomes. We exclude women with previous HPV infection, however, they might not have a smear test result recorded on data, or it could be a false negative.

Although every effort will be made to minimise confounding, there may still be confounding due to unmeasured confounders, or effect modification. Main confounders that we will be unable to measure are those related to sexual activity levels. Given these cancers may take considerable time to develop, our data follow-up time may not capture all potential cancers that may develop, so the results will be limited to the timeframes where sufficient data is available.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse events/adverse reactions will not be collected or analysed as part of this evaluation. The nature of this non-interventional evaluation, through the use of secondary data, does not fulfil the criteria for reporting adverse events, according to module VI, VI.C.1.2.1.2 of the Good Pharmacovigilance Practices

(https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf)

Only in case of prospective data collection, there is a need to describe the procedures for the collection, management and reporting of individual cases of adverse events/adverse reactions.

13. GOVERNANCE BOARD ASPECTS

CPRD GOLD, NLHR, and SIDIAP will require ethical approvals from their local Institutional Review Boards to perform this study.

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Dissemination activities to be undertaken will include mainly, although not exclusively, the creation of a final report, scientific publications, and presentations at conferences.

15. OTHER ASPECTS

N/A

16. REFERENCES

1. Office of National Statistics. Cancer registration statistics, England: 2017 2017 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2017#the-three-most-common-cancers-vary-by-sex-and-age-group.
2. European Centre for Disease Prevention and Control. Factsheet about human papillomavirus 2018 [Available from: https://www.ecdc.europa.eu/en/human-papillomavirus/factsheet.
3. World Health Organization. Cervical Cancer Elimination Initiative 2020 [Available from: https://www.who.int/initiatives/cervical-cancer-elimination-initiative.
4. Garland SM, Hernandez-Avila M, Wheeler CM, Perez G, Harper DM, Leodolter S, et al. Quadrivalent vaccine against human papillomavirus to prevent anogenital diseases. N Engl J Med. 2007;356(19):1928-43.
5. Paavonen J, Naud P, Salmerón J, Wheeler CM, Chow SN, Apter D, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. Lancet. 2009;374(9686):301-14.
6. Joura EA, Giuliano AR, Iversen O-E, Bouchard C, Mao C, Mehlsen J, et al. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med. 2015;372(8):711-23.
7. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. Lancet Oncol. 2010;11(11):1048-56.
8. European Centre for Disease Prevention and Control. Guidance on HPV vaccination in EU countries: focus on boys, people living with HIV and 9-valent HPV vaccine introduction 2023 [Available from:

https://www.ecdc.europa.eu/en/publications-data/guidance-hpv-vaccination-eu-focus-boys-people-living-hiv-9vHPV-vaccine.
9. European Medicines Agency. Gardasil 2022 [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/gardasil.
10. European Medicines Agency. Cervarix 2023 [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/cervarix.
11. European Medicines Agency. Gardasil 9 2023 [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/gardasil-9.
12. Falcaro M, Castañon A, Ndlela B, Checchi M, Soldan K, Lopez-Bernal J, et al. The effects of the national HPV vaccination programme in England, UK, on cervical cancer and grade 3 cervical intraepithelial neoplasia incidence: a register-based observational study. Lancet. 2021;398(10316):2084-92.
13. Lei J, Ploner A, Elfström KM, Wang J, Roth A, Fang F, et al. HPV Vaccination and the Risk of Invasive Cervical Cancer. N Engl J Med. 2020;383(14):1340-8.
14. Mesher D, Panwar K, Thomas SL, Edmundson C, Choi YH, Beddows S, et al. The impact of the national HPV vaccination program in England using the bivalent HPV vaccine: surveillance of type-specific HPV in young females, 2010–2016. The Journal of infectious diseases. 2018;218(6):911-21.
15. Palmer T, Wallace L, Pollock KG, Cuschieri K, Robertson C, Kavanagh K, et al. Prevalence of cervical disease at age 20 after immunisation with bivalent HPV vaccine at age 12-13 in Scotland: retrospective population study. BMJ. 2019;365:l1161.
16. Drolet M, Bénard É, Pérez N, Brisson M, Ali H, Boily M-C, et al. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. The Lancet. 2019;394(10197):497-509.
17. Stier EA, Chigurupati NL, Fung L. Prophylactic HPV vaccination and anal cancer. Hum Vaccines Immunother. 2016;12(6):1348-51.
18. Tsentemeidou A, Fyrmpas G, Stavrakas M, Vlachtsis K, Sotiriou E, Poutoglidis A, et al. Human papillomavirus vaccine to end oropharyngeal cancer. A systematic review and meta-analysis. J Sex Transm Dis. 2021;48(9):700-7.
19. Abbas KM, van Zandvoort K, Brisson M, Jit M. Effects of updated demography, disability weights, and cervical cancer burden on estimates of human papillomavirus vaccination impact at the global, regional, and national levels: a PRIME modelling study. Lancet Glob Health. 2020;8(4):e536-e44.
20. Brotherton JML, Sundström K. More evidence suggesting that 1-dose human papillomavirus vaccination may be effective. Cancer. 2020;126(8):1602-4.
21. Kreimer AR, Sampson JN, Porras C, Schiller JT, Kemp T, Herrero R, et al. Evaluation of Durability of a Single Dose of the Bivalent HPV Vaccine: The CVT Trial. J Natl Cancer Inst. 2020;112(10):1038-46.
22. Sankaranarayanan R, Prabhu PR, Pawlita M, Gheit T, Bhatla N, Muwonge R, et al. Immunogenicity and HPV infection after one, two, and three doses of quadrivalent HPV vaccine in girls in India: a multicentre prospective cohort study. Lancet Oncol. 2016;17(1):67-77.
23. Baisley K, Kemp TJ, Kreimer AR, Basu P, Changalucha J, Hildesheim A, et al. Comparing one dose of HPV vaccine in girls aged 9–14 years in Tanzania (DoRIS) with one dose of HPV vaccine in historical cohorts: an immunobridging analysis of a randomised controlled trial. Lancet Glob Health. 2022;10(10):e1485-e93.
24. Man I, Georges D, de Carvalho TM, Ray Saraswati L, Bhandari P, Kataria I, et al. Evidence-based impact projections of single-dose human papillomavirus vaccination in India: a modelling study. Lancet Oncol. 2022;23(11):1419-29.
25. Prem K, Choi YH, Bénard É, Burger EA, Hadley L, Laprise J-F, et al. Global impact and cost-effectiveness of one-dose versus two-dose human papillomavirus vaccination schedules: a comparative modelling analysis. BMC medicine. 2023;21(1):313.

26. Brotherton JM, Budd A, Rompotis C, Bartlett N, Malloy MJ, Andersen RL, et al. Is one dose of human papillomavirus vaccine as effective as three?: A national cohort analysis. Papillomavirus Res. 2019;8:100177.
27. Rodriguez AM, Zeybek B, Vaughn M, Westra J, Kaul S, Montealegre JR, et al. Comparison of the long-term impact and clinical outcomes of fewer doses and standard doses of human papillomavirus vaccine in the United States: a database study. Cancer. 2020;126(8):1656-67.
28. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, Van Staa T, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.
29. García Gil MdM, Hermosilla E, Prieto-Alhambra D, Fina F, Rosell M, Ramos Blanes R, et al. Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). Informatics in Primary Care, 2011, vol 19, p 135-145. 2011.
30. Rothman KJ, Greenland S. Planning Study Size Based on Precision Rather Than Power. Epidemiology (Cambridge, Mass). 2018;29(5):599-603.
31. Burn E, Catala M. CodelistGenerator: Generate Code Lists for the OMOP Common Data Model 2023 [Available from: https://cran.r-project.org/web/packages/CodelistGenerator/index.html.
32. Recalde M, Manzano-Salgado CB, Diaz Y, Puente D, Garcia-Gil MDM, Marcos-Gragera R, et al. Validation Of Cancer Diagnoses In Electronic Health Records: Results From The Information System For Research In Primary Care (SIDIAP) In Northeast Spain. Clin Epidemiol. 2019;11:1015-24.

17. ANNEXES

Appendix I: List of Stand-Alone documents (e.g. lists with concept definitions (conditions & drugs), validation procedures, questionnaires etc.)

Appendix II: ENCePP checklist for study protocols

Appendix III: Additional Information

Appendix I

Table 1: Preliminary concepts for HPV Vaccination

conceptId	conceptName
44132618	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension [Gardasil]
44132684	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... Injectable Suspension [Gardasil] by Merck
44132751	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML Injectable Suspension [Cervarix]
44132765	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension [Gardasil] Box of 1 by Merck
44132790	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... Injectable Suspension [Gardasil]
44132776	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... Injectable Suspension
36262935	Human Papillomavirus 0.04 MG/ML Injectable Suspension [Cervarix]
43269462	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML [Cervarix]
46275945	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 6 vaccine Injection [Gardasil]
21105145	Human Papillomavirus Prefilled Syringe [Gardasil]
21105146	Human Papillomavirus Prefilled Syringe [Cervarix]
35408015	Human Papillomavirus 0.24 MG/ML [Gardasil]
35408900	Human Papillomavirus Injectable Suspension [Gardasil]
35409079	Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil] Box of 1
36265499	Human Papillomavirus 0.54 MG/ML Injectable Suspension [Gardasil 9]
36266792	Human Papillomavirus 0.54 MG/ML [Gardasil 9]

36267065	Human Papillomavirus Injectable Suspension [Gardasil 9]
36271855	Human Papillomavirus 0.04 MG/ML [Cervarix]
36281057	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil] by MSD
36281072	0.5 ML Human Papillomavirus 0.54 MG/ML Injectable Suspension [Gardasil 9]
36281079	0.5 ML Human Papillomavirus 0.54 MG/ML Injectable Suspension [Gardasil 9] by MSD
36281140	0.5 ML Human Papillomavirus 0.04 MG/ML Injectable Suspension [Cervarix] by GSK
36281160	0.5 ML Human Papillomavirus 0.04 MG/ML Injectable Suspension [Cervarix]
36405748	Human Papillomavirus Injectable Solution [Gardasil]
36407304	Human Papillomavirus Injectable Solution [Cervarix]
36421776	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 31 Vaccine / ... Prefilled Syringe [Gardasil]
36789910	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 6 Vaccine Injectable Suspension [Gardasil]
36882767	L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine Injection [Cervarix]
36896217	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension [Cervarix]
40150716	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 6 vaccine Prefilled Syringe [Gardasil]
40167170	L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine Injectable Suspension [Cervarix]
40167172	L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine Prefilled Syringe [Cervarix]
40753446	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 31 Vaccine / ... Injectable Suspension [Gardasil 9]
40832060	L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine Prefilled Syringe [Cervarix Grk]
41144528	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 6 Vaccine Prefilled Syringe [Silgard]
42873276	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML [Cervarix]
42873514	0.5 ML human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Prefilled Syringe [Cervarix]
42873516	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Injectable Suspension [Cervarix]

42902869	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Prefilled Syringe [Cervarix]
43161166	L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension [Cervarix] Box of 1
43208813	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension [Cervarix] Box of 1
43219812	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension [Cervarix] Box of 1 by Glaxosmithkline
43258573	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix] Box of 10
36789911	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 6 Vaccine Injectable Suspension
36275003	Human Papillomavirus 0.04 MG/ML Injectable Suspension
36276830	Human Papillomavirus 0.54 MG/ML
36281062	0.5 ML Human Papillomavirus 0.04 MG/ML Injectable Suspension
36281152	0.5 ML Human Papillomavirus 0.54 MG/ML Injectable Suspension
36404949	Human Papillomavirus Injectable Solution
36893469	L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine Injection
36896175	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension
40150715	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 6 vaccine Prefilled Syringe
40167169	L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine Injectable Suspension
40167171	L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine Prefilled Syringe
40213319	human papilloma virus vaccine, bivalent
40213320	human papilloma virus vaccine, quadrivalent
40213322	Human Papillomavirus 9-valent vaccine
40753447	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Vaccine / L1 protein, Human papillomavirus type 31 Vaccine / ... Injectable Suspension
40799175	Vaccine, Human Papillomavirus Type-6,11,16,18,31,33,45,52,58
42873274	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML
42873275	human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML
42873513	0.5 ML human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Prefilled Syringe

42873515	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Injectable Suspension
42899116	human papillomavirus
42903241	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML Prefill...
42903272	human papillomavirus type 16, L1 capsid protein (residues 2-471) vaccine 0.04 MG/ML / human papillomavirus type 18, L1 capsid protein (residues 2-472) vaccine 0.04 MG/ML Prefilled Syringe
42903288	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML Prefill...
43139030	L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension Box of 1
43186919	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.04 MG/ML Injectable Suspension Box of 1
43258571	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection
43258572	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection Box of 10
43258574	L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML
43269464	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix] Box of 20
43280334	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix]
43281226	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection [Gardasil]
43281227	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix]
43285824	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection [Gardasil]
43285825	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection [Gardasil] Box of 10
43292059	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix] Box of 10
43296583	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... [Gardasil]

43297441	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection [Gardasil] Box of 10
43297443	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection [Cervarix] Box of 20
44043707	L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Injectable Suspension [Cervarix]
44056702	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 / L1 protein, Human papillomavirus type 6 Injectable Suspension [Gardasil]
44059954	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension [Gardasil] Box of 1
44059955	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension [Gardasil]
44072749	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... Injectable Suspension [Gardasil]
44085643	L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML Injectable Suspension [Cervarix]
44118574	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... [Gardasil]
44132592	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML Injectable Suspension [Cervarix] by Glaxosmithkline
44132617	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension [Gardasil] Box of 1
43264004	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML
43269463	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection Box of 10
43270323	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection
43275029	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection Box of 20
43275030	L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection
43275893	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection
43291182	L1 protein, Human papillomavirus type 6 Vaccine 0.02 MG/ML
43297440	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML / ... Injection Box of 10

43297442	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection Box of 10
44025856	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 / L1 protein, Human papillomavirus type 6 Injectable Suspension
44032271	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension Box of 1
44039553	L1 protein, Human papillomavirus type 18 0.04 MG/ML
44040913	L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML [Cervarix]
44055725	L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 Injectable Suspension
44058246	L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML Injectable Suspension
44065595	L1 protein, Human papillomavirus type 18 0.08 MG/ML
44081430	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 / L1 protein, Human papillomavirus type 31 Vaccine / ... Injectable Suspension
44091253	L1 protein, Human papillomavirus type 6 0.04 MG/ML
44091254	L1 protein, Human papillomavirus type 6 0.06 MG/ML
44109803	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension
44109804	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.12 MG/ML / L1 protein, Human papillomavirus type 18 0.08 MG/ML / ... Injectable Suspension
44132513	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension
44132565	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.04 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML Injectable Suspension
44132566	0.5 ML L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / ... Injectable Suspension Box of 1
45892474	L1 protein, human papillomavirus type 31 vaccine
45892475	L1 protein, human papillomavirus type 33 vaccine
45892476	L1 protein, human papillomavirus type 45 vaccine
45892478	L1 protein, human papillomavirus type 58 vaccine
45892497	L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML
45892498	L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML
45892499	L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML
45892500	L1 protein, human papillomavirus type 33 vaccine 0.04 MG/ML
45892501	L1 protein, human papillomavirus type 45 vaccine 0.04 MG/ML

45892502	L1 protein, human papillomavirus type 52 vaccine 0.04 MG/ML
45892503	L1 protein, human papillomavirus type 58 vaccine 0.04 MG/ML
45892504	L1 protein, human papillomavirus type 6 vaccine 0.06 MG/ML
45892506	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML...
35412967	Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil]
35414597	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil] Box of 1
35414650	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil] Box of 1 by Sanofi
35414667	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension [Gardasil]
35753734	Human Papillomavirus Injectable Suspension [Cervarix]
36259657	Human Papillomavirus 0.54 MG/ML Injectable Suspension
36264100	Human Papillomavirus 0.04 MG/ML
43259418	0.5 ML L1 protein, Human papillomavirus type 16 Vaccine 0.02 MG/ML / L1 protein, Human papillomavirus type 18 Vaccine 0.02 MG/ML Injection Box of 20
43264003	L1 protein, Human papillomavirus type 11 Vaccine 0.04 MG/ML
45892477	L1 protein, human papillomavirus type 52 vaccine
45892508	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML / L1 p...
45892510	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML...
45892511	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 31 vaccine / L1 protein, human papillomavirus type 33 vaccine...
45892512	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML...
45892513	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 31 vaccine / L1 protein, human papillomavirus type 33 vaccine...
45892514	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML...
45892515	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML / L1 p...
45892516	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML / L1 p...

46275095	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML Injection
46275096	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML Injecti...
46275097	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML / L1 p...
46275098	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.12 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 31 vaccine 0.04 MG/ML / L1 p...
46275944	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 6 vaccine Injection
46275946	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 31 vaccine / L1 protein, human papillomavirus type 33 vaccine...
46275947	L1 protein, human papillomavirus type 11 vaccine / L1 protein, human papillomavirus type 16 vaccine / L1 protein, human papillomavirus type 18 vaccine / L1 protein, human papillomavirus type 31 vaccine / L1 protein, human papillomavirus type 33 vaccine...
44040914	L1 protein, Human papillomavirus type 11 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 16 Vaccine 0.08 MG/ML / L1 protein, Human papillomavirus type 18 0.04 MG/ML / L1 protein, Human papillomavirus type 6 0.04 MG/ML [Gardasil]
44069581	L1 protein, Human papillomavirus type 11 Vaccine / L1 protein, Human papillomavirus type 16 Vaccine / L1 protein, Human papillomavirus type 18 / L1 protein, Human papillomavirus type 31 Vaccine / ... Injectable Suspension [Gardasil]
529076	L1 protein, human papillomavirus type 11 vaccine
529077	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML
529078	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML...
529079	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML...
529080	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML...
529081	0.5 ML L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML...
529112	L1 protein, human papillomavirus type 16 vaccine

529113	L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML
529114	L1 protein, human papillomavirus type 18 vaccine
529115	L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML
529116	L1 protein, human papillomavirus type 6 vaccine
529117	L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML
21075767	Human Papillomavirus Prefilled Syringe
35407381	Human Papillomavirus 0.24 MG/ML Injectable Suspension Box of 1
35407382	Human Papillomavirus 0.24 MG/ML Injectable Suspension
35411044	Human Papillomavirus 0.24 MG/ML
35412768	Human Papillomavirus Injectable Suspension
35414605	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension Box of 1
35414634	0.5 ML Human Papillomavirus 0.24 MG/ML Injectable Suspension
40213321	HPV, unspecified formulation
19093987	L1 protein, human papillomavirus type 11 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 16 vaccine 0.08 MG/ML / L1 protein, human papillomavirus type 18 vaccine 0.04 MG/ML / L1 protein, human papillomavirus type 6 vaccine 0.04 MG/ML [Gardasil]

Table 2: Preliminary concepts for cytology

conceptId	conceptName
4056397	Cervical smear - atrophic changes
4056399	Cervical smear - no inflammation
4056400	Cervical smear - trichomonas
4056401	Cervical smear - gardnerella
4056548	Cervical smear - candida
4056550	Cervical smear - herpes
4056551	Cervical smear - koilocytosis
4058746	Cervical smear - actinomyces
4155376	Cervical smear - negative
44810558	Cervical smear - human papillomavirus negative
432447	Atypical glandular cells on cervical Papanicolaou smear
433033	Abnormal cervical Papanicolaou smear
434170	Atypical squamous cells of undetermined significance on cervical Papanicolaou smear
4209489	Atypical endocervical cells on cervical Papanicolaou smear
40480043	Abnormal cervical Papanicolaou smear with positive human papillomavirus deoxyribonucleic acid test
45763589	High grade squamous intraepithelial lesion on cervical Papanicolaou smear
45773176	Low grade squamous intraepithelial lesion on cervical Papanicolaou smear
436816	Vaginal vault smear abnormal
435658	Atypical squamous cells on vaginal Papanicolaou smear cannot exclude high grade squamous intraepithelial lesion
436245	Atypical glandular cells on vaginal Papanicolaou smear

438559	Abnormal vaginal Papanicolaou smear
441691	Atypical squamous cells of undetermined significance on vaginal Papanicolaou smear
4208032	Atypical endocervical cells on vaginal Papanicolaou smear
45757384	High grade squamous intraepithelial lesion on vaginal Papanicolaou smear
45757386	Cytological evidence of malignancy on vaginal Papanicolaou smear
45772099	Low grade squamous intraepithelial lesion on vaginal Papanicolaou smear
434165	Abnormal cervical smear
441138	Cervical smear result
4056398	Cervical smear - inflammatory change
4056545	Cannot exclude glandular neoplasia on cervical smear
4058745	Cervical smear - severe inflammation
4059358	Severe dyskaryosis on cervical smear cannot exclude invasive carcinoma
4077659	Viral changes on cervical smear
4095039	Dyskaryosis on cervical smear
4147886	Cervical smear - mild inflammation
4148734	Cervical smear - viral inflammation unspecified
4149625	Cervical smear - moderate inflammation
4151556	Cervical smear - severe dyskaryosis
4152695	Cervical smear - wart virus
4155377	Cervical smear - mild dyskaryosis
4206397	Dysplasia on cervical smear
44810423	Cervical smear - high grade dyskaryosis (moderate)
44810424	Cervical smear - high grade dyskaryosis (severe)
44810425	Cervical smear - high grade dyskaryosis with features of invasive squamous carcinoma
44810429	Cervical smear - features of endocervical type glandular neoplasia
44810430	Cervical smear - features of non-cervical type glandular neoplasia
44810559	Cervical smear - human papillomavirus positive
44810846	Cervical smear - low grade dyskaryosis
45770837	Cytological evidence of malignancy on cervical Papanicolaou smear
4152694	Cervical smear - moderate dyskaryosis
437104	Abnormal smear, noted, recall delete
4056409	Vaginal vault smear negative
4056410	Vaginal vault smear - atrophic
4056547	Cervical smear - endocervical cells present
44791000	Cervical smear repeat at 60 months
44791495	Cervical smear repeat at 36 months
44805864	Cervical smear repeat at 48 months
44810155	Cervical smear repeat at 24 months
4064368	Ca cervix screening normal
44790031	Smear abnormal - patient told
4056546	Cervical smear - borderline changes
45763873	Cervical smear - borderline change in squamous cells
45763874	Cervical smear - borderline change in endocervical cells
4077517	Nuclear abnormality on smear
4164508	Cervical smear: action
4059362	Cervical smear: repeat 3 months

4059361	Cervical smear: repeat 1 month
4059360	Cervical smear - action needed
4058750	Cervical smear - cervical biopsy needed
4056555	Cervical smear: uterine curettage needed
4056554	Cervical smear: repeat 9 months
4056552	Cervical smear: repeat 4 months
4056402	Cervical smear: repeat after treatment
4058749	Cervical smear: colposcopy needed
4056553	Cervical smear: repeat 6 months
4058748	Cervical smear: repeat 12 months
2007348	Microscopic examination of specimen from female genital tract, cell block and Papanicolaou smear
4088861	Abnormal smear - 3rd recall
4085965	Smear inflamed - recall delete
4085967	Smear infected - recall delete
4215530	Cervical smear every 12 months for life
44802497	Annual cervical smear required
4235948	Sampling of cervix for Papanicolaou smear
44791066	Smear normal - patient notified
44806380	Cervical smear pus cells present
4260744	Cervical smear transformation zone cells present
4085960	Cervical smear - 1st recall
4085961	Abnormal smear - 1st recall
4085964	Smear inflamed - 3rd recall
4089332	Abnormal smear - 2nd recall
4089333	Smear inflamed - 2nd recall
4202174	Cervical smear status
35624550	Sampling of cervix for Papanicolaou smear done
44788662	Cervical cytology claim
44806381	Cervical smear red blood cells present
4062484	Screening for malignant neoplasm of cervix
435651	Atypical squamous cells on cervical Papanicolaou smear cannot exclude high grade squamous intraepithelial lesion
4085963	Smear inflamed - 1st recall
45769920	Cervical smear report received

Table 3: Preliminary concepts for cervical intraepithelial neoplasia (CIN2+)

conceptId	conceptName
4244994	Surgical endocervical margin involved by intraepithelial neoplasia
42689509	CGIN - cervical glandular intraepithelial neoplasia
192676	Cervical intraepithelial neoplasia grade 1
196165	Cervical intraepithelial neoplasia grade 2
198572	Cervical intraepithelial neoplasia

4098948	Cervical intraepithelial neoplasia grade III with severe dysplasia
4245731	Surgical exocervical margin involved by intraepithelial neoplasia
35626025	Low grade cervical glandular intraepithelial neoplasia
35626026	High grade cervical glandular intraepithelial neoplasia
4247367	Squamous cell carcinoma in situ of uterine cervix
443571	Cervicovaginal cytology: High grade squamous intraepithelial lesion or carcinoma
443570	Cervicovaginal cytology: Low grade squamous intraepithelial lesion
4116082	Carcinoma of cervix stage 0
4216500	Anaplasia of cervix
4243120	Carcinoma in situ of endocervix
4243874	Carcinoma in situ of exocervix
194611	Carcinoma in situ of uterine cervix
40486538	Endocervical adenocarcinoma in situ
4213274	Grade of intraepithelial neoplasia involving endocervical margin
4218365	Grade of intraepithelial neoplasia involving exocervical margin
40489895	Grade III squamous intraepithelial neoplasia with microinvasive squamous cell carcinoma
4331440	High-grade squamous intraepithelial lesion
4069557	Squamous intraepithelial neoplasia, high grade
4013220	Low-grade squamous intraepithelial lesion
4161592	High-grade squamous intraepithelial lesion with features suspicious for invasion

Table 4: Preliminary concepts for conization

conceptId	conceptName
4238693	Cone biopsy of cervix with dilation, curettage and with repair
4121773	Cervical electroconization
4194122	Excision of cervix by cryoconization
4273387	Excision of cervix by electroconization
4181912	Cone biopsy of cervix
4074137	Loop diathermy cone biopsy of cervix uteri
4074291	Laser cone biopsy of cervix uteri
4213044	Cold knife cone biopsy of cervix
4244994	Surgical endocervical margin involved by intraepithelial neoplasia
4213273	Surgical endocervical margin uninvolved by intraepithelial neoplasia
4245731	Surgical exocervical margin involved by intraepithelial neoplasia
4288548	Surgical exocervical margin uninvolved by intraepithelial neoplasia
4244993	Surgical endocervical margin involved by malignant neoplasm
4213272	Surgical endocervical margin uninvolved by malignant neoplasm
4244996	Surgical exocervical margin involved by malignant neoplasm, focal

4244995	Surgical exocervical margin uninvolved by malignant neoplasm
4245730	Surgical exocervical margin involved by malignant neoplasm
4262594	Surgical endocervical margin involved by malignant neoplasm, focal
4262596	Surgical exocervical margin involved by malignant neoplasm, diffuse
4288546	Surgical endocervical margin involved by malignant neoplasm, diffuse

Table 5: Preliminary concepts for invasive cervical cancer

conceptId	conceptName
4107625	International Federation of Gynecology and Obstetrics cervical cancer stage Ib
36676408	Malignant germ cell neoplasm of cervix uteri
436358	Primary malignant neoplasm of exocervix
441805	Primary malignant neoplasm of endocervix
4003682	Overlapping malignant neoplasm of uterine cervix
4047648	Neoplasm of uterine cervix
4091766	Secondary malignant neoplasm of cervix uteri
4092515	Malignant neoplasm, overlapping lesion of cervix uteri
4131009	Neoplasm of endocervix
4131759	Neoplasm of exocervix
4157449	Malignant neoplasm of endocervix
4162876	Malignant neoplasm of exocervix
4247949	Secondary malignant neoplasm of endocervix
4311881	Secondary malignant neoplasm of exocervix
4313498	Neoplasm of uncertain behavior of exocervix
4314638	Neoplasm of uncertain behavior of uterine cervix
4314644	Neoplasm of uncertain behavior of endocervix
443708	Cervical Papanicolaou smear positive for malignant neoplasm
4092514	Malignant neoplasm of endocervical gland
4095156	Malignant neoplasm of endocervical canal
4244993	Surgical endocervical margin involved by malignant neoplasm
4095746	Malignant neoplasm of cervical stump
4244996	Surgical exocervical margin involved by malignant neoplasm, focal
4245730	Surgical exocervical margin involved by malignant neoplasm
4262594	Surgical endocervical margin involved by malignant neoplasm, focal
4262596	Surgical exocervical margin involved by malignant neoplasm, diffuse
4288546	Surgical endocervical margin involved by malignant neoplasm, diffuse
198984	Malignant tumor of cervix
4310575	Carcinoma of uterine cervix, invasive
4195400	FIGO CC stage II
4199618	FIGO CC stage III
4338821	FIGO CC stage Ib occ
4304699	FIGO CC stage IV

45769633	FIGO CC stage IA1
45769645	FIGO CC stage IB1
45769669	FIGO CC stage IB2
45769680	FIGO CC stage IIA1
45769682	FIGO CC stage IIA2
45769685	FIGO CC stage IIB
45769690	FIGO CC stage IIIA
45771056	FIGO CC stage IIIB
45771057	FIGO CC stage IVA
45771058	FIGO CC stage IVB
4048735	FIGO CC stage I
4247367	Squamous cell carcinoma in situ of uterine cervix
4112314	Carcinoma of cervix
4116081	Adenosquamous carcinoma of cervix
4162134	Carcinoma of endocervix
35614875	Glassy cell carcinoma of cervix uteri
35621807	Adenoid basal carcinoma of cervix uteri
35621808	Adenoid cystic carcinoma of cervix uteri
35622692	Adenosarcoma of cervix uteri
35622770	Carcinosarcoma of cervix uteri
4243120	Carcinoma in situ of endocervix
4047649	Endocervical adenocarcinoma
37016238	Primary adenocarcinoma of cervix uteri
37016242	Primary adenocarcinoma of endocervix
4110872	Adenocarcinoma of cervix
4112875	Adenocarcinoma in situ of cervix
4112876	Fibroid polyp of cervix
4116080	Squamous cell carcinoma of cervix
4116083	Cervical fibroid
4116358	Adenomatous polyp of cervix
4201484	Local recurrence of malignant tumor of cervix
4243874	Carcinoma in situ of exocervix
4319273	Tumor of cervix affecting pregnancy
37116830	Invasive carcinoma of uterine cervix co-occurrent with human immunodeficiency virus infection
4174040	International Federation of Gynecology and Obstetrics cervical cancer (FIGO CC) stage Ia
196359	Primary malignant neoplasm of uterine cervix
4095158	Malignant neoplasm of squamocolumnar junction of cervix
45769637	FIGO CC stage IA2
194611	Carcinoma in situ of uterine cervix

4157450	Carcinoma of exocervix
40486538	Endocervical adenocarcinoma in situ

Table 6: Preliminary negative control outcomes

ConceptId	ConceptName
75860	Constipation
197304	Ulcer of lower extremity
42709838	Cellulitis of lower limb
436659	Iron deficiency anemia
4155902	Wax in ear canal
138825	Actinic keratosis
375545	Cataract
377889	Hearing loss
140673	Hypothyroidism
4026112	Rectal hemorrhage
4169905	Foot pain
197672	Urinary incontinence
4317977	Bilateral cataracts
436070	Vitamin D deficiency
4112752	Basal cell carcinoma of skin
195562	Hemorrhoids
141932	Senile hyperkeratosis
4217260	Intraocular pressure left eye
4038030	Hearing difficulty
437541	Glaucoma
380731	Otitis externa
4195039	Osteopenia
4036620	Dry eyes
378425	Blepharitis
74719	Ulcer of foot
4111921	Squamous cell carcinoma of skin
138384	Acquired hypothyroidism
374028	Age related macular degeneration
44783954	Acid reflux
4155040	Laceration of lower leg
4288544	Inguinal hernia
46287159	Traumatic wound
196456	Gallstone
135333	Pressure ulcer
4285898	Polyp of colon
374375	Impacted cerumen

443419	Laceration - injury
4053604	Open wound of lower leg
376707	Acute conjunctivitis
4016155	Prostatism

APPENDIX II: ENCEPP CHECKLIST FOR STUDY PROTOCOLS

ENCePP Checklist for Study Protocols (Revision 4)

Study title:
DARWIN EU® – Effectiveness of Human Papillomavirus Vaccines (HPV) to prevent cervical cancer in women.

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

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1. Does the protocol specify timelines for 1.1.1 Start of data collection[1] 1.1.2 End of data collection[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.	X			5. Milestones, 8.2 Data Sources

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) 2.1.2 The objective(s) of the study? 2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised) 2.1.4 Which hypothesis(-es) is (are) to be tested? 2.1.5 If applicable, that there is no *a priori* hypothesis?	X			7. Research question and objectives 8. Research methods

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)	X			8.1 Study type and Study Design

3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	X			8.2 Study Setting and Data Sources
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	X			8.8 Analysis
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))	X			8.8 Analysis
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)			X	

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	X			8.5 Study Population
4.2 Is the planned study population defined in terms of: 4.2.1 Study time period 4.2.2 Age and sex 4.2.3 Country of origin 4.2.4 Disease/indication 4.2.5 Duration of follow-up	X			8.3 Study Period 8.6.3. Other covariates 8.2 Study Setting and Data Sources 8.6.1. Exposures 8.4 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)	X			8.5 Study Population with inclusion and exclusion criteria

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)	X			8.6.1. Exposures
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)			X	
5.3 Is exposure categorised according to time windows?	X			8.6.1. Exposures
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	X			8.6.1. Exposures
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?			X	
5.6 Is (are) (an) appropriate comparator(s) identified?	X			8.8 Analysis

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?	X			8.6.2. Outcomes
6.2 Does the protocol describe how the outcomes are defined and measured?	X			8.6.2. Appendix I
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)			X	
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)			X	

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)	X			8.8 Analysis
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)			X	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	X			8.8 Analysis

Comments:

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)	X			8.8 Analysis

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)	X			8.6.1. Exposures
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	X			8.6.2. Outcomes
9.1.3 Covariates and other characteristics?	X			8.6.3. Other covariates
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	X			8.2 Study Setting and Data Sources
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	X			8.2 Study Setting and Data Sources

9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	X			8.2 Study Setting and Data Sources
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	X			8.6.1. Exposures
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	X			8.6.2. Outcomes, Appendix I
9.3.3 Covariates and other characteristics?	X			8.6.3. Other covariates
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)			X	

Comments:

Section 10: Analysis plan	Yes	No	N/A	Section Number
10.1 Are the statistical methods and the reason for their choice described?	X			8.8 Analysis
10.2 Is study size and/or statistical precision estimated?			X	
10.3 Are descriptive analyses included?	X			8.8.2 Descriptive statistics
10.4 Are stratified analyses included?	X			8.8 Analysis
10.5 Does the plan describe methods for analytic control of confounding?	X			8.8 Analysis
10.6 Does the plan describe methods for analytic control of outcome misclassification?	X			8.8 Analysis
10.7 Does the plan describe methods for handling missing data?			X	
10.8 Are relevant sensitivity analyses described?	X			8.8 Analysis

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)	X			9. Data management
11.2 Are methods of quality assurance described?	X			10. Quality Control
11.3 Is there a system in place for independent review of study results?			X	

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?	X			11. Limitations of the research methods

(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)	X			Table 8.2. Description of the selected Data Sources.

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?	X			13. Governance board aspects
13.2 Has any outcome of an ethical review procedure been addressed?			X	
13.3 Have data protection requirements been described?	X			9.2 Data storage and protection

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?	X			4. Amendments and updates

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)?	X			14. Plans for disseminating and communicating study results
15.2 Are plans described for disseminating study results externally, including publication?	X			14. Plans for disseminating and communicating study results

Comments: